

The American University in Cairo

School of Humanities and Social Sciences

THE EFFECT OF ERROR CORRECTION TYPES ON  
GRAMMATICAL ACCURACY IN STUDENT ESSAY REVISION

A Thesis Submitted to

The Teaching English as a Foreign Language Department

THE ENGLISH LANGUAGE INSTITUTE

In partial fulfillment of the requirements for the degree of Master of

Arts /Science

By

Neveen Saeed Al Saeed

June 2010

The American University in Cairo

THE EFFECT OF ERROR CORRECTION TYPES ON  
GRAMMATICAL ACCURACY IN STUDENT ESSAY REVISION

A Thesis Submitted by

Neveen Saeed Abd Al Kareem Al Saeed

To the English Language Institute

TEFL Program

June 2010

In partial fulfilment of the requirements for the Degree  
of Master of Arts in TEFL

**has been approved by**

Dr. Amira Agameya  
Thesis Committee Chair \_\_\_\_\_

Dr. Robert Williams  
Thesis Committee Reader \_\_\_\_\_

Dr. Lori Fredricks  
Thesis Committee Reader \_\_\_\_\_

\_\_\_\_\_  
Program Director                      Date                      Dean                      Date

## Table of Contents

<b>Abstract.....</b>	<b>6</b>
<b>Chapter 1; Introduction .....</b>	<b>7</b>
1.1    Historical Context and Arguments .....	7
1.2    Statement of Research Problem.....	9
1.2.1    Significance and need of the study .....	9
1.3    The Purpose of the Study.....	9
1.4    Research Question .....	10
1.4.1    Main question .....	10
1.4.2    Sub-questions.....	10
1.5    Delimitations .....	10
1.6    Definitions of Constructs.....	11
1.6.1    Theoretical definitions .....	11
1.6.2    Operational definitions.....	11
1.6.3    Definitions of variables.....	11
<b>Chapter 2; Review of Literature .....</b>	<b>12</b>
2.1    Details of the Nature of the Search.....	12
2.2    Criteria for Including Studies .....	12
2.3    The Structure of the Review .....	13
2.4    The Historical Background.....	13
2.4.1    The Truscott-Ferris debate.....	14
2.5    Studies Against the Effectiveness of Error Feedback .....	15
2.6    Studies Supporting the Effectiveness of Error Feedback .....	17
2.7    Studies Comparing Feedback and No Feedback Groups .....	21
2.8    Meta-analysis and Reviews of Literature on Feedback Studies .....	28
2.9    Conclusion.....	32
<b>Chapter 3; Methodology and Data .....</b>	<b>34</b>
3.1    Proposed Design of Study .....	34
3.1.1    Design .....	34
3.1.2    Participants.....	34
3.1.3    Data collection procedures.....	35
3.1.4    Pretest and posttest.....	36
3.2    Treatment.....	36

3.2.1	Treatment time .....	36
3.2.2	Treatment group A.....	37
3.2.3	Treatment group B .....	38
3.2.4	Control group.....	39
3.2.5	All groups .....	39
3.2.6	Some considerations .....	39
3.3	Observational Techniques .....	39
3.3.1	Rating.....	39
<b>Chapter 4; Data Analysis and Results .....</b>		<b>41</b>
4.1	Introduction .....	41
4.2	Data Analysis.....	41
4.2.1	Statistical procedures .....	41
4.2.2	Results.....	42
4.3	Analysis of Variance .....	51
4.4	Fluency.....	51
<b>Chapter 5; Discussion of Results .....</b>		<b>54</b>
5.1	Summary of the Results.....	54
5.2	Discussion of the Results.....	54
5.3	Some Extraneous Factors .....	56
5.3.1	Proficiency level .....	56
5.3.2	Control group.....	57
5.3.3	Motivation.....	57
5.3.4	H1N1 scare .....	57
5.3.5	Relation to previous research.....	58
5.4	Limitations of the Study .....	58
5.4.1	Second drafts as a measure of learning.....	58
5.4.2	Timing of the treatment .....	59
5.4.3	Duration of the study .....	59
5.4.4	Proficiency level .....	59
5.4.5	Implications for future research .....	59
<b>Appendix A .....</b>		<b>61</b>
<b>Appendix B .....</b>		<b>62</b>
<b>References.....</b>		<b>63</b>

## **ACKNOWLEDGEMENT**

I am heartily thankful to my supervisors, Dr. Amira Agameya and Dr. Robert Williams, whose encouragement, guidance and support from the beginning to the final level enabled me to develop an understanding of the subject.

Besides my advisors, I would like to thank the rest of my thesis committee: Dr. Lori Fredricks, for her encouragement and insightful comments.

I am also thankful to my family who supported me unconditionally and tolerated my mood swings all through the process of thesis writing. I also thank them for listening to my complaints and frustrations and for believing in me.

Lastly, I offer my regards to all of those who supported me in any respect during the completion of the thesis.

Neveen Al Saeed

### **Abstract**

There has been an on-going debate on whether error feedback helps students to improve their grammatical accuracy from one draft to the other (Ferris, 1999; Truscott, 1996; 1999). Some studies found that error correction was effective (Ferris & Roberts, 2001) and others refuted this argument (Semke, 1984; Kepner, 1990). Their findings showed that feedback had no or non-significant effects on accuracy. According to previous research, one area which has not been properly studied is a comparison between groups receiving feedback and a no feedback group. Therefore, the present study investigated the effect of feedback on grammatical accuracy by comparing three types of feedback: errors coded, errors underlined, and no grammar feedback. The instruments included a pretest, a posttest and two treatments. Participants were first- year students in the English Literature and Language Department at Ain Shams University. Means and standard deviations were calculated. ANOVA and dependent t-tests were also used to analyze the data. Data analysis revealed no significant differences in the improvement of grammatical accuracy in the two treatment groups. It was however found that both treatment groups outperformed the no-feedback group in editing verb and noun-related errors. However, the control group outperformed the two other groups in correcting their article-related errors. The researcher concluded that even if students do not receive grammar feedback, their writing improves because of the rewriting process itself.

## Chapter 1

### Introduction

#### 1.1 Historical Context and Arguments

Although the topic of error correction and its effect on EFL writing has been researched since the 1960's, it has become a topic of controversy since Truscott published his seminal review of the effect of grammar correction on L2 writing in 1996. He argued that error correction should be abandoned because it is not effective in helping improve the accuracy of students' writing. According to him, error correction even harms their writing. Since then, researchers have tried to either refute or support Truscott's argument.

Ferris (1999), for example, countered by arguing that Truscott's work was biased because he focused only on studies that supported his opinion and ignored those that contradicted his position. She asserted that more compelling evidence was needed before claiming that error correction should be abandoned. In fact, a great number of studies that investigated the effect of error correction on accuracy had conflicting results regarding the effectiveness of error feedback (e.g. Ashwell, 2000; Chandler, 2003; Ferris & Roberts, 2001; Ferris, 2003; Kepner, 1991; Lee, 2004; Polio, Fleck, & Leder, 1998). Some of these studies concluded that grammatical error correction is effective (e.g. Ashwell, 2000; Ferris & Roberts, 2001); while other studies found grammatical error correction to be ineffective (e.g. Polio et al., 1998; Kepner, 1991)

Most research on feedback compared the effect of different types of feedback (direct, indirect, coded or uncoded) on the grammatical accuracy of the writing of students (e.g. Robb, Ross, & Shortreed, 1986) but did not explore the difference between groups who receive feedback and those who do not. Therefore, there is a need for further experimental research that compares groups receiving different types of feedback with a no correction group. According to Ferris (2004), only six studies (Ashwell, 2000; Fathman & Whalley,

1990; Ferris & Roberts, 2001; Semke 1984; Polio et al., 1998; Kepner, 1991) made this kind of comparison. These studies investigated the effect of error correction on the grammatical accuracy of L2 writing, but had conflicting results. Ashwell (2000), Fathman & Whalley (1990), and Ferris & Roberts (2001), agreed that error feedback improves the grammatical accuracy of students' writing. For example, Ferris & Roberts (2001) had three groups in their study: a coded feedback group, an uncoded feedback and a no feedback one. They found no significant differences between the two groups that received coded and uncoded feedback in terms of grammatical accuracy. However, these two error feedback groups outperformed the no feedback group in grammatical accuracy. Their findings supported those of Ashwell (2000) and Fathman and Whalley (1990). (Please see the review of literature section for reviews of these studies.)

In contrast, the three other studies found feedback to be ineffective. Semke (1984), for example, had four groups in her study. Three groups received error correction, (direct feedback, indirect coded feedback, and a mixture of form and content feedback), and the fourth group only received feedback on the content. Semke found that error correction had no effect on writing accuracy. Chandler (2003) suggested that the reason for Semke's findings could be that the no feedback group wrote twice as much as the other groups. Hence, they improved in their writing.

Similarly, Polio et al.'s study (1998) consisted of two groups, one using error correction and one that did not. Again, the no correction group wrote twice as much as the correction group. In this study, the findings suggested that all groups improved in accuracy and there were no significant differences between the groups. In Kepner's study (1991), the error correction group did not produce fewer errors than the no feedback group. This might be because the students were not asked to make use of the feedback in the following drafts

(they wrote journal entries). In the review of literature section, the above-mentioned studies are discussed in detail.

## **1.2 Statement of Research Problem**

The aim of the present study is to further investigate the effect of two different types of feedback (coded and uncoded feedback) on developing the grammatical accuracy of students' writing. In addition, these feedback groups were compared to a "no feedback" control group to explore which is more useful.

### **1.2.1 Significance and need of the study**

This topic has important implications for writing Academic English as a Foreign Language (EFL), because writing is a complex skill that is essential for academic success. Most EFL learners face difficulties in this skill. Even if their presentation of ideas is fluent, grammatical errors detract from the quality of their writing.

Egyptian EFL university students were chosen as participants in this study, because they often have severe problems with grammatical structures. The researcher, who works at an Egyptian national university, has noticed that writing teachers pay special attention to grammatical errors while correcting their essays. However, students are not given the chance to rewrite their essays. If the study's findings suggested that error feedback and rewrites proved to be effective in writing classes, then teachers could be advised to use multiple drafts. This type of feedback helps the students to identify their errors and correct them. However, if the no correction feedback turned out to be more effective, this could be a starting point for further research that explores whether teachers should abandon error correction.

## **1.3 The Purpose of the Study**

Both Ferris (1999) and Truscott (1996; 1999) agreed that present research done on the effect of error correction on student writing is still insufficient to determine whether error

correction improves grammatical accuracy of students' writing. In addition, they stressed the need for more experimental studies comparing groups that receive feedback with groups that do not. This study investigated these issues by comparing the effect of feedback on grammatical accuracy of writing in the Egyptian EFL classroom.

The Egyptian EFL university students in the study are intermediate freshman students who are majoring in English Literature in an Egyptian national university. Their writing is usually full of grammatical errors. Some of their common errors are article usage, subject verb agreement, verb tenses, and run on sentences. Their teachers lower the grades of students who commit these errors because they are English literature students. This study explored the effect of feedback on grammatical accuracy using a different kind of subjects and in a different place.

## **1.4 Research Question**

### **1.4.1 Main question**

What are the effects of different types of error feedback or no feedback on the Egyptian EFL students' ability to revise their grammatical errors in their writing?

### **1.4.2 Sub-questions**

- i. Does the use of coded feedback in academic writing instruction have an effect on grammatical accuracy in the revision process?
- ii. Does the use of uncoded feedback in academic writing instruction have an effect on grammatical accuracy in the revision process?
- iii. Do students receiving no grammar feedback on their academic writing improve in grammatical accuracy in the revision process?

## **1.5 Delimitations**

This study will focus on grammatical error correction, excluding word choice errors. It does not focus on students' or teachers' perspectives of the correction. Moreover, it is a

cross-sectional study, not a longitudinal one, because the treatment will be administered in one semester. This study does not deal with peer correction; it only focuses on teacher feedback, how it differs from giving no feedback and how it affects the students' writing.

## **1.6 Definitions of Constructs**

### **1.6.1 Theoretical definitions**

*Grammatical accuracy of writing* is the ability of the students to write essays that are free from grammatical errors and to edit their errors (Ferris, 2001).

*Error feedback* has been defined as the indication of the grammatical errors found in a student's text to help the student produce accurate writing (Ferris, 2003).

### **1.6.2 Operational definitions**

*Grammatical accuracy of writing* is the ability of the student to edit grammatical errors. It is calculated as the total number of errors per essay.

*Error feedback* is the marking of the most common grammatical errors in a students' writing either by coded or uncoded feedback (indicating the type of error or only underlining it). In this study, three types of error feedback were used:

1. *No feedback group* is the group of students who receive no grammatical feedback on their essays. They only receive content feedback and are asked to rewrite their texts.
2. *Coded feedback group* is the group of students who receive codes, for example VT (i.e. verb tense), to refer to the type of error he/she did. (See Appendix A).
3. *Uncoded feedback group* is the group of students whose grammatical errors are only underlined. They have to find out the type of error and try to edit it.

### **1.6.3 Definitions of variables**

For the purpose of the study, there are three types of feedback (coded, uncoded, and no feedback) used. They will be measured as three independent variables. There is also a dependent variable: grammatical accuracy.

## **Chapter 2**

### **Review of Literature**

The present review summarizes and critiques the different studies in the field of EFL writing, especially those dealing with the effect of feedback on students' writing. In other words, this review is an attempt to present a complete picture of the research done in the field of error feedback and EFL writing.

#### **2.1 Details of the Nature of the Search**

To find studies for this literature review, several sources were used. Four electronic preliminary sources were used to find primary research studies: JSTOR, ERIC, LLBA, and Academic Search Premier. The keywords that helped in finding studies suitable for the search included error correction, feedback, writing, EFL writing, ESL writing, EFL, ESL and language learning. The search was limited to journal articles found in full text on the databases used. Some combinations of key words such as EFL and error correction resulted in zero hits. In addition to the online journals, the American University in Cairo library was searched, both the books and the journals it is subscribed to. This review covers research studies from 2000 to 2008. However, older studies (e.g. Ferris, 1999; Semke, 1984; Truscott, 1996) were used as part of the historical background of the field.

#### **2.2 Criteria for Including Studies**

The selection of the studies to be used in the literature review was based on certain criteria. The studies which dealt with the topic of error correction as a technique in EFL/ESL writing classrooms or which explored the effect of different types of feedback on writing were reviewed thoroughly. The review only focused on college or graduate students; therefore, studies that dealt with children were excluded. Many studies that explored the perceptions of the teachers and students to the usefulness of error correction were found in the search. Only two studies dealing with perceptions were included in the review to support

the view that students appreciate feedback on their work. However, the researcher only used two because it is not the focus of the study. In addition, the year of publication was another element that helped in choosing which articles to include. Studies that were reviewed ranged from year 2000 to 2008. After discussing the criteria for choosing the articles for the review, the next section will clarify the structure of the review and a rationale for this organization.

### **2.3 The Structure of the Review**

The first part of the review describes the historical background of the field. In other words, the introduction shows the development of the feedback, its uses and the ways it was perceived. Then the first section of the review consists of the studies that opposed the effectiveness of error feedback on the students' writing. As for the second section, studies supporting feedback will be reviewed. The third part of the review is more specific in its nature and is the most important part of the literature review because it is directly related to the problem of the study. The articles reviewed in this section compare students who receive feedback on their writing and those who do not. This section includes six studies (Ashwell 2000, Semke 1984, Fathman & Whaley 1990, Ferris and Roberts 2001, Polio et al 1998, Kepner 1991).

Other studies were discussed in the historical background of the study. At the end, the findings of these studies are compared to reach a conclusion.

### **2.4 The Historical Background**

Researchers' views about error feedback and its importance have undergone a lot of changes over the years. According to Ferris (2003), error correction was first thought of as an essential means of helping students to use or practice their language. In other words, the main focus of the teachers at that time was the students' grammatical accuracy. In the 1970's there was a shift in this focus; teachers started to focus on the process of writing rather than the accuracy. This shift in focus resulted in a total neglect of accuracy and complete attention of

creativity and originality in writing. As a response to the process approach, some researchers such as (Horowitz 1986) thought that the total neglect of form was counterproductive, since academic writing needs to follow a certain mode of presentation. He also argued at that time that the writing process is not enough to familiarize students with the rules of academic writing. As a result, the interest in error feedback was renewed and many studies explored the effectiveness of error feedback on the writing (e.g. Semke, 1984).

#### **2.4.1 The Truscott-Ferris debate**

In 1996, Truscott published his seminal review in which he explored and questioned the effect of grammar correction on L2 writing, especially its effect on grammatical accuracy. His focus was on the role of error correction in helping students restructure their IL grammar. He based his argument on second language acquisition theories. In other words, he argued that feedback could be harmful because students could be developmentally not ready to correct the error. In addition, Truscott argued that he did not find enough evidence to support the effectiveness of error feedback in improving writing. He claimed that since feedback is affected by teachers' and/or students' motivation and readiness to give and/or edit errors, this debate was useless.

Ever since Truscott published his review, researchers have tried to either refute or support Truscott's findings. Ferris (1999) was one of the researchers who disagreed with Truscott. She argued that Truscott's work was biased because he focused only on the findings that supported his opinion and ignored those that contradicted him. She added that all the research cited by Truscott had different designs and participants. Therefore, she asserted that more research was needed before claiming that error correction should be abandoned.

Most of the studies that investigated the effect of error correction on accuracy had conflicting results regarding the effectiveness of error feedback. Some of these studies concluded that grammatical error correction is effective (Ashwell, 2000; Ferris & Roberts,

2001); while others found grammatical error correction to be ineffective (Fazio, 2001; Kepner, 1991; Polio et al., 1998). In the next two sections, studies that explored these two views will be discussed.

## **2.5 Studies Against the Effectiveness of Error Feedback**

Due to the ongoing debate on the effectiveness of teacher feedback and whether it improves grammatical accuracy, researchers became more interested in investigating this topic. Some of these researchers concluded that feedback had no benefits (Fazio, 2001; Kepner, 1991; Lalande, 1982). While some of these researchers compared the effect of three types of feedback on students' writing (Fazio, 2001; Lalande 1982), others compared between students who receive feedback and those who do not (Kepner, 1991).

In his study on 60 intermediate students, Lalande (1982) investigated the outcome that different feedback mechanisms had on grammatical accuracy. The study was quasi experimental and longitudinal. The pretest and posttest were 45 minute in-class essays. Participants were divided into four groups: two control and two experimental groups. All groups received the same instructional material. The only difference between the groups was the kind of feedback they received. Control groups received direct feedback and teachers provided them with the corrections. All they had to do was copy the corrections in their rewrites. Treatment groups, on the other hand, received codes as their feedback types. In their rewrites, students were asked to correct their mistakes and rewrite the whole essay. In addition, they were asked to fill an error awareness survey to keep track of their most frequent errors. Experimental groups outperformed the control ones. However, Lalande pointed out that there was no statistically significant difference between the two groups and that more longitudinal studies were needed before these findings could be generalized. This finding comes in agreement with other research findings that corrective feedback has no significant effect on accuracy (Fazio, 2001; Truscott, 1996, 1999).

According to Truscott (1996) and Ferris (1999), in order to test the true effects of feedback, researchers have to compare the accuracy of students receiving different types of feedback. Accordingly, Fazio (2001) examined the effect of three types of feedback (corrections, commentaries, and a combination of the two) on grammatical accuracy. She argued that error feedback had no significant effect on students' writing accuracy. For her study, Fazio used 112 students who were in grade 5 in a French school. Students were divided into three groups according to the type of feedback they received. Fazio used in-class journal writing, class observations and interviews as data for her study. Findings showed that none of the three groups improved in their accuracy. However, it is worth noting that Fazio attributed lack of improvement to the short treatment time which could have affected the results.

Kepner (1991) investigated the question of whether there are any differences in grammatical accuracy between groups receiving error feedback and those who do not which is an important gap in the literature. Using a quasi-experimental design, Kepner used treatment (who received feedback) and control (who received no grammatical feedback) groups. His study used 60 participants enrolled in a Spanish class, who were divided into four groups. Kepner asked two teachers to teach these groups (i.e. each teacher taught two groups.) In other words, Kepner gave two groups feedback on their grammatical errors, while the other two groups received no such feedback. Participants were required to write journal entries as a response to each of their eight assignments. After collecting the data, the researcher analyzed it statistically. Results showed that the error feedback group improved more than the control group by 15%, which for him was not a significant improvement. Kepner concluded that error correction did not help students avoid the sentence-level errors. These findings support other research that claimed that error feedback is ineffective (Truscott, 1996, 1999).

Rob, Ross and Shortreed's study (1986) is different than the previously reviewed studies in that they found that while direct feedback is effective, indirect feedback has insignificant benefits. They examined the effect of four types of error feedback: direct, coded, uncoded, and the number of errors per line. There were 134 Japanese participants in this quasi- experimental study. Participants were divided into four groups depending on the type of feedback they received. The researchers aimed at testing the hypothesis that the direct feedback has the strongest impact on the writing of the students. Each participant produced five narrative compositions over a nine-month academic year. The researchers analyzed the data statistically to assess the writing accuracy, complexity and fluency. The findings of this study showed no statistically significant differences between the four feedback groups on any of the three measures (complexity, accuracy or fluency). However, the accuracy of all four groups improved. Rob et al. (1986) concluded that there was no use for indirect feedback and that direct feedback was both less time consuming for teachers and helpful for the students and their writing.

The above reviewed studies highlighted the non-significant effect of feedback on grammatical accuracy of students' writing. While these studies concluded that feedback is not beneficial, others found that it helped in improving students' grammatical accuracy (Chandler, 2003; Ferris, 1997; Ferris & Roberts, 2001; Gascoigne, 2004). These benefits of error feedback are highlighted in the following section.

## **2.6 Studies Supporting the Effectiveness of Error Feedback**

A number of studies were carried out to explore the effects of different types of feedback to refute Truscott's attack on error feedback (1996). Most of their findings showed that error feedback helps improve students' writing accuracy. Based on Truscott's recommendation, studies in this section explored two main points: the effect of different types of feedback and the difference between feedback and no feedback groups. Accordingly,

Ferris and Roberts (2001) explored three different types of feedback (coded, uncoded, and no feedback). In their study, they had four research questions. In the first question they explored the effect of different types of feedback (coded, uncoded and no feedback) on the ability of students to self edit their errors. The second question examined whether students corrected certain types of errors more than other ones. The third question dealt with the students' perspectives about the types of error feedback they need the most. Finally, the researchers explored the influence of prior grammar knowledge on the students' ability to edit their texts.

Ferris and Roberts (2001) used a quasi-experimental design for their study. They had three treatment groups (coded, uncoded, and no feedback groups). Their 44 participants were randomly assigned to these groups. To answer the first and second questions, they gave a pretest (a 50-minute in-class essay) to all of their participants and after that, the researchers corrected their essays using the three types of feedback and gave them back to the students. The students, then, attempted to self-edit their essays. To answer the third research question, Ferris and Roberts (2001) used a five-item questionnaire to survey students' opinions about their experiences in studying English grammar, the problems they faced when writing essays and feedback type they preferred.

Statistical procedures were used to analyze the data. The findings of this study showed that the two groups who received feedback significantly outperformed the no feedback group. However, there were no significant differences between the coded and uncoded feedback. The participants succeeded in editing "treatable" errors like verbs and nouns more than untreatable errors such as word choice. As for the students' perception of error feedback, all students agreed that they expected feedback from the teacher and most of them preferred the coded feedback. Although the questionnaire showed that students were all familiar with grammar, they did not know how to use their knowledge while writing their pretest essays. This knowledge was used in the self-editing exercise.

Another study that targeted the same topic was Chandler (2003). He published a two-study article exploring the effect of error feedback on students' writing. In the first study, Chandler examined whether teachers should give feedback or not. He had three research questions. First, he investigated whether feedback allows students to make fewer errors at the end of the semester. Second, he explored whether students who do not correct their errors make fewer errors in their next assignment. Third, Chandler compared between the grammatical accuracy of students who correct their errors and those who do not.

Chandler's study was an experimental study. He had a treatment and a control group. The two groups were asked to write five written assignments about their lives. Each assignment was five pages long. Then the teacher gave the students error feedback on their essays. The experimental group was asked to correct all the errors that the teacher underlined before submitting the next assignment. On the other hand, the control group corrected their errors at the end of the semester. Chandler (2003) used statistical procedures to analyze his data. He found that students who did not correct their errors after each assignment did not improve in accuracy. However, the experimental group increased in accuracy. Both groups increased in fluency.

In the second study, Chandler (2003) aimed to explore the effect of four different types of error correction on students' writing. This is similar in a way to Ferris and Roberts' study (2001). The research question explored how the teacher should give feedback to the students to improve their writing. In this study, Chandler used an experimental design again. The students were asked to write five assignments, eight pages each. The teacher corrected the students' assignments, of both groups, using four types of error correction: correction, underlining describing the error, describing error only, and underlining only. Statistical analysis was used to analyze the data. Chandler found that the students' accuracy and fluency improved over the semester. Moreover, results showed that feedback in the form of teacher

correction and underlining had a statistically significant effect on students' essays, whereas the other two feedback types did not. This finding, according to Chandler, could be because students found it easier and quicker to correct the errors using their teachers' commentary or underlining. The quality of the writing remained complex in all groups.

The findings of these studies are supported by other studies that explored the same issue (Ferris, 1997; Gascoigne, 2004). Ferris (1997) explored the effect of teacher commentary on student essays. She investigated how teacher comments written in the margins or at the end of students' essays improved the quality of students' second drafts both grammatically and content. For the purpose of the study, Ferris analyzed comments on both grammar and content. She also explored the type of commentary whether they were positive or negative and if they were in the form of questions, imperatives, or requests for more information or clarifications. The research questions that Ferris aimed at answering targeted the types of teacher feedback (comments in the form of questions, requests, or imperatives) that are most effective on students' rewrites and the extent to which students made changes in their papers (due to teacher feedback). To answer these questions, Ferris used 47 freshmen and sophomore students who were enrolled in an ESL composition class. The students were asked to write 4 essays and a minimum of three drafts for each. Ferris used 110 pairs of essays and their rewrites for her study. Four coders were trained to code the essays to see which types of teacher feedback were most influential on students' rewrites in terms of accuracy and fluency. In other words, the researcher explored the type of commentary that helped students the most to correct their grammatical and content errors in the rewrites.

In addition, Ferris used a rating scale to explore how these comments were incorporated in the rewrites. Regarding the most influential type of teacher feedback, Ferris found out that longer feedback had a greater effect on the rewrites than short or general comments. Another finding was that marginal requests for information and grammar

comments helped students write better drafts. Therefore, teacher feedback was found to positively affect revised drafts.

Gascoigne (2004) based his study on Ferris' 1997 research. He investigated whether teacher feedback helped students improve, and what factors of corrective feedback influence beginner students' writing. His study included 25 freshman students who were native speakers of English enrolled in a French class. For the sake of the study students were asked to write 8, 50 minute in-class essays. Teachers wrote their comments and gave them back to the students to revise. Then, like Ferris, he calculated, on a scale of 0 to 6, the effect of teacher commentary on students' revisions. He supported Ferris' findings that corrective feedback helps improve students' writing.

After reviewing the different viewpoints regarding the effectiveness of feedback, the next section explores studies that compared between students who receive feedback and those who do not.

## **2.7 Studies Comparing Feedback and No Feedback Groups**

It is clear from the studies reviewed in the previous sections that more studies that compare groups receiving feedback to those who do not are needed. According to Truscott (1996), without this kind of comparison no one can be sure about the benefits of feedback. This section reviews most studies that did this comparison. While some of these studies found that feedback is effective (Ashwell 2000, Bitchener, Young, & Cameron 2005, Fathman & Whalley (1990), and Ferris & Roberts 2001), others supported Truscott's view of the ineffectiveness of feedback (Semke 1984; Polio et al., 1998; Kepner, 1991; Truscott & Hsu, 2008).

Ferris and Roberts (2001) conducted this comparison in their study and had positive results about the benefits of feedback. Other studies like Bitchener, Young, & Cameron (2005) and Fathman & Whalley (1990) had the same findings. For example, Bitchener,

Young, & Cameron (2005) designed their study to compare between groups receiving feedback with those who do not. Their study explored the effect of two different types of feedback (direct feedback and no feedback) on grammatical accuracy. For their study they focused on three target forms: past simple, prepositions, and definite articles. They had three groups: two treatment and one control group. Group One received direct feedback and five minute conferences to discuss their essays. As for Group Two, they received only direct feedback. Group Three did not receive any feedback on the three target forms. Findings of the study proved that direct feedback did not help students improve in their accuracy. On the other hand, the group that received both direct feedback and conferences improved significantly in their production of articles and past simple tense but not prepositions.

Fathman and Whalley's (1990) study is another recent research which explored the effect of four types of feedback on students' essays. The researchers designed their study to look at feedback on form and content and to find out the most effective type of feedback on students' writing. They also investigated the best stage for teachers to give students feedback on form versus feedback on content. Fathman and Whalley used 72 intermediate ESL students. For the purpose of the study, they divided them into four groups. Each group received a different kind of feedback. Group One received no feedback on their essays. Group Two received only grammar feedback. Students found their errors underlined and they had to correct them. As for Group Three, they were provided with content feedback. Lastly, the essays of Group Four were marked with both grammar and content feedback.

Fathman and Whalley found that both grammar and content feedback were effective. However, grammar feedback was more effective than content feedback alone because general content feedback did not point out the errors to the students. Another important finding was that students who rewrote their essays without receiving feedback improved both in fluency

and content. It can be assumed that rewriting in itself helps improve students' writing. In addition, students who received grammar feedback also improved their content.

The findings of the two previous studies are consistent with a study conducted by Ashwell (2000). He wanted to explore two research questions. First, he examined whether mixing content and form feedback was more beneficial for writers than giving only one type of feedback. Second, he investigated whether teachers should give form feedback alone without any comments on content on the paper. For the sake of his study, Ashwell used 50 students who were enrolled in 2 writing classes. These classes met once a week for an hour and a half. Both classes had the same proficiency level and were taught by the same teacher (in this case the researcher himself). It was their first time to take a college-level writing class. Ashwell's study was a longitudinal study because the progress of students was measured after a one-year period. Each class was asked to write four assignments and they produced three drafts for each assignment. In addition, students had a textbook that helped them with their sentence structures and they were asked to write diaries to help them with their fluency. For this study, Ashwell got his data from the third writing assignment in the semester and its three drafts.

Ashwell's study had a quasi experimental design where he had three treatments (i.e. three types of feedback) and one control group who did not receive feedback on any of their drafts. One group of students received content feedback on their first draft and form feedback on the second. Another group received form feedback first and then content. The last treatment group received both content and form feedback on their drafts.

To control for the extraneous variable that one student might receive more feedback on his essay than others; Ashwell gave a twelve minute time limit in which each teacher wrote feedback on students' papers. After collecting the data, the researcher measured the

formal accuracy (number of errors divided by the number of words) and asked native speakers to measure the content quality.

His findings were that there were no significant differences between the three feedback groups. However, all three feedback groups outperformed the control one in formal accuracy. On the other hand, like Fathman and Whalley (1990), and Rob et al. (1986), Ashwell found that the control group improved like the other groups because the rewriting helped them. In addition, Ashwell pointed out that the group which received both types of feedback on all their drafts slightly improved in their writing than the other ones (but this difference was not statistically significant). Another finding was that mixing both types of feedback did not harm students' writing. This supports Fathman and Whalley's (1990) research findings.

Ashwell, however, pointed that there were some limitations to his study. First, the researcher was the one who gave all feedback. Second, there was no significant inter-rater reliability in the content quality check. Third, the sample size was small (only two classes). Last, scorers were not provided with enough training.

Polio, Fleck and Leder (1998) tried to further explore the difference between feedback and no feedback groups. Polio et al.'s study had conflicting results with Ashwell (2000), Bitchener, Young, & Cameron (2005), Fathman & Whalley (1990), and Ferris & Roberts (2001). Polio et al. examined whether students could edit their grammatical errors if they were given time to revise. This study had four hypotheses that the researchers wanted to test. First, they hypothesized that no differences would be found in students' linguistic accuracy at the end of the semester. Second, they argued that no differences would be found in the linguistic accuracy between the students' revised essays and their original ones. Third, they hypothesized that no differences would be found in the accuracy of students between the beginning and end of the semester. Fourth, they believed that no differences would occur

between groups receiving training in grammar and editing and those who do not in terms of linguistic accuracy.

The participants in Polio et al.'s (1998) study were students of English for academic purposes. The researchers used in-class essays written by the students and their revisions for data collection. For the pretest, participants were asked to write a 30-minute essay at the beginning of the semester. After two days, students were asked to edit their essays in 60 minutes. This same process was repeated in week 15 of the semester to collect data for the posttest.

Error feedback was the treatment used by the researchers. The control group was asked to write journal entries all through the semester without receiving any kind of feedback. The treatment group, on the other hand, wrote journal entries, received feedback, reviewed grammar and was trained to edit texts. To test their hypotheses, the researchers used statistical procedures to analyze their data. Like Kepner's study (1991), the findings of this study showed that there were no significant differences between the accuracy of students who received error feedback and those who did not. In addition, their accuracy did not change significantly from the first week to week 15 (end of the semester).

Other studies had findings that supported those of Polio et al. (1998). According to Truscott, Semke's (1984) study supports the view that error feedback does not have any significant effect on students' writing. In this study, Semke compared four types of feedback: feedback on content in the form of questions and comments and no grammatical feedback, direct feedback on form (teacher marks all the errors and provides the student with their correction), both types of feedback (form and content), and finally coded feedback (use of codes to highlight the error without correcting it). Truscott (1996, 1999) argued that to know the real effect of error feedback, researchers must compare groups of students who do not receive feedback on their essays and those who do. Therefore, Semke's study fits this

criterion because she is comparing between a no grammatical feedback groups with groups who receive different types of feedback.

Her findings showed that grammatical feedback did not help students improve in their written accuracy. However, she suggested that this happened because the content feedback group wrote twice as much as the grammatical feedback group which might have helped improve their writing skills. Both Polio et al. and Semke's research findings came in accordance with Kepner's (1991) which gave more support to the argument that grammar feedback is not essential to improve grammatical accuracy.

As mentioned above, Ferris (1999) argued that Truscott's claims should not be taken for granted because Truscott did not build his argument on a study that he made. Also, she recommended that more studies needed to be done to resolve this debate, because even the previous research was incomparable in terms of design, participants or methodology. Therefore, in 2008, Truscott, together with Hsu, carried out a study to find out whether findings of this study will support Truscott's previous views or not. Truscott and Hsu (2008) claimed that because researchers previously focused on the ability of students to use feedback to edit their first drafts, they found that error feedback had positive effects on grammatical accuracy. According to Truscott and Hsu, this is not a sign of learning. In other words, this kind of progress could not guarantee that students will write grammatically accurate essays in the future. Therefore, in their study they decided to explore whether error correction helps students write better essays in the future.

For their study, the researchers had 47 graduate students who were enrolled in a basic writing class. In this 18 week course, students met once a week with their instructors for three hours. To make sure that the participants have the same proficiency level, they took a diagnostic test. They were asked to write a 40 minute essay. Students who scored between 30 and 42 were included in this study. To collect the data, Truscott and Hsu used two instructors.

They had two groups: one control and the other experimental. The control group had their errors underlined by their teachers. As for the experimental group, they did not receive any kind of feedback on their work. Both groups shared the same course objectives, content, class activities and materials.

Before collecting their data, the researchers allowed the participants 11 weeks to settle in, and get used to the course system. On week twelve, participants were asked to write a 30-minute narrative essay. On week thirteen, teachers returned the essays to the students and gave them 30 minutes to rewrite their first drafts using feedback, if appropriate. A week later, students were asked to write a new narrative to find out if students will improve more than the first narrative. Means, standard deviations, ANOVA and Wilcoxon test were used to analyze the data.

When comparing the first essay with the revision, it was found that the experimental group significantly improved from the first to the second draft. The control group, however, did not show any significant improvement. In spite of their improvement between the first and second draft, the experimental group did not improve in terms of grammatical accuracy on their third writing task (which was on a new topic). Therefore, Truscott and Hsu concluded that researchers could not consider students doing well on their second drafts or revisions as evidence of the effectiveness of error feedback. They need to explore students' performance on new writing tasks after receiving their feedback.

Looking at the above reviewed studies, it is clear that corrective feedback research has not reached a conclusive end. All studies had conflicting results. In the next section, different meta-analyses and reviews are discussed. These reviews are important because they tried to explain the reason of the conflicting research in previous research.

## 2.8 Meta-analysis and Reviews of Literature on Feedback Studies

Based on Truscott's (1996, 1999) controversial papers and Ferris' (1999) response to Truscott, there has been an agreement that more research needs to be done to be able to form a view as to whether error correction is useful to students. Therefore, Ferris found it important to conduct a review of all the studies that looked at the effect and role of error feedback in relation to grammatical accuracy. The findings of her review supported what she and Truscott previously agreed upon: more research is still needed. Ferris' review also emphasized that all previous research findings were incomparable. In other words, they were incomparable in terms of their participants, treatments, or research design. Therefore, researchers should start replicating studies to be able to compare the findings. Ferris stressed the fact that she did not have a strong claim that corrective feedback is effective and all teachers must use it, nor could she prevent teachers from using it by claiming that it is harmful. All that she claimed is that most of the previous research hints that corrective feedback has a positive effect. That is why she argues that more studies are needed to support or refute this finding.

One of the reasons that Truscott claimed in his case against grammar feedback was that there were no studies that compared between groups of students who received feedback and those who did not. He also argued that this comparison was essential because it will determine whether feedback really had a good effect. Again, Ferris agrees with Truscott and she pointed out that only six studies had the above-mentioned comparison. At the end of the review, she sums up the gaps found in the literature done on corrective feedback. She recommended that researchers do more longitudinal studies because most of the previous studies were short-termed. Another recommendation was that researchers replicate the designs of previous studies to be able to compare their findings and reach conclusive results about the effectiveness of error feedback.

Like Ferris (2004), Guenette (2007) believed that all previous research findings are incomparable because researchers used inconsistent designs and methodologies. Guenette, however, reviewed studies using a different perspective. Because she found that previous research on the value of corrective feedback in improving written grammatical accuracy had conflicting results, Guenette did another review to attempt to answer the question whether these conflicting results were due to the different designs, methodologies, and other extraneous variables. In other words, her main purpose was to attribute the conflicting results of the studies, not to the effectiveness or ineffectiveness of feedback, but to the different designs, methodologies and extraneous variables used in these studies.

To answer her question, Guenette reviewed all the research included in the studies of Truscott (1996) and Ferris (2004) to look for reasons of contradictions in results. She found that research designs were one of the factors affecting the results. The researchers used different or incomparable populations, groups (correction vs. no correction groups), and designs (short-term versus long-term). Guenette argued that populations (i.e. students' proficiency levels) were not clearly reported or measured in previous studies. For example, Ashwell (2000) was both the teacher and the researcher. However, he did not clarify how he controlled for proficiency levels. In the conclusion of this section, Guenette stressed that difference in participants' proficiency levels lowers the reliability and validity of the studies' results.

Guenette, then, stated that the significance of comparing correction with no correction groups was to test (support or refute) the effect of feedback (Ferris & Roberts, 2001). This suggested design is the same as the one used in the present study. To test this effect, it is necessary for the researchers to make sure that both control and treatment groups should have the same proficiency levels, instructions context and writing conditions. She also suggested

that more longitudinal studies were needed to measure the improvement of students' overtime (because most short-term studies had positive results).

Moreover, Guenette suggested that more research comparing types of feedback on form with content feedback is needed. In addition, she stressed that all participants should receive the same classroom activities and incentives (grades) to get reliable results. Therefore, in the present study, students in all groups received the same feedback on content and the same type of essays.

Other extraneous variables that could have contributed to the confusing findings of the studies reviewed were highlighted to be considered in future research. Some of these variables were inconsistency and unintelligibility of teacher feedback and its inappropriateness to the students' proficiency levels. She advised the researchers to use similar research designs, similar data collection procedures and to test the effect of feedback over a long period of time. In addition, Guenette stressed that the type of feedback used must be suitable for the students' proficiency levels. Another conclusion was that more descriptive studies that examines and controls for different variables are needed.

Like Guenette (2007), Russell and Spada (2006) carried out a meta-analysis to review previous research on corrective feedback. They argued that there were many extraneous variables that could influence the effectiveness of error feedback like the type of feedback and amount. Previous research suggested that the relationship between feedback and grammatical accuracy is affected by different elements. Due to these findings, Russell and Spada found it necessary to do a meta-analysis to investigate the different factors (e.g. types, the amount and sources of feedback given) that could influence corrective feedback and its effect on grammatical accuracy. In other words, the researchers tried to examine whether corrective feedback helps improve grammatical accuracy. Russell and Spada looked at both

oral and written feedback, but for the purpose of the present study, only their findings on written feedback are discussed.

According to previous research, it was suggested that to be able to have significant results about the effectiveness of feedback, a specific definition of corrective feedback and its measurement is needed (Truscott, 1996; Gunette 2007). Based on this suggestion, Russell and Spada defined corrective feedback as any kind of feedback (oral or written) given to students by their teachers or peers. Their findings supported Guenette's conclusion that corrective feedback is effective for grammatical accuracy and helps students' writing. They also found that it had durable effects (i.e. not short-termed), because the studies had both a pretest and posttest. Nevertheless, Russell and Spada recommended that more research need to be done to support this finding because only five out of fifteen studies that they analyzed had a delayed posttest. In this meta-analysis, they reviewed other studies such as Ashwell (2000), Fazio (2001) and Kepner (1991). In addition, the researchers explored different variables and methodologies that could have affected the effectiveness of corrective feedback in previous research. This again was related to Guenette's review (2007).

At the end of their analysis, Russell and Spada recommended that more studies were needed to compare between the explicit and implicit feedback and on other aspects before any decision is taken on the effectiveness of feedback.

Like Russell and Spada (2006), Truscott (2007) conducted a meta-analysis with the aim of exploring the effect of corrective feedback on learners' accuracy. However, contrary to what Russell and Spada found, Truscott still believed that corrective feedback had no value and added nothing to students' writing.

After looking at the different research findings of the studies reviewed in the meta-analysis, Truscott claimed that error correction is not useful and even if it has any effect on writing, the effect is minimal. He argued that Russell and Spada (2006) found positive effects

because they did not review studies that found negative effects of corrective feedback. Moreover, Truscott (2007) believed that Russell and Spada (2006) included studies that looked at students' ability to revise their essays or take a grammar knowledge test. According to Truscott, these studies do not measure certain learning. In addition, he claimed that no one argues with the fact that neither grammar helps students to revise their essays, nor that students can pass a grammar test.

At the end of the review, Truscott recommended that what is needed to fill gap in research are more meta-analyses and studies that include delayed posttests. Truscott believed that these two things are the only way to test the effect of error feedback.

## **2.9 Conclusion**

A variety of studies related to the different effects that error feedback has on grammatical accuracy were covered in this review of literature. Based on the review, studies on this topic had conflicting findings. Some researchers claimed that this is due to the inconsistent designs and methodologies. They recommended that more studies with comparable designs need to be done (Guenette, 2007). Another important gap in literature was the small number of studies comparing between feedback and no feedback groups. Ever since Truscott stated that there is a need for such a comparison, more studies were done to fill this gap. However, in 2004, Ferris found that only six studies did this comparison (Ashwell, 2000; Fathman & Whalley, 1990; Ferris & Roberts, 2001; Semke 1984; Polio et al., 1998; Kepner, 1991). Like other studies, they had conflicting results which made researchers call for the need of more studies examining this issue. Only one study (Bitchener et al., 2005) further researched this gap.

The present study is an attempt to fill the gap in literature (studies with comparable designs and studies that compare feedback and no feedback groups). To do this, the study

examines whether students who receive feedback on their writing improve in the area of grammatical accuracy more than students who do not receive such feedback.

## Chapter 3

### Methodology and Data

#### 3.1 Proposed Design of Study

##### 3.1.1 Design

This study is a quasi-experimental, applied, exploratory study. In this study, two treatment groups and one control group were used. The two treatment groups received different types of feedback (coded and uncoded), while the control group received no grammar feedback on their essays.

The present study explored the effect of error feedback on the grammatical accuracy of students' writing. This means, that the independent variable (feedback) was manipulated (because three types of feedback were used) to examine any possible change in the grammatical accuracy (dependent variable). Therefore, this study is a quasi-experimental study.

Since the study tries to address writing problems, accuracy and feedback in Egyptian EFL writing classes, hence, this study is applied. Moreover, it is an exploratory study because it did not have any hypotheses. It just explored the effect of the treatment on students' later drafts (in terms of grammatical accuracy). Finally, this is a qualitative study because the data (i.e. student essays) was analyzed verbally, and frequencies of grammatical errors were also used.

##### 3.1.2 Participants

The target sample in this study was freshman students enrolled in an intermediate-level essay course in the Faculty of Arts at Ain Shams University. Their ages ranged between 16 and 18. Three intact writing sections were randomly chosen to serve as a sample for this study. The rationale for using such a sample was that students in these sections usually have their grammatical errors directly pointed out for them. For that reason, it was

important to explore the effects of the different kinds of feedback on their writing and on their grammatical accuracy. Moreover, this is the institution where the researcher works. Therefore, this was a convenient sample.

Students did not take a placement test to determine their proficiency levels because it is not part of the department's policy. This means that the sample had mixed proficiency levels. These non-homogenous groups could have had an impact on the results. To control for proficiency level as an extraneous variable, the study focused on the three most common errors that students did in their essays.

Since students at Ain Shams University are not required to attend, subject attrition was a risk in this study. To control for this attrition, the researcher used three intact sections. Each section included 200 students. The researcher, then, randomly chose 20 students from each section. These students were the ones who attended the pretest, treatment and posttest. This study consisted of one control group and two treatment groups. The study had significant findings which are discussed in full detail below.

### **3.1.3 Data collection procedures**

#### *3.1.3.1 Procedures*

The procedures in this study had the following order:

#### *3.1.3.2 Pretest*

The essays of students in the three groups were collected.

#### *3.1.3.3 Administration of treatment*

Teachers corrected the essays using the three different feedback types. Students in the treatment Group A received their essays with codes on grammar errors and general content feedback. Those in treatment Group B had their errors underlined for them with no reference to the type of error and general content feedback. Students in the control group received general comments on their content without any grammar feedback.

#### *3.1.3.4 Posttest*

Students wrote their second drafts in-class using the feedback on their essays.

### **3.1.4 Pretest and posttest**

#### *3.1.4.1 Pretest*

The pretest was a timed writing test in which students were asked to write an in-class essay in one hour. The same test was administered to both the control group and the two treatment groups. The control group took the pretest to control for the pretest effect. Therefore, the results were due to the treatment not the pretest. This test was used to determine the grammatical errors found in the essay of each individual student.

After administering the treatment, the students received the posttest. Like the pretest, the posttest was a one-hour, in-class essay. Students were asked to edit their essays using the feedback their teachers gave them on their essays.

## **3.2 Treatment**

For the purpose of the study, three groups were used: a control group and two treatment groups: a coded feedback group (treatment Group A) and an uncoded feedback group (treatment Group B). Three types of feedback were used because, according to Ferris (2003), more studies are needed to compare between feedback and no feedback groups. This is related to the research question that compares between the effects of these three types of feedback.

### **3.2.1 Treatment time**

Two of the writing classes: treatment Group A and the Control Group were held on Mondays and Wednesdays (2 hours each) and treatment Group B was held on Tuesdays and Thursdays (2 hours each). Table 1 illustrates the timeframe of the procedures in the control and treatment groups.

Table 1  
*Time Frame of the Study*

<b>Procedure</b>	<b>Treatment Group A</b>	<b>Treatment Group B</b>	<b>Control Group</b>
<b>Collecting first drafts</b>	In class (week 2)	In class (week 2)	In class (week 2)
<b>Training on codes</b>	15 minutes in 2 <sup>nd</sup> class	N/A	N/A
<b>Treatment</b>	Teacher codes grammar errors and content feedback.	Teacher underlines grammar errors and content feedback.	Teacher gives content feedback ONLY.
<b>Writing 2<sup>nd</sup> drafts</b>	One class (week4)	One class (week 4)	One class (week 4)
<b>Collecting 2<sup>nd</sup> drafts</b>	In class	In class	In class
<b>Total time</b>	2weeks	2 weeks	2 weeks

### 3.2.2 Treatment group A

#### 3.2.2.1 Training

This group received feedback on their grammatical errors in the form of codes. Before starting to write essays, the instructor of treatment group introduced the students to the codes and how they can use or interpret them. The training was in the form of an interactive workshop to make sure students knew what exactly they had to do when they see the codes. The teacher gave students sample sentences with coded grammatical errors. Students were, then, asked to identify what the codes referred to. This training was essential because without training students would not have been able to use to improve the quality of their writing.

#### 3.2.2.2 Procedures

First draft (pretest)

In the second week of classes, students were asked to write an essay in one hour. It was an argumentative essay about the difference between traditional and love marriage. The teacher collected all the essays at the end of the class. Teachers were given one-day training on how to code student essays using the codes (see appendix A for the full code list). After the pretest, class teachers collected students' essays to code their grammatical errors.

### In-class editing

At the first class of week 3, students were given back their first drafts with all their grammatical errors labeled with correction codes (such as VT: verb tense error) (see appendix A for the full code list). They also received general content feedback. Using the codes, students were asked to rewrite their essays in an hour. The teacher collected the second drafts at the end of the class. Students were given their second drafts the next class.

### 3.2.3 Treatment group B

#### 3.2.3.1 Procedures

#### First draft

In the second week of classes, students were asked to write an essay in one hour. The same topic was used as for Group A. The teacher collected all the essays at the end of the class.

#### In-class editing

At the first class of week 3, students were given back their first drafts with all their grammatical errors underlined. They also received general content feedback. Using the underlined parts, students were asked to rewrite their essays in an hour. The teacher collected the second drafts at the end of the class. Students were given their second drafts the next class.

It is worth noting that treatment Group A (Coded feedback) and treatment Group B (uncoded feedback) were chosen in this study because they represent two different kinds of cognitive engagements (Ferris & Roberts, 2001; Ferris, 2003). In the coded feedback, the code serves as a clue. In other words, the error is already identified and the student only attempts to correct it. On the other hand, the uncoded feedback has two levels of cognitive engagement. First, the student will identify the error. Second, he /she will try to correct it. This study tests which type of feedback is more effective.

### **3.2.4 Control group**

Students in this group did not receive any grammar feedback. They received their first draft with only content feedback. Students were asked to rewrite their first drafts.

### **3.2.5 All groups**

Based on Guenette's (2007) suggestion that to get accurate results, both control and treatment groups should have the same instruction context and writing conditions. Therefore, all groups in the present study received feedback on content. The instructors gave general content feedback on the students' essays. Content feedback included comments on topic sentences, conclusion, thesis statements and other form-related comments.

### **3.2.6 Some considerations**

One very important ethical problem is that the control group was not exposed to the same teaching materials as the treatment groups (Perry 2005). To avoid this issue, the teacher of the control group was introduced to the two other feedback types. The researcher promised to share the findings of the research (which type was more effective) so that the teacher could use it with the control group in the rest of the semester.

Another important problem to consider was treatment fidelity. To ensure that both treatments were carried out in the proper way, the researcher trained the teachers to spot and code errors.

## **3.3 Observational Techniques**

### **3.3.1 Rating**

Three raters were used as an observational technique. They were asked to look at the students' pretest and posttest essays. Before administering the pretest, each rater was given a list of definitions of 10 grammatical errors, which was adapted from Ferris (2001 & 2003). They were trained for one day to use this list in their rating. The training was expected to help

raters identify these errors in the students' writing. In addition, raters were given an error analysis form (see Appendix B) to fill in the errors that they found in each student's essay.

The three raters first looked at the pretest. Each rater flagged the grammatical errors found in the essay of each student. Then the data of the three raters were compared to verify whether they spotted the same errors or not. The raters, then, highlighted the problematic (grammatical) areas of each individual student. These areas were examined in this study. In other words, the three most common errors were focused on in the study.

The raters repeated the same process with the posttest data. The pretest and posttest data were compared to assess the effect of the treatment on the writing of each individual student. In other words, it tested the effect of the three types of feedback on the grammatical accuracy (the number of grammatical errors per word).

To control for subjectivity, several precautions were taken into consideration (Perry, 2005). First, three raters were used in the study to make sure that they identified the same errors and used the same criteria. In addition, this helped to compare their severity, which will affect the results. The raters were given a coding list of the grammatical errors and used it to calculate the number of the errors in each of the students' writing.

Secondly, the observers were trained to spot the grammatical errors in the students' writing using the coding list of the errors until they were able to identify the same errors. The use of raters is the best way to analyze the data. Otherwise, the researcher would have been the one collecting and analyzing the data, which could have caused the researcher effect.

## Chapter 4

### Data Analysis and Results

#### 4.1 Introduction

The current study examined the effect of feedback on the students' ability to correct their grammatical errors in the revision process. The main focus of the study was to find out which type of feedback would help students overcome their grammatical problems. Essays were written by participants, and frequencies of errors were calculated for data analysis. The results of data analysis are presented in this chapter.

#### 4.2 Data Analysis

The data analysis in this study was done using the Stat 4 U software. Data in the present study was mainly comprised of students' essays in the pretest, and posttest. Grammatical accuracy was measured by calculating the number of errors per essay in each of students' essays. In the following section, data of the pretest and the posttest are analyzed.

##### 4.2.1 Statistical procedures

After obtaining the first and second drafts of treatment groups and the control group, the data was tabulated. The three most common errors: nouns, verbs and articles were focused on to eliminate extraneous variables, such as proficiency levels, effect on the data. Table 2 gives a detailed description of the three errors.

Table 2

*Description of error categories*

Nouns	All errors in plural or possessive endings
Verbs	All errors in verb tenses or verb forms
Articles	All errors in article or determiner usage, omission or unnecessary use.

The study looks at the ability of students to spot any kind error within each error category as a whole. This is replicating the structure of Ferris and Roberts (2001) and Fathman and Whalley (1990) who, for example, gathered all noun errors under the category of nouns and did not deal with each type individually.

The data of each feedback group was analyzed as follows:

First, a series of dependent *t*-tests were performed to compare between the means and standard deviations of frequencies of errors in the pretest and the posttest. This was done to measure students' improvement in the posttest in comparison with pretest. Second, a series of dependent *t*-tests were used to compare the frequencies of the three error types in the pretest and the posttest. Lastly, in order to measure the group which improved the most, an analysis of variance, ANOVA, was used. These data were important to answer the research question. The next section shows the results of the statistical procedures.

## **4.2.2 Results**

### *4.2.2.1 Treatment group A (coded feedback)*

Table 3 displays the number of errors in the pretest and posttest of students. It also includes the descriptive statistics (mean and standard deviation) done on the data. The total number of errors decreased significantly in comparison to those in the pretest. The total number of means decreased from ( $M= 21.50$ ,  $SD = 10.013$ ) to ( $M= 14.05$ ,  $SD = 9.052$ ). However, some students' errors remained the same in both essays (e.g. Student 1), while others' errors increased in the posttest such as Student 17. The majority of students, on the other hand, showed a significant decrease in the numbers of their errors (e.g. Students 2 and 4).

Table 3  
*Total number of errors of coded feedback group*

<b>Total</b>		
<b>Students</b>	<b>1<sup>st</sup> draft</b>	<b>2<sup>nd</sup> draft</b>
1	30	30
2	39	37
3	32	5
4	27	13
5	9	4
6	28	18
7	23	8
8	38	18
9	12	10
10	14	10
11	17	11
12	10	4
13	30	20
14	9	7
15	35	25
16	12	10
17	15	23
18	17	12
19	16	9
20	17	7
<b><i>M</i></b>	21.50	10.01
<b><i>SD</i></b>	14.05	9.052
<b><i>t</i> = 4.365</b>	<b>p = 0.0003</b>	

As it is clear from Table 3, the total number of means decreased from ( $M = 21.50$ ,  $SD = 10.013$ ) to ( $M = 14.05$ ,  $SD = 9.052$ ). However, some students' errors remained the same in both essays (e.g. Student 1), while others' errors increased in the posttest such as Student 17. The majority of students, on the other hand, showed a significant decrease in the numbers of their errors (e.g. Students 2 and 4).

The dependent  $t$ -test results in Table 3 showed that the difference between the two essays was statistically significant with  $t = 4.365$  and  $p = 0.0003$ . This meant that students in the coded feedback group improved in terms of grammatical accuracy.

As for the difference in the number of errors of verbs, nouns, and articles, Tables 4 and 5 display the means, standard deviations of these three types of errors in both the pretest and in treatment Group A.

Table 4  
*Frequencies of different error types*

Students	Nouns		Verbs		Articles	
	1 <sup>st</sup> draft	2 <sup>nd</sup> draft	1 <sup>st</sup> draft	2 <sup>nd</sup> draft	1 <sup>st</sup> draft	2 <sup>nd</sup> draft
1	14	14	9	9	7	7
2	18	12	16	11	5	14
3	18	2	7	1	7	2
4	10	4	9	6	8	3
5	4	1	2	1	3	2
6	18	12	5	5	5	1
7	16	7	3	0	4	1
8	30	14	4	0	4	4
9	8	4	2	3	2	3
10	11	5	3	4	0	1
11	12	8	2	2	3	1
12	3	1	4	1	3	2
13	17	11	8	6	5	3
14	6	4	2	1	1	1
15	26	16	5	4	4	5
16	5	2	3	4	4	3
17	13	7	0	6	2	6
18	10	9	3	2	4	1
19	10	3	3	1	3	1
20	6	3	6	2	5	2

Table 5  
*Means, standard deviations and t-test scores*

	Verbs		Nouns		Articles	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Pretest	4.80	3.62	12.75	7.10	3.95	1.99
Posttest	3.45	3.00	6.95	4.78	3.15	3.10
t-score	2.269		6.016		1.116	
P value	0.031		0.001		0.278	

It is clear from Table 4 above, which shows the frequency of the errors of each student, that students greatly improved in terms of verbs and nouns, but not articles. This finding was supported by the *t*-scores found in Table 5. Students' verb errors decreased

significantly with  $t = 2.269$  and  $p=0.031$ . Similarly, the number of noun endings errors dropped with  $t = 6.016$ ,  $p=0.001$ . However, articles showed no significant difference with  $t=1.116$ ,  $p= 0.278$ .

To sum up, the results for treatment Group A revealed that students improved in grammatical accuracy from the pretest to the posttest. This was supported by the findings of the descriptive statistics and dependent  $t$ -tests which showed statistical significance between the total number of scores and the number of verb and noun errors in both pretest and posttest. In addition, results suggested that students did not show great improvement in article errors.

#### 4.2.2.2 Treatment group B (uncoded feedback group)

Table 6 displays the number of errors in the pretest and posttest of students. It also includes the descriptive statistics (mean and standard deviation) done on the data.

Table 6  
*Total number of errors of un-coded feedback group*

Students	Total	
	1 <sup>st</sup> draft	2 <sup>nd</sup> draft
1	55	21
2	17	7
3	10	4
4	23	16
5	16	11
6	8	7
7	25	22
8	27	26
9	18	12
10	17	14
11	25	14
12	18	14
13	12	7
14	35	11
15	7	3
16	16	8
17	12	3
18	11	15
19	7	0
20	16	10
<b><i>M/ SD</i></b>	18.75/11.22	11.25 / 6.76
<b><i>T=</i></b> 4.060	<b><i>P=</i></b> 0.0007	

Table 6 suggests that the total number of errors in the posttest decreased significantly in comparison to those in the pretest. This was clear in Table 6 because the total number of means decreased from ( $M= 18.75$ ,  $SD = 11.22$ ) to ( $M= 11.25$ ,  $SD = 6.76$ ). However, unlike the previous group: Treatment Group A (coded), none of the students' errors remained the same in both essays nor increased in the posttest. All students showed a significant decrease in the numbers of their errors (e.g. Students 1 and 2).

The dependent  $t$ -test results, found in Table 6, showed that the difference between the two essays for treatment Group B was statistically significant with  $t = 4.060$  and  $p= 0.0007$ . This meant that students in the uncoded feedback group improved in terms of grammatical accuracy.

As for the difference in the number of errors of verbs, nouns, and articles, Tables 7 and 8 display the means, standard deviations of these three types of errors in both the pretest and posttest in treatment Group B (uncoded). It is clear from Table 7 that students greatly improved in terms of verbs and nouns, but not articles.

Table 7  
*Frequencies of different error types*

Students	Nouns		Verbs		Articles	
	1 <sup>st</sup> draft	2 <sup>nd</sup> draft	1 <sup>st</sup> draft	2 <sup>nd</sup> draft	1 <sup>st</sup> draft	2 <sup>nd</sup> draft
1	19	10	22	8	14	3
2	5	1	4	2	8	4
3	8	3	2	1	0	0
4	15	12	5	3	3	1
5	6	5	4	3	6	3
6	5	2	2	4	1	1
7	16	9	4	3	5	10
8	14	21	6	3	7	2
9	10	8	4	3	4	1
10	15	9	2	4	0	1
11	15	8	5	3	5	3
12	13	11	3	1	2	2
13	5	3	6	3	1	1
14	26	6	5	2	4	3
15	5	1	2	1	0	1
16	11	6	3	2	2	0
17	7	0	1	1	4	2
18	7	5	3	5	1	5
19	5	0	0	0	2	0
20	12	0	3	2	1	2

Looking at the number of errors in the pretest and the posttest in Table 7, it is clear that students in this group (uncoded) were able to edit their verb and noun errors but were not as successful in editing their article related errors.

Table 8  
*Means, standard deviations and t-test scores*

	Verbs		Nouns		Articles	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Pretest	4.30	4.46	10.95	5.72	3.50	3.426
Posttest	2.70	1.750	6.0	5.2	2.250	2.268
t-score	2.16		4.277		1.667	
P value	0.0430		0.004		0.112	

This finding was supported by the *t*-scores found in Table 8. Students verb errors decreased significantly with  $t = 2.16$  and  $p=0.0430$ . Similarly, the number of noun errors dropped with  $t = 4.277$ ,  $p=0.004$ . However, articles showed no significant difference with  $t=1.667$ ,  $p= 0.112$ .

To sum up, the results for treatment Group B revealed that students improved in grammatical accuracy from the pretest to the posttest. This was supported by the findings of the descriptive statistics and dependent *t* –tests which showed statistical significance between the total number of scores and the number of verb and noun errors in both pretest and posttest. In addition, results proved that students did not show great improvement in article errors.

#### 4.2.2.3 Control Group (content only feedback)

After collecting all the essays, frequencies of the errors were calculated in both the pretest and posttest. Table 9 displays the number of errors in the pretest and posttest of students. It also includes the descriptive statistics (mean and standard deviation) done on the data.

Table 9

*Total number of errors of control group*

<b>Total</b>		
<b>Students</b>	<b>1<sup>st</sup> draft</b>	<b>2<sup>nd</sup> draft</b>
1	23	16
2	8	6
3	8	8
4	14	11
5	11	13
6	15	9
7	8	11
8	9	20
9	15	8
10	16	15
11	21	18
12	25	17
13	33	28
14	15	20
15	44	37
16	9	6
17	26	24
18	19	23
19	11	6
20	20	20
$t=1.562$	$p=0.1348$	

Table 9 shows that the dependent *t*-test results showed that there was no statistical significance between the pretest and posttest of the control group with  $t= 1.562$ ,  $p = 0.1348$  with 19 degrees of freedom. This meant that students in the control group did not significantly improve in terms of grammatical accuracy between the two essays. This finding was supported by the fact that the total number of means and standard deviations slightly decreased from pretest to posttest. The total number of means decreased from (M= 17.50, SD=9.33) to (M=15.80, SD= 8.23).

Moreover, errors increased in some essays (e.g. Students 8, and 18). However, they decreased in some essays such as those of Students 19 and 12, and remained the same in some (e.g. Student 20). As mentioned above, the present study focused on the three most common errors in students' essays: nouns, verbs, and articles. Table 10 displays the numbers of each error type in both the pretest and posttest in the control group.

Table 10  
*Frequencies of different error types*

Students	Nouns		Verbs		Articles	
	1 <sup>st</sup> draft	2 <sup>nd</sup> draft	1 <sup>st</sup> draft	2nd draft	1 <sup>st</sup> draft	2 <sup>nd</sup> draft
1	16	14	6	1	1	1
2	4	4	3	2	1	0
3	3	3	3	3	2	2
4	8	6	5	4	1	1
5	5	12	4	1	2	1
6	9	7	4	2	2	0
7	4	8	2	2	2	1
8	7	16	0	0	2	4
9	9	5	3	3	3	0
10	11	11	5	4	0	0
11	16	14	10	4	1	1
12	6	6	3	2	1	0
13	13	3	3	3	7	4
14	8	6	5	4	1	1
15	5	12	4	1	2	1
16	9	7	4	2	2	0
17	4	8	2	2	2	1
18	7	16	0	0	2	4
19	9	5	3	3	3	0
20	11	11	5	4	0	0

Table 11  
*Frequencies of different error types*

	Verbs		Nouns		Articles	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Pretest	3.70	2.15	8.20	3.78	1.85	1.46
Posttest	3.75	2.31	8.70	4.22	1.10	1.37
t-value	3.50		-0.463		2.38	
P value	0.24		0.648		0.028	

The calculation of the *t*-scores, mean scores and standard deviations of each error in both essays showed that students improved significantly in the posttest in only one of the errors: articles. As shown in Table 11, the number of errors significantly decreased with  $t=2.38$ ,  $p=0.028$ . As for the verbs and nouns, the change was not statistically significant with a  $t= 3.50$ ,  $p= 0.24$  and  $t= - 0.463$ ,  $p=0.648$  respectively.

To sum up, the results for the control group revealed that students did not improve significantly in grammatical accuracy from the pretest to the posttest. This was supported by

the findings of the descriptive statistics and dependent t-tests which showed no statistical significance between the total number of scores and the number of verb and noun errors in both pretest and posttest. In addition, unlike the two treatment groups, results proved that Control Group students were able to improve their article errors.

### 4.3 Analysis of Variance

In order to answer the research questions which explore the type of feedback that influences grammatical accuracy the most, ANOVA was used to compare between the posttest data of the three groups. The results in Table 12 showed that there were no significant differences between the two treatment groups, the coded and uncoded feedback groups. However, both groups significantly outperformed the control group in the total number of errors, verbs and nouns. The control group, on the other hand, significantly outperformed the treatment groups in correcting their article errors. In the following chapter, the findings are discussed.

Table 12  
*Analysis of Variance*

<b>Total Number of errors in Second Drafts</b>			
<b>Source</b>	<b>Df</b>	<b>F</b>	<b>P</b>
Group 1- Group 2	2.80	1.62	0.519
Group 1- Group 3	1.750		0.077
Group 2- Group 3	2.07		0.018
<b>Number of Verb Errors in Second Drafts</b>			
Group 1- Group 2	0.750	1.38	0.518
Group 1- Group 3	1.100		0.023
Group 2- Group 3	0.350		0.081
<b>Number of Noun Errors in Second Drafts</b>			
Group 1- Group 2	0.950	1.65	0.80
Group 1- Group 3	1.750		0.04
Group 2- Group 3	2.10		0.018
<b>Number of Article Errors in Second Drafts</b>			
Group 1- Group 2	0.900	3.81	0.45
Group 1- Group 3	2.050		0.0212
Group 2- Group 3	1.15		0.027

#### 4.4 Fluency

It is clear from the results mentioned above that when given feedback students were able to edit their grammatical errors better than when they were only given content feedback. However, it is worth noting that students in the Control Group significantly improved the content of their essays. They wrote longer essays and improved the quality of their writing (e.g. topic sentences, word choice and organization of their essays.). Table 13 shows the mean number of words of the three groups.

Table 13  
*Mean number of words in original essays and rewrites*

	Original	Rewrite	Increase
Control	281	340	59
Treatment A (Coded)	220	242	22
Treatment B (Uncoded)	235	245	10

As it is clear from Table 13, the type of feedback given to students has an effect on the number of words they wrote. The uncoded feedback showed the least improvement in fluency, whereas the control group showed the most improvement. According to Fathman and Whalley (1990), some researchers argued that grammar feedback “inhibits” the fluency of the students. Students in the present study did write longer essays when they received no grammar feedback than the two other groups who received grammar feedback. This might suggest that when students receive grammar feedback on their work, they do not focus on the content as when they do not. This is also supported by the fact that the uncoded feedback group wrote the least number of words. Since the uncoded feedback needs more thinking skills (Ferris, 2003), students focused more on finding out their errors and correcting them than on their content.

Participants in the Control Group succeeded to improve the content of their essays in comparison to the two Treatment Groups. This improvement was in the form of adding topic

sentences which were missing in the first drafts and/or organizing ideas in a better way. Table 14 displays the percentage of students who improved their content (i.e. added topic sentences in second drafts and organized ideas) in all three groups.

Table 14  
*Percentage of students changing content in the second draft*

Scores Description	Control Group		Coded Group		Uncoded Group	
	Topic Sent.	Organization	Topic sent.	Organization	Topic sent.	Organization
Higher score	80	70	20	10	25	10
Same score	10	30	40	57	30	40
Lower score	10	0	40	33	45	50

Table 14 shows the percentage of students who improved their content scores, those who had the same score, and those who scored lower than the first drafts. The two content errors that were focused on are topic sentences and organization of ideas. Students who scored higher on topic sentences are the ones who added topic sentences on their second drafts. Those who scored the same are the students who did not make any changes in the number of content errors. As for the group who scored lower on content than the first draft, they are the students who made more content errors on the second draft than on the first draft. It is clear from Table 14 that the Control group students significantly improved in terms of content in comparison to the two Treatment Groups.

Detailed discussion of the results reported above and its relationship to previous research are presented in chapter five.

## Chapter 5

### Discussion of Results

#### 5.1 Summary of the Results

The research questions explored the outcome of using different feedback types on students' ability to write grammatically accurate essays. Findings showed that there are significant differences between the two treatment groups and the control group. On the other hand, there was not any mentionable difference between treatment Group A receiving coded feedback and treatment Group B receiving uncoded feedback. Another interesting finding was that when asked to rewrite their essays, both treatment groups were able to edit most of their errors related to verbs and nouns. However, they were not able to successfully edit errors related to articles. Students in the control group, on the other hand, were able to locate and successfully correct their article errors but not their verb and noun errors.

#### 5.2 Discussion of the Results

The findings suggest that there was not a significant difference between using codes for feedback and underlining the errors. This substantiates Ferris & Roberts (2001), and Robb et al.'s (1986) findings regarding the fact that there were no significant differences between coded and uncoded feedback types. In other words, both types of feedback affect students' writing similarly. This finding could mean that if teachers want students to use their feedback to produce grammatically accurate essays, it may be enough to underline the errors without writing any codes which will save the time of teachers. According to Truscott (1996), one of the problems with grammar feedback is that teachers may make mistakes when correcting students. This could be avoided based on the findings of the present study.

Although the finding that uncoded feedback has a similar effect to the coded feedback suggest that teachers should use uncoded feedback, one should keep in mind that the present study was short-termed. To be sure that uncoded feedback is the best strategy for helping

students produce grammatically accurate essays; a long term study is needed to make sure that it helps students improve in accuracy over a long period of time (Lalande, 1982).

Another interesting finding was that the Control Group students were not able to locate and correct all their grammatical errors. This finding was similar to what Bitchener, Young, & Cameron (2005), Fathman and Whalley (1990), and Ferris and Roberts (2001). Like these studies, students in the no grammar feedback group edited some of their grammatical errors but they were outperformed by the treatment groups who received grammar feedback. It is worth noting, however, that the treatment groups were successful in editing the noun and verb errors (Table 4 and 7) which Ferris and Roberts (2001) called the treatable or easy to correct errors; but unlike Ferris and Roberts (2001) they were not able to correct their article errors. Surprisingly, the control group significantly improved in their article usage (Table 10). A reason for this finding could be that the instructor of the Control Group explained the usage of articles before the rewrite and this did not happen with any of the experimental groups.

To sum up, answers to the research questions could help teachers decide on the effectiveness of teacher error feedback:

- i. Does the use of coded feedback in academic writing instruction have an effect on grammatical accuracy in the revision process?

The use of codes to locate errors for the students did help students significantly improve their noun and verb errors.

- ii. Does the use of uncoded feedback in academic writing instruction have an effect on grammatical accuracy in the revision process?

Underlining errors in students' texts helped students show improvement in editing their verb and noun errors. There were no significant differences between the two Treatment groups mentioned above.

iii. Do students receiving no grammar feedback on their academic writing improve in grammatical accuracy in the revision process?

When students received general commentary on the content of their essay, they improved in terms of fluency. They were also able to improve the quality of their writing. For example, some students added topic sentences which were missing in their first drafts. Others managed to organize their ideas better. In addition, students in no grammar feedback group were able to correct their errors related to article usage. This finding was surprising because according to Second Language Acquisition development theories, articles are acquired after nouns and verbs. Therefore, it was surprising that these students were able to edit article errors which are global errors (Ferris, 2003) and not verbs and nouns which are local errors (Ferris & Roberts, 2001; Ferris, 2003). One possible reason for these results could be that the Control Group was at a higher proficiency level than the two Treatment Groups A and B. Another possible reason could be that participants in this group were focused on content and not grammar. Therefore, they focused on errors that affected the content or ideas and not on local errors. This possibility supports the findings of other researchers such as Fathman and Whalley (1990). To sum up, the rewriting process proved to be very beneficial for the students even if they do not receive grammar feedback from their teacher.

In the next section, some extraneous variables that might have affected the results of the study are discussed.

### **5.3 Some Extraneous Factors**

#### **5.3.1 Proficiency level**

One possible reason that the Treatment Groups outperformed the Control Group is the proficiency level. According to Truscott (1999) and Ferris (2004), students with low proficiency level will find it very hard to locate and correct their errors without help from their teachers. In the present study, participants were freshman students at a public university

who do not have a strong language background. This could have affected the performance of the Control Group. They could have found it hard or frustrating to receive content feedback and no grammar feedback.

### **5.3.2 Control group**

As discussed above, the control group wrote more words in their rewrites. This increase in word numbers could have revealed more grammatical problems in students' essays.

Another interesting point is that students in this group improved the quality of their sentences and ideas. They tried to develop their ideas probably because they had to read the whole essay and correct their errors (for details please refer to section 4.4 in chapter 4).

### **5.3.3 Motivation**

Students who were enrolled in the essay classes did not receive any grades on their drafts. Therefore, according to their teachers, students did not show willingness to edit or redraft their essays. The insignificant results could have been because of the students' reluctance to correct or write proper essays.

### **5.3.4 H1N1 scare**

Medical conditions were an issue affecting the attitude of students in all three groups and their motivation. Students were reluctant to attend classes and participate in class activities due to the H1N1 scare. This affected the design of the study because the researcher intended to do a delayed posttest to measure whether grammar feedback had a long-term effect. However, out of the 60 participants, only 15 attended the delayed test. The findings of the delayed posttest were very interesting and significant but were not added in the study due to the small number of participants.

### **5.3.5 Relation to previous research**

Few studies in the literature examined the difference between feedback and no feedback groups and their effect on grammatical accuracy. In other words, previous research mainly studied the effect of different feedback groups. The findings of the present study supports those of Ashwell 2000, Bitchener, Young, & Cameron 2005, Fathman & Whalley (1990), and Ferris & Roberts 2001. Similar to the findings of the present study, they found no advantage of coded feedback over the uncoded one. Therefore, teachers could start giving students clues of their errors and students could self-edit their papers.

This study, however, refutes the findings of studies such as Semke 1984; Polio et al., 1998; Kepner, 1991, and Lalande, 1982. These studies found that error feedback does not have any significant effect on students' writing. However, one should keep in mind that these findings could have been due to differences in participants, designs, or period (long term vs. short term), according to Guenette (2004). Limitations of the study are discussed in the following section.

## **5.4 Limitations of the Study**

### **5.4.1 Second drafts as a measure of learning**

Truscott and Hsu (2008) argued that although students who receive feedback on their essays are usually successful in editing their errors on the second drafts, this improvement does not mean that students have learned and will not commit errors in the future. Therefore, they recommended that all research should include a section where students are asked to write on a new topic other than the one they received feedback on. Unfortunately, in the present study, the researcher planned to have a delayed posttest. However, there was subject attrition on that day, due to H1N1 precautions which is discussed later in this section. This resulted in having three to five participants in each group out of 20. Therefore, the researcher

had to discard the findings because they were not of any significance. Having a delayed posttest would have added more value to the finding of the study.

#### **5.4.2 Timing of the treatment**

The treatment was carried out on the second week of classes which could have affected the students' performance in the study. In other words, the students did not have time to get used to the system in the class, especially that they were freshmen. Therefore, researchers need to keep in mind the time when they introduce their treatment.

#### **5.4.3 Duration of the study**

The duration of this study was very short: three weeks. This could have affected the findings. Therefore, it is recommended that a longitudinal study should be carried out to explore whether the treatments have the same effects they did on the long run.

#### **5.4.4 Proficiency level**

Since students at Ain Shams University do not sit for a placement test, students participating in the study were of different proficiency levels. Although the researcher tried to control for this by choosing the three most common errors found in students' essays, it could have been an extraneous variable affecting the results.

#### **5.4.5 Implications for future research**

The findings of this study need further research to support or refute them. There is a need to have more studies comparing between feedback and no feedback groups. However, researchers should make the treatment period longer to measure long term effects of feedback types. Another thing that needs to be included in future studies is other types of errors such as run-ons, fragments, and word choice. This would be important to find out whether there are untreatable errors (Ferris and Roberts, 2001).

Moreover, in the present study, the different types of noun, verb, and article errors were treated as one category. It would be useful for future research to check the effect of

feedback on each type of error under each category. For example, one could explore the effect of feedback on noun endings, possessive and noun form errors individually to find out if it had a different effect on each type.

To sum up, teacher feedback will always be an important topic for both teachers and students. Therefore researchers still need to investigate different feedback strategies to help students and teachers. The present study is a short termed and experimental study that has limitations, but it highlighted the possibility that some feedback strategies work better than others. However, it suggests that more research still needs to be done.

## Appendix A

### Coding System

To be able to identify the grammatical errors in the students' writing, raters will be given a list of 10 grammatical errors, their definitions and examples. This list will facilitate the process of spotting the grammatical errors in the essays.

Grammatical error	Definition	Example
Verb tense	It includes missing or wrong tense markers and the modals, which mark tense such as, can/could.	I <b>attend</b> the concert last night. (attended)
Verb form	Any kind of error in the formation of the verb phrase (not including time or tense marking) (e.g. infinitives, and ill formed passives)	They hope <b>find</b> happiness. (to find)
Subject verb agreement	It is an error in noun or verb form leading to lack of agreement in number.	Building houses <b>are</b> tiring. (is)
Articles/determiners	It includes unnecessary, wrong, or missing articles or determiners.	Put <b>book on a table</b> . (The book on the table).
Noun endings	Includes missing or ill-formed plural or possessive markers.	<b>Magician</b> are tricky people (magicians)
Pronouns	Pronouns that do not agree in number or case with its referent or that have no antecedent. (only personal pronouns)	We should let <b>he</b> know before it is too late. (him)
Fragments	It is either a dependent clause used as a sentence or clauses with no subjects.	She said I hated the game. <b>Which is not true.</b>
Run-ons	The use of run-on sentences or comma splices.	In Egypt people are afraid to speak about their personal life, people from Africa like to express their feelings.
Word choice	Includes the use of wrong verb, modal, preposition and relative pronoun.	In addition <b>of</b> money...(to)
Miscellaneous	Other errors not mentioned in the previous categories	

Appendix A is adapted from Ferris (2003).

## Appendix B

### Student's Error Analysis Form

<b>Error type</b>	<b>Number of errors</b>	<b>Ranking *</b>
Subject verb agreement		
Noun endings		
Article/determiner		
Verb tense		
Verb form		
Fragments		
Run-ons		
Pronouns		
Word choice		
Miscellaneous		

\* The rater will give the number 1 to the most frequent error and 10 to the least frequent.

Appendix B is adapted from Ferris (2003).

## References

- Ashwell, T. (2000). Patterns of teacher response to student writing in a multiple-draft composition classroom: Is content feedback followed by form feedback the best method? *Journal of Second Language Writing* 9, 227-258. Retrieved December 4, 2006 from Science Direct.
- Bitchener, J., Young, S., & Cameron, D. (2005). The effect of different types of feedback for improvement in the accuracy and fluency of L2 student writing. *Journal of Second Language Writing* 14, 191-205.
- Chandler, J. (2003). The efficacy of various kinds of error feedback for improvement in the accuracy and fluency of L2 student writing. *Journal of Second Language Writing* 12 (3), 267-296. Retrieved December 4, 2006 from Science Direct.
- Chandler, J. (2004). A response to Truscott. *Journal of Second Language Writing* 13, 345-348.
- Ellis, R., Sheen, Y., Murakami, M., & Takashima, H. (2008). The effects of focused and unfocused written corrective feedback in an English as a foreign language context. *System* 36, 353-371.
- Fathman, A. & Whalley, E. (1990). Teacher response to student writing: Focus on form versus content. In B. Kroll, *Second language writing: research insights for the classroom* (pp.178-190). Cambridge: Cambridge University Press.
- Fazio, L. (2001). The effect of corrections and commentaries on the journal writing accuracy of minority- and majority-language students. *Journal of Second Language Writing* 10, 235-249.
- Ferris, D. R. (1995 a). Can advanced ESL students be taught to correct their most serious and frequent errors? *CATESOL Journal* 8, 41-62.

- Ferris, D.R. (1995 b). Student reactions to teacher response in multiple-draft composition classrooms. *TESOL Quarterly* 29, 33–53.
- Ferris, D.R. (1995 c). Teaching ESL composition students to become independent self-editors. *TESOL Journal* 4, 18–22.
- Ferris, D. R. (1997). The influence of teacher commentary on student revision. *TESOL Quarterly* 31 (2), 315-339.
- Ferris, D. R. (1999). The case for grammar correction in L2 writing classes: A response to Truscott (1996). *Journal of Second Language Writing* 8 (1), 1-11.
- Ferris, D. & Roberts, B. (2001). Error feedback in L2 writing: How explicit does it need to be? *Journal of Second Language Writing* 10 (3), 161-184. Retrieved December 4, 2006 from Science Direct.
- Ferris, D. R. (2003). *Response to student writing: Implications for second language students*. Mahwah, NJ: Lawrence Erlbaum.
- Ferris, D. R. (2004). The grammar correction debate in L2 writing: Where are we, where do we go from here? *Journal of Second Language Writing* 13 (1), 49-62. Retrieved December 2, 2006 from Science Direct.
- Frantzen, D. (1995). The effects of grammar supplementation on written accuracy in an intermediate Spanish content course. *The Modern Language Journal* 79, 329-344.
- Gascoigne, C. (2004). Examining the effect of feedback in beginning L2 composition. *Foreign Language Annals* 37 (1), 71-76.
- Guenette, D. (2007). Is feedback pedagogically correct? Research design issues in studies of feedback on writing. *Journal of Second Language Writing* 16, 40-53.
- Hedgcock, J., & Lefkowitz, N. (1996). Some input on input: Two analyses of student response to expert feedback in L2 writing. *The Modern Language Journal* 80 (3), 287-308.

- Hendrickson, J.M. (1978). Error correction in foreign language teaching: Recent theory, research and practice. *The Modern Language Journal* 62 (8), 387-398.
- Hendrickson, J.M. (1980). The treatment of error in written work. *The Modern Language Journal* 64 (2), 216-221.
- Hyland, F. (2000). ESL writers and feedback: Giving more autonomy to students. *Language Teaching Research* 4 (1), 33-54.
- Hyland, F. (2003). Focusing on form: Student engagement with teacher feedback. *System* 31 (2), 217-230.
- Kepner, C. G. (1991). An experiment in the relationship of types of written feedback to the development of second-language writing skills. *Modern Language Journal* 75 (3), 305-313. Retrieved January 16, 2007 from JSTOR.
- Lalande II, J. F. (1982). Reducing composition errors: An experiment. *Modern Language Journal* 66 (2), 140-149.
- Lee, I. (2004). Error correction in L2 secondary writing classrooms: The case of Hong Kong. *Journal of Second Language Writing* 13 (4), 285-312.
- Leki, I. (1990). Coaching from the margins: Issues in written response. In B. Kroll, *Second language writing: research insights for the classroom* (pp.57-68). Cambridge: Cambridge University Press.
- Leki, I. (1991). The preferences of ESL students for error correction in college-level writing classes. *Foreign Language Annals* 24, 203-218.
- O'Leary, Z. (2004). *Essential guide to doing research*. London: Sage Publications, Incorporated.
- Perry, F. L. (2005). *Research in applied linguistics: Becoming a discerning consumer*. Mahwah, NJ: Lawrence Erlbaum.

- Polio, C. (1997). Measures of linguistic accuracy in second language writing research. *Language Learning* 47, 101–143.
- Polio, C., Fleck, C., & Leder, N. (1998). If only I had more time: ESL learners' changes in linguistic accuracy on essay revisions. *Journal of Second Language Writing* 7, 43-68. Retrieved January 8, 2007 from Science Direct.
- Robb, T., Ross, S., & Shortreed, I. (1986). Salience of feedback on error correction and its effect on EFL writing quality. *TESOL Quarterly* 20, 83-93. Retrieved February 7, 2007 from JSTOR.
- Semke, H. (1984). The effects of the red pen. *Foreign Language Annals* 17 (3), 195-202.
- Russell, J., & Spada, N. (2006). The effectiveness of corrective feedback for the acquisition of L2 grammar: A meta-analysis of the research. In J. Norris & L. Ortega, *Synthesizing research on language and teaching* (pp. 133- 164). Amsterdam: John Benjamins Publishing Company.
- Truscott, J. (1996). The case against grammar correction in L2 writing classes. *Language Learning* 13 (4), 337-343. Retrieved December 8, 2006 from ERIC.
- Truscott, J. (1999). The case for "the case against grammar correction in L2 writing classes": A response to Ferris. *Journal of Second Language Writing* 8 (2), 111-122.
- Truscott, J. (2007). The effect of error correction on learners' ability to write accurately. *Journal of Second Language Writing* 16, 255-272.
- Truscott, J., & Hsu, A. Y. (2008). Error correction, revision, and learning. *Journal of Second Language Writing* 17, 292-305.
- Yates, R., & Kenkel, J. (2002). Responding to sentence level errors in writing. *Journal of Second Language Writing* 11, 29-47.
- Zamel, V. (1985). Responding to student writing. *TESOL Quarterly* 19, 79–101