CHAPTER III

RESEARCH METHOD

This chapter discusses the research method which is used in this study. It includes research method, research design, population, sample and sampling technique, technique for collecting data, technique for analyzing data, and statistical hypothesis.

A. Research Design

The research design that was used by researcher in the study was experimental research. As Ary, *et.al* (2010) have already stated, experimental research is a study of the effect of the systematic manipulation of one variable on another. It is usually regarded as the research method to test the hypothesis started with a question about the relationship between two variables or more. It is in line with what Mackey and Gass (2005: 2) have highlighted that "quantitative research generally starts with a hypothesis and is followed by the quantification of data and some sort of numerical analysis is carried out". Meanwhile, according to Nunan (1992:25), experimental research is carried out in order to explore the strength of relationship between variables.

As Ary, *et.al* (2010) described, sometimes researcher or experimenter cannot randomly select subjects for experimental treatment. Therefore, the experimenter must use already assembled groups, such as classes. In this case, then the research was called quasi-experimental.

In this current study, the researcher only took the intact two classes assigned as experimental group and control group. Experimental group was a group which experienced the treatment that was indirect corrective feedback, while the control group was a group which received direct corrective feedback as a conventional method in giving correction to students' error.

B. Variables of the Study

As aforementioned, quasi-experimental research was selected since the study tried to reveal the effect of indirect corrective feedback on students' writing accuracy as related to the different level of grammatical sensitivity in order to provide a better explanation for the outcome. The level of grammatical sensitivity was seen as a variable that influenced the method used to give feedback to the students' writing. Hence, this current study involved three kinds of variables; written corrective feedback, students' writing accuracy, and students' different level of grammatical sensitivity.

The first variable was independent variable. Independent variable is the variable which the researcher expects to influence the other variable (Nunan, 1992:25). Accordingly, the independent variable of the study was indirect corrective feedback. Meanwhile, the second variable was called as dependent variable that was the variable upon which the independent variable was acting (Nunan, 1992:25). Thus, the dependent variable in this study was the students' writing accuracy. Further, the third variable in this study was grammatical sensitivity level. It was called as moderator variable which was defined as the

variable that might give result or could modify the interaction between an independent variable and other variables (Mackey & Gass, 2005: 103).

A moderator variable was a type of independent variable that may not be the main focus of the study, but might modify the relationship between the independent variable and the dependent variable. The moderator variable was embedded into this research by assigning subjects to groups based on such existing variable; in this study was the students' level of grammatical sensitivity. The independent variables of either type are knows as factors. Therefore, a quasi-factorial design is used. Figure 3.1 below shows the variables of the study.

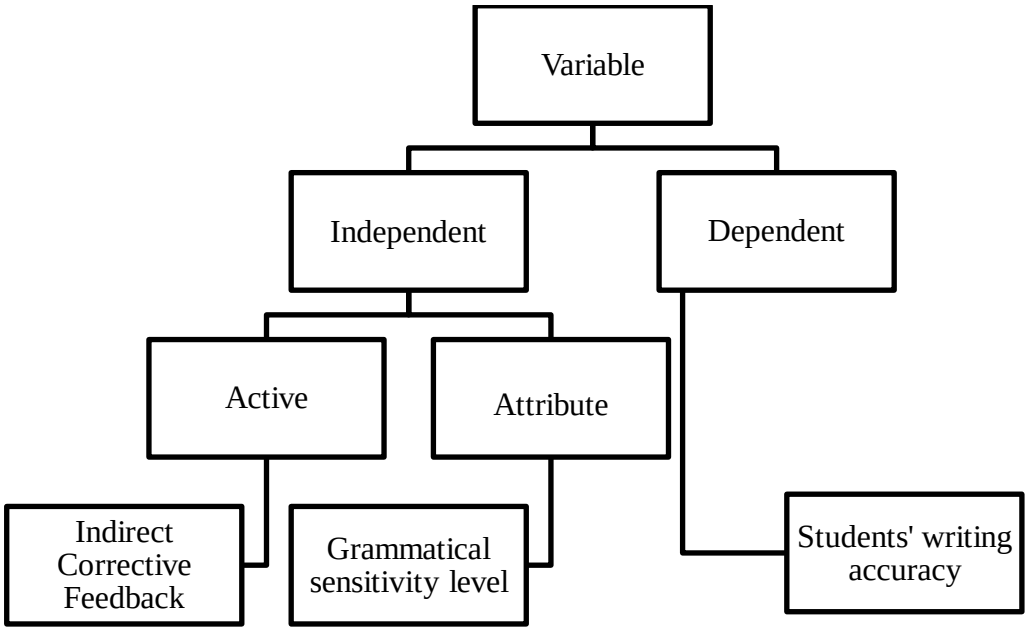

Figure 3.1 Variables of the Research

According to the number of variables involved in the study, this research employed simple factorial design 2x2, which was further read as factorial design 2 by 2. 2x2 factorial design was illustrated in Table 3.1 below:

Table 3.1. Simple Factorial Design 2 X 2

Main Effect / Simple Effect	Indirect Corrective Feedback (Group A1)	Direct Corrective Feedback (Group A2)
High Grammatical Sensitivity (B1)	Students having high grammatical sensitivity taught using Indirect Corrective Feedback (Group A1B1)	Students having high grammatical sensitivity taught using Direct Corrective Feedback (Group A2B2)
Low Grammatical Sensitivity (B2)	Students having low grammatical sensitivity taught using Indirect Corrective Feedback (Group A1B2)	Students having low grammatical sensitivity taught using Direct Corrective Feedback (Group A2B2)

The table 3.1 above shows us that:

1) By comparing the observations under treatment variable, (A_1) and the observation (A_2), it was possible to see the difference of the effectiveness of the methods to teach writing skill for the students viewed from their grammatical sensitivity levels.
2) By comparing high grammatical sensitivity level (B_1), it was possible to see the difference of the effectiveness of the method to teach writing skill for the students with high grammatical sensitivity.
3) By comparing low grammatical sensitivity level (B_2), it was possible to see the difference of the effectiveness of the method to teach writing skill for the students with low grammatical sensitivity.

4) By comparing the individual effects, group A_1B_1 versus A_2B_2, group A_1B_2 versus A_2B_2, it was possible to identify the interaction of indirect corrective feedback and students' grammatical sensitivity level in teaching writing.

The issue of extraneous variables which might affect the result of the experimental research somehow needed to be considered. Campbell and Stanley (1963) as cited by Borg and Gall (1983) stated that there were some extraneous variables in experimental research that should be controlled unless they would be a threat for the research. Controlling the extraneous variables would ascertain that the observed change was merely due to the treatment rather than due to the extraneous variables. The extraneous variables in question were; history, maturation, testing, measuring instruments, statistical regression, differential selection of subjects, treatment, experimenter effect and diffusion.

History deals with the events or conditions that may occur between the beginning of the test and the posttest measurement and may produce changes in the dependent variables. To avoid this threat, the experimental and control groups were chosen from the same level and experience. In this study, the students of eleventh grade of SMA Mamba'us Sholihin were chosen as the sample of the study. Both group, experimental and control group, were taken from the same grade level to avoid history treat.

Second, to avoid maturation, researcher treated both experimental and control groups at the same period. The researcher conducted treatment at the same number of meeting. Treatment for control group was done for six meetings. The number of meeting to give treatment for experimental were also consisting of

six meetings. The procedures were also the same; the difference was only on the technique given during applying corrective feedback. Experimental group were experiencing indirect corrective feedback, while the other one experienced direct corrective feedback.

Third, to minimize the threat in the form of testing, the researcher administered post-test which was using different but equal topic compared to the pre-test. The topic being given between pretest and posttest were adjusted to the level of the students and were cultivated to have the equal topics. The topics for pretest were: English as compulsory subject at school, social media for students, and the importance of breakfast. Meanwhile, the topics for the posttest were: the use of laptop in the classroom, the prohibition of smartphone in a school, and the use of plastic bag.

Fourth, the possibility of the lack of validity and reliability had been avoided by conducting try out for writing pre-test, writing post-test, as well as grammatical sensitivity test. Expert validation was also employed to validate the instruments to check both face validity and construct validity. Meanwhile, the content validity was based on the agreement of the topic being tested to the syllabus used by the sample of the study. The questionnaire was also distributed to the try out-test takers to make sure the clarity of the instructions. Meanwhile, inter-rater reliability was used to ascertain the reliability of the instruments.

Fifth, statistical regression was avoided by selecting the samples of the study which were having equal characteristics. The population where the study was conducted, they were already grouped into certain classes and there were no

any special classification to group them. Therefore, it was assumed that the groups were equal.

Sixth, differential of subject selection threat was avoided by comparing the result of pre-test to make sure that both experimental and control groups were having similar or equal character. Next, mortality threat was avoided by conducting the treatment within six meetings so that there were no students who dropped out from the experiment. Then, experimenter effect was avoided by teaching the material which was based on the syllabus used at the school so that it was appropriate with the level of the students. Last, diffusion threat was avoided by not telling the subjects that they were grouped into either the experimental group or control group.

C. Population, Sampling Technique, and Sample

1. Population of the Study

A population is a set of items or individuals each of which can be assigned values of one or more characteristics (Jobson, 1999: 12). Meanwhile, according to Ary et.,al (2010:148) population is defined as all members of any well-defined class of people, events, or objects.

The population of this study was the students of the eleventh grade at SMA Mamba'us Sholihin Blitar in the academic year of 2018/2019. The total numbers of the population were 144 students divided into 4 classes. All of the classrooms were having the similar characteristic especially when being seen from the

students' competency because there was no any placement test considered to assign them into certain classroom.

2. Sampling Technique of the Study

Sampling technique was a step to select the sample. Since the researcher was not able to draw a random sample of students for a study, the researcher used nonprobability sampling in order to select the members of the sample. The sampling technique in this study was cluster random sampling. It was the selection of groups or cluster, of objects rather than individuals (Ary et.,al, 2010).

There were four classes of eleventh grade students SMA Mamba'us Sholihin Blitar to choose as one experimental group and one control group. All of those seven classes were having the same chance to be chosen as sample. The technique of taking the sample was by using lottery. The steps to take sample were as follows:

a. Writing down the name of each class on small piece of paper.
b. Rolling the paper and put it into a straw, then put into a can
c. Shaking the can to mix the piece of paper.
d. Dropping two pieces of paper.

The lottery containing the name of the two classes chosen then were mixed, shake, and dropped again to assign which class belongs to an experimental group and which class belongs to a control group.

Figure 3.2 below shows the two groups from intact classes being selected as the sample of the study by employing the cluster sampling technique:

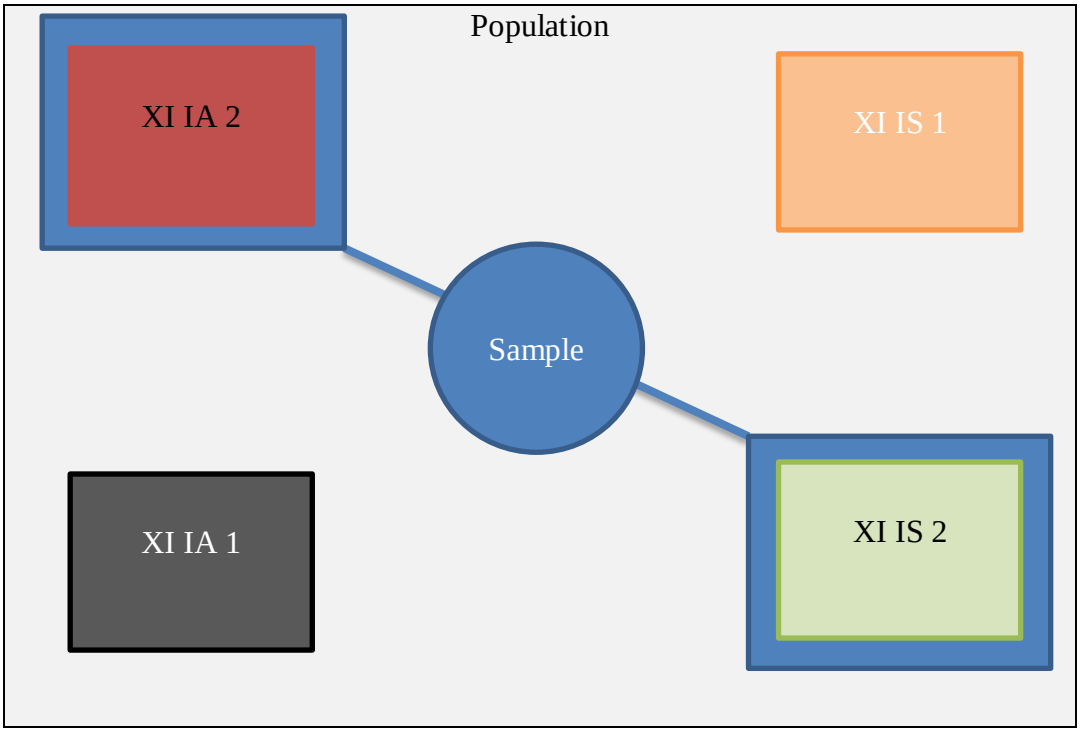

Figure 3.2 The Sampling Technique of the Study

The Figure 3.2 above shows that there were two groups from all of four groups selected as the sample of the study. All of the members of each class were following the study.

3. Sample of the Study

The sample was a subset of the population and is chosen in such a way that the degree of uncertainty in the sample values are known (Jobson, 1999: 12). Sample was also defined as the small group which was observed as a part of the population about the generalization was made (Ary, *et.al*: 2010).

There were two classes selected as the sample of the study. Each class consists of 30 and 29 students divided into two groups; the students having high

grammatical sensitivity and students having low grammatical sensitivity. Thus, the total number of the sample in the study was 59 students.

D. Treatment Procedure

The treatment given in the study was in the form of indirect corrective feedback. The treatment of the experimental group was in the form of indirect corrective feedback. Meanwhile, for the control group they were experiencing the traditional strategy in giving feedback that was direct corrective feedback. Thus, indirect corrective feedback was implemented on experimental group and the direct corrective feedback was implemented on control group.

Pretest on writing and grammatical sensitivity test were conducted to both groups. Writing pretest was conducted in order to know their initial ability in writing. It was also aimed at assuring that both groups were equal so that the result of the study truly represent what it supposed to yield that was to verify the effectiveness of indirect corrective feedback on the students' writing accuracy.

The treatment was given within six meetings; excluding the meeting for pretest and grammatical sensitivity test. After both groups were having six meetings of the treatment, the posttest on writing was then conducted. Below will be presented the treatment schedule for experimental and control group.

Table 3.2 The Treatment Schedule for Both Groups

Meeting	Activity	Date	
		Experimental	Control
1	Writing Pretest	March 11, 2019	March 14, 2019
2	Modeling the Text	March 18, 2019	March 21, 2019
3	Prewriting I	March 25, 2019	March 28, 2019
4	Drafting I	April 1, 2019	April 4, 2019
5	Revising I	April 8, 2019	April 12, 2019
6	Prewriting II + Drafting II	April 15, 2019	April 18, 2019
7	Revising II	April 22, 2019	April 25, 2019
8	Posttest	April 29, 2019	May 2, 2019

Table above shows that both experimental and control group were having the same amount of meeting. There were eight meetings; two meetings for pretest and posttest writing and six meetings for the treatment.

The lesson plan (see appendix 1 and 2) used for both the experimental and the control groups were designed by the researcher and were consulted to the thesis advisors as well as English teacher. The lesson plan for experimental and control group was different only on the way hoe researcher gave corrective feedback. Experimental group was given indirect corrective feedback and control group was given direct corrective feedback. Direct corrective feedback was chosen to be implemented in control group as it was common technique given by teacher in order to correct the students' piece of writing.

Due to the new technique applied for experimental group, the researcher first explained to the students regarding the symbols which were used to correct

the error. Below will be presented the procedures to teach both experimental and control group.

Table 3.3 Teaching Procedures for Both Groups

Experimental	Control
➢ Teacher explains and gives model of hortatory exposition text.	➢ Teacher explains and gives model of hortatory exposition text.
➢ Teacher explains and gives the example of the symbols used to give correction.	➢ Teacher asks the students to make an outline.
➢ Teacher asks the students to make an outline.	➢ Teacher together with the students discuss one of the student's outline
➢ Teacher together with the students discuss one of the student's outline	➢ Teacher gives correction on the students' outline
➢ Teacher gives correction on the students' outline	➢ Teacher gives back the outline
➢ Teacher gives back the outline	➢ Teacher asks the students to develop the outline into draft
➢ Teacher asks the students to develop the outline into draft	➢ Teacher makes correction on the students' draft.
➢ Teacher makes correction on the students' draft.	➢ Teacher gives back the students' draft.
➢ Teacher gives back the students' draft.	➢ Teacher asks the students to revise their first draft.
➢ Teacher asks the students to revise their first draft.	➢ Teacher asks the students to read their draft one more time.
➢ Teacher asks the students to read their draft one more time.	➢ Teacher asks the students to submit their final draft.
➢ Teacher asks the students to submit their final draft.	

Table above shows that both groups experiencing the same topic for learning. The only difference was on the use of indirect corrective feedback given by teacher to the students of experimental group.

1. Teaching Procedures in the Experimental Group

There were eight meetings for the experimental group, including pretest and posttest. The teaching activities consisted of six meeting starting from modeling the text, prewriting, drafting, and revising. The students were experiencing two times of producing hortatory exposition text. The first topic for their writing was breakfast and the second topic was the ex-school activity.

The first meeting of the treatment, researcher as a teacher explained the hortatory exposition text; definition, generic structures, language features, and example. Teacher explained the topic and conducted shared-discussion to discuss the text; identifying its generic structure, language features, and content of the text.

The next meeting, the teacher asked the students to make an outline for the first topic that was breakfast. Teacher together with the students discuss common error in writing hortatory exposition text as well as reviewing the symbols used as a means of giving correction. In the third to the sixth meeting, the teacher asked the students to write hortatory exposition text by their own. For full description can be found on appendix 1.

In giving the corrective feedback, teacher only gave the indication on the error without directly giving the correct form. The students had to repair their error based on the indication given by teacher. Below will be shown the example of indirect corrective feedback which was given to the students.

S≠V Prep sing/plural

‘Breakfast also ~~give~~ you a chance to get ✓ some vitamins and ~~nutrient~~ from healthy foods like ……’

Art Art

‘Breakfast provides ✓ body and brain with fuel after ✓ overnight fast. …’

Further, the errors that were given symbols were supposed to be corrected by the students themselves. The sentences above should be written as follows:

‘Breakfast also gives you a chance to get in some vitamins and nutrients from healthy foods like …’

‘Breakfast provides the body and brain with fuel after an overnight fast. …’

2. Teaching Procedures in the Control Group

There were eight meetings for the control group, including pretest and posttest. The teaching activities consisted of six meeting starting from modeling the text, prewriting, drafting, and revising. The students were experiencing two times of producing hortatory exposition text. The first topic for their writing was breakfast and the second topic was the ex-school activity.

The first meeting of the treatment, researcher as a teacher explained the hortatory exposition text; definition, generic structures, language features, and example. Teacher explained the topic and conducted shared-discussion to

discuss the text; identifying its generic structure, language features, and content of the text.

The next meeting, the teacher asked the students to make an outline for the first topic that was breakfast. In the third to the sixth meeting, the teacher asked the students to write hortatory exposition text by their own. For full description can be found on appendix 2.

In giving the corrective feedback, teacher gave the correct form of the error made by students. The students had to repair the error based on the correction given by teacher. Below will be shown the example of direct corrective feedback which was given to the students.

. Research

'Breakfast can be good for waistline too~~, research~~ shows those who eat

be to be

breakfast are less to ✓ overweight and more likely ~~tobe~~ ideal.'

Further, the errors that were given the correct form were supposed to be corrected by the students. The sentences above should be written as follows: 'Breakfast can be good for waistline too. Research shows those who eat breakfast are less to be overweight and more likely to be ideal.'

E. Instrument and Instrumentation

The instruments used in the study were in the form of test. There were two tests used in the study. The first was writing test and the second was grammatical sensitivity test.

1. Writing Test

Test is a set of questions or problems to measure ones' skill, ability, or knowledge. According to Arikunto (2004: 139), test is a set of questions or exercises or other means used to measure skill, knowledge, intelligence, ability, or talent of an individuals or group of people. Accordingly, writing test in this current study is defined as a task given to students in order to measure their skill in writing. Type of test which was used in the study was an essay test.

a. Development of Writing Test

To develop and construct the writing tests; pre- and post-test, there were some steps employed. The first step was developing test content specification by making blueprint. Making blueprint covered the activity of identifying the syllabus, determining the object of the test, kind of test, topic of test, time allocation, and scoring. The second was writing the test. Writing the test was done by preparing writing prompts and direction. The third was expert validation. Expert validation was conducted by asking the expert to review and make suggestion toward the test constructed by researcher. The validator for writing test in this study were; the thesis advisor, the English

lecturer, and the teacher of English at the school in which the study was conducted.

The instrument validator mentioned there was one thing needed to be revised on the writing test readability questionnaire. On the third question on questionnaire, the researcher had to change the term paragraph into the term words, so that the item question been like "*Apakah perintah soal tersebut menyebutkan secara tersurat berapa banyak kata yang harus dibuat dalam teks?*" rather than "*Apakah perintah soal tersebut menyebutkan secara tersurat berapa banyak paragraph yang harus dibuat dalam teks*". The next problem was on the third direction on the test. The clarity of the aspect would be evaluate according to the validator was not common but it was acceptable. Hence, the researcher decided not to omit the third direction, and just left it as what it was. Further, the English teacher suggested to lessen the words which must be write by the students from 250 words to 150 words with consideration that the students were not really good at writing.

After receiving some suggestions from the validator, then the last step was revising the test. The researcher then made revision on the test based on the suggestions given by the validators.

b. Tryout of Writing Test

In order to set valid test for resulting valid data, the researcher next tried out the tests. Since the pivotal validity of test is on its content validity (Brown, 2004), then try out is important as one of the efforts to determine its content validity. Some aspects showing the evidence of the validity are good logic and good wording (Latief, 2016).

Try out was conducted to one class at SMA Bustanul Muta'alimat. The test takers chosen were the eleventh grade students. They were outside of the sample of the study. Try out was conducted on March, 15th 2019 and it was followed by 19 students.

While trying out the test, the researcher distributed a questionnaire in order to know whether or not the instruction of the test was clear enough for students. Hughes (1996: 39-40) states that a good instruction have some requirements such as: (1) it must be clear and explicit; (2) it should avoid the supposition that all students know what is intended; and (3) it should not rely on the students' power of telepathy to elicit the desire behavior.

After trying out the test, researcher asked the students which part of the writing test which was not clear. Most of the students had no problem regarded the writing test. There were only some students which were confused on the use of the term "banned" on the test topic. It was due to their lack of vocabulary. As consequence, the researcher removed the term "banned" and replace it with the word "forbidden". The topic then became "The Prohibition of Smartphone at School".

In order to assure the test readability, during the process of trying out the test, the researcher distributed the questionnaire containing 5 items questions adapted from Hughes (1996) to check test readability. The questions contained questions as follow:

Table 3.4 Writing Test Questionnaire

No	Questions	Yes	No
1	Apakah perintah soal tersebut cukup jelas dan tidak bermakna ambigu?		
2	Apakah perintah soal tersebut menyebutkan secara tersurat bagaimana cara mengerjakan soal tersebut?		
3	Apakah perintah soal tersebut menyebutkan secara tersurat berapa banyak kata yang harus dibuat dalam teks?		
4	Apakah perintah soal tersebut meyebutkan secara tersurat hal-hal apa yang akan dinilaai dalam teks yang dihasilkan?		
5	Apakah perintah soal tersebut menyebutkan secara tersurat jenis teks yang harus dibuat?		

c. Validity and Reliability of Writing Test

The type of the writing test was in the form of essay test. Therefore, rational validity was fulfilled in this study. That was why the validity of the writing test as the instrument of the study was based on the approval of the expert on the use of an essay test in order to gain the data on the students' writing accuracy performance. All of the validators approved the instrument

used by researcher. It then could be concluded that the instrument used in this study was feasible to collect the data.

Meanwhile, in order to ensure the reliability of the data in the form of students' score taken from writing test, the researcher used inter-rater reliability. There were two raters for the same result of writing test. The raters were the English school teachers. The researcher provided the raters with the writing scoring rubric adapted from Brown (2007), concerning on three aspects of writing accuracy which were; grammar, vocabulary, and mechanics. Each component was described and was given score in order to ease the raters to give the scores. The score for each description of the components were also used to avoid subjectivity from the test raters. The full rubric can be seen on Appendix 6.

After having two scores from two different raters, the researcher then tested the reliability of the test using Intraclass Correlation computed by using SPSS v.23. It was the most common used statistic for assessing inter rater reliability for ordinal, interval, and ratio data number. If the Intraclass Correlation Coefficient was close to 1, it indicated perfect agreement between the raters.

The result of the computation showed that the value of Cronbach's Alpha coefficient was .963. It indicated very high reliability. Thus, the writing test of the present study was categorized as a test which having very high reliability. Complete computation can be seen in appendix 15.

2. Grammatical Sensitivity Test

Grammatical sensitivity test was test to measure the students' ability to distinguish grammatical functions. The test items contained 20 items questions adapted from Longman TOEFL written by Deborah Philips (2001). The test items were in the form of error recognition test.

In order to categorize the students' level of grammatical sensitivity, the researcher looked into the result of grammatical sensitivity test. The theory used to classify the students level of grammatical sensitivity was from Piraud (2006). According to Piraud (2006), the students with high level grammatical sensitivity were those who correctly answer 65% of 20 questions given. While those who weren't able to correctly answer the minimum criteria, were categorized as students with low level of grammatical sensitivity.

a. Development of Grammatical Sensitivity Test

Grammatical sensitivity test was constructed under several steps. The first, researcher made blueprint (see Appendix 9). Then, researcher wrote the test by doing adaptation from TOEFL book. The adaptation was in the form of simplifying the words or sentences used in the grammar recognition test item.

Afterwards, the researcher came to see the expert validator. Validator suggested that the grammatical sensitivity test should at least containing one clear example on how to do the test. Based on the expert's suggestion, the researcher then gave one example on how the test takers should do the test.

b. Tryout of Grammatical Sensitivity Test

To make the test appropriate to test the students' level, the grammatical sensitivity test was also being tried out. Trying out the grammatical sensitivity test was also aimed at ascertaining its validity and its reliability. The tryout had been carried out before the real test was administered. The tryout was employed at two classes at SMA Bustanul Muta'allimat consisted of 32 students as test takers. It was conducted on March, 16th 2019. The students of SMA Bustanul Muta'alimat were chosen as test takers participants due to their similar characteristics with the sample to whom the study would be employed. Therefore, they were assumed to be able to be genuinely representative for the population of test takers for whom the test was designed.

c. Validity and Reliability of Grammatical Sensitivity Test

After conducting the try out, the results of the grammatical sensitivity test were analyzed using SPSS program v.23. The analysis was intended to estimate the degree of its item difficulty, item discrimination, item validity, and the test reliability. SPSS v.23 was used to analyze each item of questions due to its effectiveness and efficiency.

1) Item Difficulty

This is defined simply as the proportion of test takers who answer an item correctly (Fulcher & Davidson, 2007). Test item should not be too easy or too difficult for the population for whom the test was designed. Henning (1987:50) as cited in Fulcher & Davidson (2007) mentioned that the acceptable range being around from 0.3 to 0.7, and values around 0.5 was

categorized as ideal item question. In this study, the categorization was based on the recommendation from Djiwandono (2008). Table 3.3 below shows the categorization guideline:

Table 3.5 Index of Item Difficulty Categorization

Index range	Category	Interpretation
.80 – 1.00	Very easy	Revised
.60 -- .79	Easy	Possible to be retained
.40 -- .59	Moderate	Possible to be retained
.20 -- .39	Difficult	Possible to be retained
.00 -- .19	Very Difficult	Revised

Based on the guideline above, the item difficulty of grammatical sensitivity used in this study was described in Table 3.4 below. The full result of the computation can be seen on Appendix 17.

Table 3.6 The Result of Item Difficulty Analysis

Index range	Category	Item Number	Status
.80 – 1.00	Very easy	4, 5, 7, 8, 20	Revised
.60 -- .79	Easy	1, 3, 9, 10, 16, 17	No revision
.40 -- .59	Moderate	2, 11, 12, 14, 15, 18, 19	No revision
.20 -- .39	Difficult	6	No revision
.00 -- .19	Very Difficult	13	Revised

Based on Table 3.4 above there were 5 items which were very easy, six items were easy, 7 items were moderate, 1 item was difficult, and 1 item was very difficult. Hence, items which were categorized into very easy and very difficult were revised to be implemented in the sample of the study.

2) Item Discrimination

The test should also be able to discriminate between higher ability and lower ability test takers (Fulcher & Davidson, 2007:103). The ability to discriminate is important in an approach to scoring that assumes that getting more correct answers is directly related to more of the ability in question, and that getting fewer correct answers is directly related to less of the ability in question.

The most commonly used method of calculating item discrimination is the point biserial correlation. This is a measure of the association between responses to any specific item (i.e. a 0 or a 1) and the score on the whole test (a continuous rather than a binary variable). Heaton (1989:180) mentioned that the items of discrimination ranged from +1 to -1. The higher the discrimination index, the better the item can determine the difference. Table 3.5 below guided the categorization of the items based on its discrimination capability as cited from Djiwandono (2008:24):

Table 3.7 Index Range for Item Discrimination

Index range	Category	Status
.40 – up	Very good	Accepted
.30 - .39	Good	Accepted
.20 - .29	Fair	OK
.19 - less	Poor	Revised or rejected

Based on the guideline above, the item difficulty of grammatical sensitivity used in this study was described in Table 3.4 below. The full result of calculation can be seen on Appendix 17.

Table 3.8 The Result of Item Discrimination Analysis

Index range	Category	Item Number	Status
.40 – up	Very good	2, 3, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19	Accepted
.30 - .39	Good	1, 4, 20	Accepted
.20 - .29	Fair	-	-
.19 - less	Poor	-	-

Based on Table 3.6 above there were 17 items which were very good in discriminating the ability of the students. There were also 3 items which were good and accepted as the items questions given to the sample of the study.

3) Item Validity

Validity is the most important consideration in developing and evaluating measuring instruments. Validity defined as the extent to which an instrument measured what it claimed to measure (Ary *et.al*, 2010:225). The focus of recent views of validity is not on the instrument itself but on the interpretation and meaning of the scores derived from the instrument. Validity of the grammatical sensitivity in this study was analyzed using SPSS v.23. The validity value (shown in *Sig*) must be lesser than the significance value $\alpha = .05$. Full calculation result was shown in Appendix 17.

Based on the result of computation, the validity of each item of grammatical sensitivity could be summarized as follows:

Table 3.9 Validity Coefficient per Item

Item Number	Sig. (two tailed)	Category
1	.048	Valid
2	.002	Valid
3	.018	Valid
4	.045	Valid
5	.001	Valid
6	.000	Valid
7	.003	Valid
8	.002	Valid
9	.000	Valid
10	.000	Valid
11	.000	Valid
12	.000	Valid
13	.017	Valid
14	.000	Valid
15	.000	Valid
16	.000	Valid
17	.004	Valid
18	.000	Valid
19	.001	Valid
20	.042	Valid

.

Table above showed that the item of questions were all valid. Therefore, it was possible to employ the grammatical sensitivity test to group the students based on their grammatical sensitivity level.

4) Reliability

Whenever a test is administered, the test user would like some assurance that the results could be replicated if the same individuals were tested again under similar circumstances. This consistency means reliability (Fulcher & Davidson, 2007:104). The classical formulation of

this was put by the British psychologist Charles Spearman in his work on correlation between 1907 and 1913. He argued that an observed score on any test was a composite of two components: a true score and an error component.

According to Ary *et.al.*, (2010), the reliability index ranged from 1.00 for a perfect reliable test to 0.00 for totally unreliable test. The categorization was described in Table 3.8 below:

Table 3.10 The Reliability Index

Index range	Category
.800 – 1.00	Very High
.600 - .799	High
.400 - .500	Fair
.200 – .399	Low
< .200	Very Low

Based on the result of computation (see Appendix 17), the reliability of the grammatical sensitivity test was .750. Thus, the test was categorized into test with high reliability.

F. Technique of Collecting Data

Referring to the variables of the study, the technique for collecting the data in this current study was by administering test. The data were collected through administering both writing and grammatical test. Grammatical sensitivity test was done before the treatment being given, while writing test was employed after the treatment was began.

Grammatical sensitivity test was done on March 19 2019 for the experimental group, and on March 21 2019 or the control group. The test was consisted of 20 items question in the form of error recognition test. The students were having 60 minutes to do the test. The test was started at 07.30 AM and finished at 08.30 AM.

After grammatical sensitivity test employed, the researcher then check the students' question. All of the students' question were recapitulated and scored in order to categorize the students into their level of grammatical sensitivity. The students who could answer correctly ≥ 13 items questions were categorized into the students with high level of grammatical sensitivity. Meanwhile, the students who were able to correctly answer less than 13 items were categorized into the students with low level of grammatical sensitivity.

Further, writing posttest was done on May 14 2019 for the experimental group. Meanwhile, for the control group it was done on May 17 2019. The test was in the form of essay test. The students were being given some topics to be selected as the theme for their writing of hortatory exposition text. The time allotment for doing the test was 60 minutes. The test was started at 07.30 AM, and it was finished at 08.30 AM. The teacher distributed the test and the students directly do the test as the instructions given.

After the writing posttest was done, the researcher herself together with the English teacher evaluated and scored the students' writing. The scoring was done based on the scoring rubric which has been set previously. The scoring was divided into three aspects focusing on the element of writing accuracy. The two

raters gave specific points for each writing accuracy elements. After the score from the two raters were collected, both scores from the two raters for each components of writing accuracy were added up and averaged to have one single score for each component. Then, the score from each component were added up and calculated in order to have final score for the students' writing. Full description can be seen on appendix 22 for experimental group, and appendix 23 for control group.

G. Technique of Analyzing Data

Techniques of analyzing the data for current study were descriptive and inferential statistic. Descriptive statistic was used to know mean, median, mode, and standard deviation of the writing test score of each group. Meanwhile, inferential statistic was used to test the hypothesis, whether or not it can be used to make generalization for the population.

However, before testing the hypotheses, the researcher conducted prerequisite tests covering; normality testing and homogeneity testing. Normality testing was conducted in order to make sure that the data was normally distributed, while homogeneity testing was conducted to ensure whether the data are homogenous or not. The data in this study was called normally distributed and homogenous because the significance values of both groups were greater than the level of significance $\alpha = .05$. These two prerequisite tests were computed using SPSS version 23.

Then, after completing normality and homogeneity testing, the researcher came up to the hypotheses testing using statistical calculation. For the research question no 1 the researcher used independent T test, while for research question no. 2 and 3, the researcher used statistical calculation called Two-way ANOVA to check whether or not there was interaction between corrective feedback and level of grammatical sensitivity. All of the computations were using SPSS version 23. Two-way ANOVA was used as a consequence that in this study, there were two groups that have been split on two independent variables (called factors). The primary purpose of a two-way ANOVA was to understand if there was an interaction between the two independent variables on the dependent variable, in this case was the interaction between corrective feedback and grammatical sensitivity level on the students' writing accuracy score.

H. Statistical Hypotheses

In this study the researcher proposes four hypotheses. These hypotheses are based on the formulation of the problem as presented in the previous chapter.

1. Ho_1 : $\mu A_1 = \mu A_2$ (there is no significant difference score on writing accuracy between the students taught by using indirect corrective feedback and those taught by using direct corrective feedback)

   Ha_1 : $\mu A_1 > \mu A_2$ (The students who are taught using indirect corrective feedback have significantly better writing accuracy skill than those taught using direct corrective feedback)

2. Ho_2 : $\mu B_1 = \mu B_2$ (There is no significant difference score on writing accuracy of students based on the different level of grammatical sensitivity)

   H_{a2} : $\mu B_1 > \mu B_2$ (The students with high level of grammatical sensitivity have significantly better writing accuracy than those with low level of grammatical sensitivity)

3. H_{03} : A x B = 0 (There is no interaction between indirect corrective feedback and students' grammatical sensitivity level)

   H_{a4} : A x B > 0 (There is an interaction between indirect corrective feedback and students' grammatical sensitivity level)