

**THE EFFECT OF DIRECT AND INDIRECT WRITTEN CORRECTIVE FEEDBACK
ON THE ACQUISITION OF RULE-BASED AND ITEM-BASED LINGUISTIC
FEATURES**

Fatimah Idris Alkhawajah

**A Dissertation submitted to the Faculty of Graduate Studies in Partial Fulfillment of the
Requirements for the Degree of Doctor of Philosophy**

**Graduate Program in Linguistic and Applied Linguistics
York University
Toronto, Ontario**

August, 2016

© Fatimah Alkhawajah, 2016

Abstract

One of the most important tasks for second language (L2) writing teachers is providing their students with corrective feedback (CF) on their writing (Hyland & Hyland, 2001). Teachers and students agree that written CF is “both desirable and helpful” (Goldstein, 2004). Recent research suggests that written CF has a positive effect on the acquisition of certain linguistic features. However, many questions related to written CF remain in need of further investigation including questions related to the impact of CF types on linguistic features.

The current study investigated whether there exists a differential effect of direct and indirect CF on the acquisition of rule-based features (simple present) and item-based features (prepositions). It also examined one of the possible factors impacting the effectiveness of written CF, which is students’ preferences for CF types. Fifty students enrolled in an EFL writing class were divided into four groups. Each group received one of the following treatments: direct CF on simple present, indirect CF on simple present, direct CF on prepositions, or indirect CF on prepositions over three sessions. In this pretest/immediate posttest/delayed posttest design, students received written CF, revised writing tasks, and completed new tasks and tests. They also responded to a questionnaire exploring their preferences for different CF practices.

Students’ accuracy scores in the writing tests were subjected to three-way mixed ANOVA to investigate whether there was a differential effect between direct and indirect CF on the target features. The questionnaire data were analyzed to examine the impact of students’ preferences on accuracy in writing. Results showed that simple present, a type of rule-based feature, responded better to indirect CF while prepositions, a type of item-based feature, responded better to direct CF. The analysis of the questionnaire data revealed that the students had different preferences in terms of how CF should be implemented, but no relationship was

found between students' preferences of direct and indirect CF and their performance in writing. The findings suggest that teachers should consider addressing different types of linguistic features through different types of CF and that teachers engage their students in discussion about the effectiveness of different CF types.

Dedication

To Ali, Alhawra, and Alzahra

Without you, this... could not be achieved!

Acknowledgments

I would like to acknowledge and express my heartfelt gratitude to my thesis supervisor Dr. Antonella Valeo for the continuous support for my work throughout my PhD journey. Her guidance, suggestions for better ideas, and valuable feedback have improved my research greatly. I could not have imagined having a better supervisor and mentor for my PhD study, and if I had to do this all over again, I would choose her all over again.

Besides my supervisor, I would like to thank the rest of my thesis committee, Dr. Linda Steinman and Dr. Khaled Barkaoui for their support and insightful comments starting from the proposal stage to the last versions of the thesis. Their valuable suggestions have improved my work and made the completion of this project possible.

Last but not least, I would like to thank my family starting with my parents, Idrees and Nawal, for their support and continuous prayers for my success. I would also like to thank my small family: my husband, Ali, for his support, patience and encouragement, and my two beautiful girls, Alhawra and Alzahra, for just being in my life!

Table of Contents

Abstract.....	ii
Dedication	iv
Acknowledgments	v
List of Tables.....	viii
List of Figures	ix
Chapter One: Introduction	1
Rationale	1
Research Questions	5
Thesis Organization	5
Chapter Two: Literature Review.....	7
Theoretical Frameworks for CF.....	7
SLA Theories	7
Effect of CF in language learning.	7
Effect of the type of the target features in language learning.....	11
Writing Theories	13
Empirical Research	19
Research on Direct CF.....	21
Research on Indirect CF.....	27
Research on the Differential Effect of Direct and Indirect CF.....	31
Research on how Different Types of Linguistic Features Respond to Different Types of CF ...	36
Research on Students' Preferences for Different CF Types	39
Implications for the Present Study	44
Chapter Three: Methodology	46
Context	46
Participants	48
Target Structures	50
What Constitutes an Accurate Use of the Simple Present Tense?	51
What Constitutes an Accurate Use of Prepositions?	52
CF Types	52
Operationalizations	54
Direct written CF.	54
Indirect written CF.	54
Design	54
Procedure for Data Collection.....	56
Tools and Instruments.....	59
Pilot Study.....	61
Participants.....	62
Procedure for Data Collection.....	63
Results of Piloting the Design and the Instruments.....	64
Instruments.....	64
Design.....	68
Procedure for Data Scoring and Coding	70

Scoring Present Simple Tense Tests.....	71
Scoring Prepositions Tests	73
Scoring Reliability	73
Procedure for Data Analysis	77
Chapter Four: Data Analysis and Results	79
Groups' Performance at Baseline	80
Group Differences at Baseline	80
Pretest 1.....	80
Pretest 2. S	82
Pretests' Differences.....	83
Groups' Performance after Treatment	85
Descriptive Statistics	85
Statistical Assumptions	87
Inferential Statistics	89
Students' Preferences for Different CF Practices	95
Students' Preferences.....	95
Relationship between Students' Preferences of Direct CF or Indirect CF and Accuracy	103
Descriptive statistics.....	103
Statistical assumptions.....	104
Inferential statistics.....	104
Summary of the Results	106
Chapter Five: Discussion and Conclusions	108
Effect of CF on Accuracy	109
Effect of Different Types of CF on the Acquisition of Different Types of Linguistic Features ...	112
Students' Preferences for Different CF Types	120
Limitations	126
Conclusions and Implications.....	129
Directions for Future Research	132
References	135

List of Tables

Table 1: Study Groups	55
Table 2: Design and Timeline	59
Table 3: Descriptive Statistics for Pretest 1	81
Table 4 : One-way ANOVA analysis for Pretest 1	82
Table 5: Descriptive Statistics for Pretest 2	82
Table 6: One-way ANOVA Analysis for Pretest 2	83
Table 7: Paired-Samples T-Test Descriptive Statistics	84
Table 8: Paired Samples T-Test Analysis	84
Table 9: Descriptive Statistics for Accuracy Scores by Treatment Groups	86
Table 10: Levene’s Test of Equality of Error Variances	88
Table 11: Three-way Mixed ANOVA: Treatment and Accuracy [Tests of Within-Subjects Effects]	90
Table 12: Pairwise Comparisons for the Treatment Groups	91
Table 13: Pairwise Comparisons for Three Time Points: Pretest 2, Immediate, and Delayed Posttests	94
Table 14: Three-way Mixed ANOVA: Treatment and Accuracy [Tests of Between-Subjects Effects]	95
Table 15: Descriptive Statistics for the Students’ Responses to Questionnaire 2	96
Table 16: Descriptive Statistics for Accuracy Scores by Preference Group	104
Table 17: Two-way Mixed ANOVA: CF Preferences and Accuracy [Tests of Within-Subjects Effects]	105
Table 18: Two-way Mixed ANOVA: CF Preferences and Accuracy [Tests of Between-Subjects Effects]	106
Table 19: Summary of the Results	107

List of Figures

Figure 1: Profile Plots for the Four Treatment Groups.....	93
---	-----------

Chapter One: Introduction

Rationale

Researchers argue that there are two types of evidence required for language learning: positive evidence and negative evidence (e.g. Gass, 2003; Long, 1996). Positive evidence is the spoken or written L2 input to which the learner is exposed (Gass, 2003). It is the first and most important condition for language learning to take place. Negative evidence, on the other hand, explains what is not acceptable in L2 (Long, 1996). It can be preemptive, occurring before errors are made, such as in a planned grammar lesson, or reactive, occurring as a result of language errors made by the learner (Long & Robinson, 1998). The reactive negative evidence is what is commonly known as corrective feedback (CF). For the purpose of the current study, CF is defined as “a teacher’s reactive move that invites learners to attend to the grammatical accuracy of something they have... written” (Sheen, 2011, p. 1). CF is the teacher’s response to an incorrect linguistic form used by the learner in writing.

Research on CF has produced different results in terms of its effectiveness. While the early research on CF produced either negative or inconclusive results in regard to the efficacy of CF (Kepner, 1991; Lalande, 1982; Polio, Fleck & Leder, 1998; Semke, 1984), the more recent research has suggested that written CF has a positive effect on the acquisition of certain linguistic features (e.g. Bitchener & Knoch, 2010a, 2010b; Ellis, Sheen, Murakami, & Takashima, 2008; Sheen, 2007; Sheen, Wright, & Moldawa, 2009). However, many questions related to written CF are still in need of further investigation. One of these questions is the effect of different types of CF on the acquisition of different types of linguistic features, which is an important question for reasons related to theory, pedagogy, and the current state of the research on CF. First, most of the current research on CF has either investigated the effect of one CF

approach, direct or indirect, on the acquisition of different types of linguistic features, or it has investigated the effect of both CF approaches, direct and indirect, on the acquisition of one type of linguistic feature, usually rule-based features. Very limited research, almost none to the best of my knowledge, has examined the effect of both direct and indirect CF on the acquisition of both rule-based and item-based linguistic features.

However, there are some studies that are closely related to this specific question.

Bitchener, Young, and Cameron (2005) investigated the effect of one approach, direct CF, on the acquisition of three linguistic features, two of them were rule-based and the third one was item-based. Ferris (2006) examined how writing teachers provided CF to their students, what types of CF they used, direct or indirect, and with what type of linguistic features, rule-based versus item-based features. Also, in a study by Van Beuningen, Jong, and Kuiken (2012), a separate analysis was conducted on the students' writing to determine how grammatical and non-grammatical errors respond to different types of CF that targeted all error types. Aside from these studies, there is not much research investigating the relationship between type of CF and type of linguistic feature treated. Therefore, little research has examined the differential effect of direct and indirect CF on the acquisition of different types of linguistic features, rule-based and item-based features.

The effect of CF for learners in an EFL context (i.e. English as a Foreign Language) is another important area that requires more investigation. Most CF studies have been conducted in ESL contexts (i.e. English as a Second Language). How much of the ESL research is applicable in the EFL context is unclear. There are numerous differences between the two contexts that may impact the effect of CF on L2/FL acquisition. One of these differences is the possible daily use of the language offered by the ESL environment in contrast to the EFL environment in which

English is not used on a daily basis. This difference suggests that there might be some factors helping CF in the ESL context to have an effect, such as the rich opportunity to practice L2 and possibly the opportunity to receive corrections from others, in comparison to the EFL context which does not provide much opportunity for the learners to practice the language. Investigating the effect of written CF on language improvement in a country where English is not much used is an important area of research as EFL is widely taught all over the world.

Moreover, the present study was situated in a nursing bachelor program in Saudi Arabia which adds to the significance of the context of the study. English has a significant importance in the medical educational programs and the medical profession in Saudi Arabia as it is the medium of instruction and communication in these places. English is the language of use in college assignments and medical reports, and accuracy is considered essential and expected in the medical field. It is important, therefore, to prepare the students in medical colleges to write accurately for their college assignments and their future career.

Second, the question of how different types of CF affect language learning and writing development is an important pedagogical question. Providing direct and indirect CF on different linguistic forms in students' writing is common practice in language classrooms. Research suggests that both types of CF are useful when certain conditions are met. One of these conditions is targeting the linguistic features that can respond to CF instead of wasting time and energy on features that might be not amenable to CF. Investigating how the learning of a certain type of linguistic feature might be affected through certain types of CF might provide teachers with answers and tips on how to provide more effective CF on their students' writing.

Third, there are many theoretical arguments on how direct and indirect CF may affect language learning. For example, it is argued that direct CF triggers 'noticing' easily because of

its explicitness (Ellis, Loewen, & Erlam, 2006). Noticing is one of the conditions required for language learning to take place (Schmidt, 1990, 1994, 1995). On the other hand, it is argued that indirect CF promotes guided and self-learning (Chandler, 2003; Lalande, 1982). To what extent these theoretical arguments are accurate when dealing with not only treatable errors but also untreatable errors is an area in need of further investigation. Previous research has provided support for the positive effect of direct and indirect CF on treatable errors (e.g. Bitchener et al., 2005; Sheen et al., 2009); but future research still needs to provide evidence for, or against, the benefit of direct and indirect CF on untreatable errors. If research cannot present a case for CF being effective for the acquisition of complex, untreatable, or less treatable linguistic features, many of the arguments against CF may be still valid.

An additional question related to the effect of CF is concerned with the variables that may have an impact on the degree to which some students benefit from CF while other students do not appear to benefit at all. One of these possible variables is individual differences. Students have different characteristics that may affect their learning outcomes, such as motivation, aptitude, attitude, learning styles, and learning preferences. Schmidt (2010) argues that the characteristics of the learner can affect what and how much the students notice when learning L2. Research also suggested a possible relationship between learners' beliefs about language learning and learning outcomes (Mori, 199; Peacock, 1999) and between learners' beliefs and what is noticed in CF (Kartchava & Ammar, 2013). While it is important to know the effect of CF, positive or negative, it is also essential to know the factors that may contribute to creating this effect. For this reason, it was decided to examine students' preferences of CF types as one of the possible factors that may impact the effectiveness of CF.

Based on this rationale, the current study had three main goals. The first goal was to investigate the effect of CF, regardless of type, on the grammatical accuracy of EFL learners' use of the target features. The second goal was to investigate if there is a differential effect between direct and indirect CF on the acquisition of item-based and rule-based linguistic features. The third goal was to investigate if students' preferences of CF types affect the effectiveness of the given CF on the grammatical accuracy of the target features.

Research Questions

Based on the goals of the study, three questions were investigated in the study:

1. What is the effect of written CF on the grammatical accuracy of EFL learners' use of the target features over an eight-week period?
2. What is the differential effect of direct CF and indirect CF on the accurate use of different types of linguistic features, rule-based and item-based linguistic features?
3. What is the impact of individual differences, specifically preferences of CF types, on the effectiveness of written CF on the grammatical accuracy of students' writing?

Thesis Organization

The thesis is composed of five chapters. This first chapter presents the rationale of the study and the goals guiding the study followed by the research questions. In the second chapter, I review the literature related to written CF including theoretical frameworks and empirical studies investigating CF. In the third chapter, I describe the methodology of the study including the procedure for data collection, the procedure for scoring and coding, and the pilot study and how the main study was reshaped in some parts based on the results of the pilot study. In the fourth chapter, the results of data analyses are reported. In the fifth, and last, chapter, the study results

are discussed in light of previous research and related theories. The limitations of the study, the implications for the ESL/EFL writing class, and possible directions for future research based on the study's results are also discussed in the last chapter.

Chapter Two: Literature Review

The focus of this chapter is on how different theoretical frameworks, within the SLA field and the writing research field, construct the role of CF in second language acquisition and in L2 writing development. The chapter also includes a review of related empirical research and discusses the limitations and gaps of previous research which then led to the current study.

Theoretical Frameworks for CF

SLA Theories

Effect of CF in language learning. An early theoretical perspective on the teaching of grammar draws on the notion of Universal Grammar (UG), a theory of linguistics proposed by Noam Chomsky. This theory suggests that all human languages share certain characteristics and that grammar does not need to be taught in class because the linguistic ability is an innate knowledge in the human being. Researchers and linguists who hold these views believe that negative evidence provided by grammar teaching and CF is not a necessary condition for language acquisition. For example, Schwartz (1993) asserted that only positive evidence is needed for language learning while negative evidence plays little or no role in language acquisition. Based on UG, CF is viewed as playing no role in language learning.

This unsupportive position for CF can also be found in some of the early L2 learning theories, such as Krashen's Second Language Acquisition Theory (1982). Krashen suggested a number of hypotheses which all imply that CF should be avoided. The basis of Krashen's theory is the Acquisition-learning Hypothesis which states that there are two independent systems of L2 performance, acquisition and learning. Whereas acquisition is a subconscious process and happens naturally, learning is a conscious process that involves formal instruction and results in

knowledge about the language. According to Krashen, learning is less important than acquisition for L2 performance. Consequently, CF, which is part of formal instruction, is not needed as it does not affect acquisition. Also, the Monitor Hypothesis suggests that learning acts as a monitor functioning only in certain conditions. If these conditions are not met, learning, in the form of grammar teaching and/or error correction, should be avoided. The same position towards CF can be found in the Input Hypothesis which asserts that comprehensible input, i.e. positive evidence, is all that L2 learners need for acquisition. Further, the Natural Order hypothesis suggests that certain grammatical structures are acquired early whereas others are acquired later, and this natural order of acquisition cannot be changed through formal teaching. Thus, according to Krashen, CF is best avoided, and if CF is to have any role in language learning, it is very limited.

All the five hypotheses of Krashen's second language acquisition theory confirm the importance of the positive evidence for language learning while they eliminate, or give little credit to, the role of negative evidence. Krashen's theory, along with UG, may provide some answers to questions related to language learning, but they leave many other questions unanswered. For example, it has been noted that students in French immersion programs fail to acquire certain grammatical features despite the positive evidence presented by rich comprehensible input (Harley & Swain, 1978). Also, relying solely on positive evidence to acquire a language may take a long time to achieve results in L2 development (Sharwood Smith, 1991). For these and other reasons, early SLA theories were later challenged by many researchers drawing on alternate theories and empirical research.

Two such theories are the Noticing Hypothesis (Schmidt, 1990) and the Output Hypothesis (Swain, 1985). Both of these theories support a facilitative role for CF in helping L2 learners to notice any mismatch between their language production and the target-like

production. The Noticing Hypothesis (Schmidt, 1990, 1994, 1995) stated that noticing is a necessary condition for language learning; what language learners notice in input is what becomes intake. Although Schmidt does not claim that noticing is the only condition needed for learning, he asserts that it is an important first stage in language learning. Moreover, Schmidt suggested that noticing alone is not enough to convert input into intake; noticing should be accompanied by conscious comparisons between the given input and the (learners') current interlanguage. This concept is called "noticing the gap". When the learner notices a mismatch in his/her current language and compares it to the target language, learning may take place (Schmidt, 2001).

Similarly, the role of noticing in language learning is a key concept in the Output Hypothesis (Swain, 1985, 1991). The hypothesis states that the learner's production as output is a necessary condition for language learning because of the mental processes that the learner engages in when producing L2. Swain suggested three functions for the learner's output. First, the noticing function: when students use L2, they may notice differences between their production and the target-like production and, in turn, they may adjust their language production to match the target language. Second, the hypothesis-testing function: through output, the learner has the opportunity to test their hypotheses about the target language. When the students speak in an L2 class, they may receive positive or negative feedback on their output. This feedback can help them confirm, or disconfirm, their language hypotheses. Thirdly, the metalinguistic function: when using L2, the learner has the opportunity to reflect on the structure of the language they produce and, possibly, develop their linguistic knowledge.

All of the three functions of output may not take place if the learners are not provided with sufficient feedback on their output (Swain, 1991). Thus, according to this theory, CF is

essential for noticing the differences between the learner's current language and the target language, reflecting on L2 structure, and testing the learner's L2 hypotheses. The Noticing Hypothesis and the Output Hypothesis were first developed to address SLA processes through L2 speaking and oral CF. They are, however, relevant to written CF with one difference. In oral production, the opportunity to reflect on the language structure, test any language hypotheses, and notice a mismatch (language errors) might be more limited than the opportunities provided through written production. Writing allows more time to reflect and achieve the three functions of output, making written CF a useful environment for language learning.

The role of noticing was further explored by Schmidt (1995, 2001) who argued that there are two levels of awareness, a low level of awareness at which noticing occurs, and a higher level of awareness where understanding takes place. Understanding why a certain structure is considered an error and how to correct it may lead to internalizing the linguistic feature and using it unconsciously in the future.

The Noticing Hypothesis has received much attention in the field of applied linguistics and is widely accepted. Truscott (1998), however, argued that the hypothesis is based on weak grounds in cognitive psychology and has no grounds in any coherent theory of language. He also suggested that noticing is only needed for the acquisition of metalinguistic knowledge. Therefore, CF is not needed for language learning because triggering the learners' noticing is not a necessary condition for language acquisition. Still, Truscott did not eliminate the useful role of CF in constructing the learner's metalinguistic knowledge. He argued that CF has a potential role in improving the students' metalinguistic knowledge but does not play any role in the process of acquiring L2 for communicative purposes.

Effect of the type of the target features in language learning. How different types of linguistic features are affected by CF is another important question addressed in research and in the current study. It is argued that some types of linguistic errors are more amenable to CF than others (e.g. Ferris, 1999; Ferris, 2002). A distinction has been made in the literature between two types of linguistic features. The first type of features functions according to clear grammatical rules, such as verb tenses, and thus is called “rule-based” or “treatable” linguistic features (Bitchener et al., 2005; Ferris, 1999). The second type of features does not follow clear usage rules and is used differently depending on the context, such as idioms and word choice, and thus is called “item-based” or “untreatable” linguistic features (Bitchener et al., 2005; Ferris, 1999).

It is suggested that morphological, syntactic knowledge and lexical knowledge are acquired differently (Schawartz, 1993). This may imply that rule-based features and item-based features are learned differently. That is, some linguistic features might be learned easier through explicit instruction of the related rules in the language class while other linguistic features may benefit more from the implicit means of language learning, such as practicing the English language in an ESL context. The resulted knowledge from explicit and implicit learning is called implicit knowledge and explicit knowledge. Explicit knowledge is also referred to as declarative knowledge whereas implicit knowledge is referred to as procedural knowledge which can be accessed easily in online language use (Van Beuningen, 2010).

Explicit knowledge is knowledge resulting from the conscious learning and memorization of a series of facts, and thus the learner can verbalize those facts (i.e. language rules) (Ellis, 2009a). Conversely, implicit knowledge is knowledge resulting from implicit learning that happens without awareness or any demands on the learner’s attention, and thus the learner cannot verbalize what was learned (Ellis, 2009a). Implicit knowledge is available through

automatic processing, such as in conversations, whereas explicit knowledge is available through controlled processing only, (Ellis, 2009a) such as in answering language exercises or editing a piece of writing.

The role of explicit and implicit learning/knowledge in SLA processes is still a controversial topic (Ellis, 2005). Krashen in his Acquisition-learning Hypothesis (1982) made a distinction between the explicit ‘learning’ and the implicit ‘acquisition’ and argued that the benefit of explicit knowledge is limited to monitor use if enough time is available to the learner. DeKeyser (2003) also distinguished between explicit and implicit learning using the criterion of how much awareness is present in the learning process. Implicit learning is learning without awareness while explicit learning is learning with awareness; there is however little evidence that any learning can take place without awareness (DeKeyser, 2003).

How each type of knowledge possibly feeds, or does not feed, into the other type is also a question answered differently by different researchers. Krashen (1982) argued that explicit and implicit knowledge are acquired differently and that explicit knowledge and implicit knowledge are entirely separate systems, and thus they do not feed into each other. DeKeyser (1998), on the other hand, suggested that explicit and implicit knowledge are partially connected, and the gap between them can be bridged by language practice. There are also other researchers who argue that the value of explicit knowledge exceeds monitor use (e.g. Ellis, 2005) and that explicit knowledge may convert into implicit by noticing the gap (Ellis, 1997).

If the knowledge of rule-based features and the knowledge of item-based linguistic features are acquired differently, one through explicit learning and the other through implicit learning, and they represent two distinct systems, it is possible that different types of linguistic features respond differently to different types of triggers. Thus, they should be treated by

different types of CF. If, on the other hand, there is an interface connecting the two systems of knowledge, it raises the possibility that some types of CF under certain circumstances can affect both types of knowledge equally. The current study investigated whether different types of CF are needed to affect the learning of different types of linguistic features, or if different linguistic features respond in the same way to different types of CF.

Writing Theories

Whereas the focus in SLA theories is on how CF may help second language learning, writing theories are concerned with how CF may help to improve the overall quality of a certain text (Ferris, 2010; Sheen, 2011). The main concern in writing theories, such as the process model, is how students develop their writing of a text from the first stages, prewriting, to the last stages, editing. In these writing theories/models, one can notice that CF is given a very limited role and mentioned only in the last stage of the writing process as a tool helping the writer to improve the final version of his/her text. However, writing instruction can be conceptualized and viewed as a means needed to develop L2 acquisition rather than an end in itself (Reichelt, 2001). That is, writing instruction in which CF is provided regularly can be considered an “instructional technique” to help the students notice the target linguistic features and, consequently, learn those features in L2; this view of writing instruction is one way to bridge the gap between the SLA field and the L2 writing field (Sheen, 2011) because it combines both roles of CF in one view. It considers CF as a tool to improve both L2 acquisition and writing.

Since CF in L2 writing theories is viewed “as a means of helping L2 writers to revise an initial draft” (Sheen, 2011), this section presents two revision theoretical models that suggest how writers revise their work and what strategies they use in these revision processes. Before exploring those revision models, we must distinguish two types of revision/evaluation situations

in writing. The first one is where the writer evaluates his/her own text and has access to both the text and the intentions behind the text, whereas the second one is where the text is evaluated by someone else, such as a teacher providing written feedback to a student or a student providing written feedback to a peer (Hayes, Flower, Schriver, Stratman, & Carey, 1987). The following revision models reflect both types of evaluation with differences only in using, or not using, all the suggested strategies in the models.

One of the first proposals that addressed the revision process in writing is a model developed by Scardamalia and Bereiter (1983). The model suggests that the revision process consists of three main operations: Compare, Diagnose, and Operate (C.D.O.). These operations are initiated when the writer discovers a mismatch between the text written so far and the text intended to be written. The Compare operation evaluates the written text and determines if it differs from the intended text. If discrepancy is found, the Diagnose operation defines the nature of the problem and how to fix it. Then, any needed corrections or modifications are carried out through the Operate operation.

Scardamalia and Bereiter's revision model is more elaborated than the reviewing process mentioned in Flower and Hayes' writing process model (1980, 1981). However, when examining Flower and Hayes' writing model and Scardamalia and Bereiter's revision model more closely, we can find that the main operations suggested in both models are roughly the same. The evaluating sub-process in the Flower and Hayes model is the equivalent of the Compare and Diagnose operations in the Scardamalia and Bereiter model as they both imply the writer reading, comparing the current text with the intended one, and determining the problem. Also, the revising sub-process in the Flower and Hayes model is the equivalent of the Operate process in the Scardamalia and Bereiter model as they both involve the act of executing any needed

corrections or modifications in the text. Therefore, whereas Scardamalia and Bereiter's revision model seems more complex than what is suggested about revision in Flower and Hayes writing model, it does not actually offer more possible revision processes in writing than Flower and Hayes.

Moreover, it has been suggested that Scardamalia and Bereiter's model can be considered an educational tool that may facilitate revision more than a procedural model of revision (Alamargot & Chanquoy, 2001). This is probably because the model provides a description of the revising stages that all writers are supposed to go through, not what the writers actually do when revising a text. Because of these limitations, other revision models were proposed to account for what writers/teachers/peers do when revising to improve a certain text or to provide feedback to another writer. One of these models is a model developed by Hayes, Flower, Schriver, Stratman, and Carey (1987).

Hayes et al.'s model (1987) is based on protocol data, i.e. thinking-aloud protocol that requires the participants to say whatever they are thinking and doing while carrying out a certain task, in this case revising a text. What distinguishes Hayes et al.'s model is that it does not only suggest the possible processes and strategies involved in revising, but it also suggests the possible types of knowledge needed for revising (Alamargot & Chanquoy, 2001). This includes several processes and sub-processes for revising, different types of knowledge needed for revising, different types of problem representations, and several revising strategies depending on the problem type and the knowledge needed to execute the revision process. For this reason, Hayes et al.'s model is considered "the more complete and complex model" suggested for the revision process (Alamargot & Chanquoy, 2001, p. 108).

According to the Hayes et al. model, the revision process starts with a process called “Task Definition”. It is what the writer sets as the objectives of revision and how to achieve these objectives. Task Definition is fundamental in the revision process because it determines which aspects to look for, global issues or sentence-level issues, while revising. After the goals of revision have been set, the writer starts the “Evaluation” process by reading the text to understand, assess, and determine the text’s problems. This process is supposed to result in “Problem Representation”. Problem representation varies in how specific it is. The problem can be “ill-defined”, “well-defined”, or in between. At one end of the spectrum, the reviser “feels” that there is a problem with the text but cannot define the nature of it. This is called “Detection” in the model. At the other end of the spectrum, the reviser knows the exact problem and how to fix it. This is called “Diagnosis” in the model. After defining the task, reading the text, and determining the text’s problems, the reviser selects a strategy appropriate to deal with the defined problems. The model suggests five strategies that are usually used by revisers: ignore the problem (if it is superficial, for example), delay action, search for more information to better understand the problem, rewrite the text while keeping the same ideas, or revise the text to fix the problems. Choosing between rewriting and revising the text depends on many factors, such as the nature of the problem and how experienced the reviser is. Hayes and his colleagues suggest that more experienced writers usually choose the “revise” strategy over the “rewrite” strategy because their “means-end repertory” (i.e. problem-solving procedures) is more elaborated.

As mentioned previously, the Hayes et al. model is the more complex revision model among the available theoretical models as it includes all the possible procedures and strategies that could be used by writers in a revision process and the types of knowledge required for

revision. It illustrates how knowledge and revising processes interact with each other at every level/stage of revision. In addition, the model does not present the revision process as a linear process, like most previous models do, but as a recursive process in which the processes and sub-processes are interrelated and can occur at any point of the revision process. Another aspect that distinguishes this model from other models is that it explains how experienced and novice writers differ in their knowledge and revision strategies which, in turn, affect what and how much they revise.

One limitation of these revision models, however, is that they are all descriptive in nature (Alamargot & Chanqoy, 2001). They are based on descriptions of what the participants said they were doing in their revising process. Moreover, some of these models cannot be validated experimentally. For example, because Hayes et al.'s model is very complex and includes many possible strategies, it cannot be experimentally validated (Alamargot & Chanqoy, 2001). Scardamalia and Bereiter's model, on the other hand, has been validated through a class experiment and proved to increase students' attempts of revisions; however, the quality of the participants' texts did not improve as a result of these revisions. In addition, the models do not clarify where the needed information for revision is stored or how it is retrieved (Alamargot & Chanqoy, 2001). In the Flower and Hayes' writing model (1980), it is proposed that the knowledge of topic, audience, and writing plans is stored in the long-term memory. The Hayes et al. revision model (1987), on the other hand, does not indicate where the needed information for the revision process is maintained.

Another limitation is that these models do not explain how the suggested strategies and processes might be prompted in the writing class (Alamargot & Chanqoy, 2001). Addressing how these processes are prompted for writers can help language teachers to better facilitate

certain writing activities, such as peer feedback and self-editing. Also, in these writing theories, CF on accuracy plays a minor role whereas CF on content and organization is viewed to be more important to the development of students' writing skills (Sheen, 2011). In academic writing, however, accuracy is as essential and important as fluency and content, which makes CF on accuracy an important component that should be addressed and given more attention in writing theories.

There are other theoretical frameworks that address how revision is processed in writing, such as the procedural model of revision for Butterfield, Hacker and Albertson's (1996). One aspect common in all of these revision models is that they present the revision process as a complicated, purposeful act based on many skills and a series of decisions. Laying out the structure of the revision process in this detailed way provides practitioners with the types of processes, strategies, and knowledge that can be taught in class to help the students carry out more effective revisions for their writing and more productive sessions of peer CF. Another advantage of these elaborated, complex, and detailed models is that they can help writing teachers to predict where exactly in the revising process their students face problems, so that they can determine the kind of help and instruction needed for their students (Hayes et al., 1987).

As we have seen, the role of CF is viewed differently in SLA and L2 writing theories. CF in L2 writing theories is viewed as an aid to revise a text rather than a possible way to acquire L2 (Ferris, 2010; Sheen, 2011). Researchers who draw on L2 writing theories are concerned with the process of writing and the improvement of the entire text. The goal of CF in writing studies is not linguistic accuracy but the improvement of the learners' ability to write and the improvement of their texts.

The current study investigates the effect of CF on the accurate use of a specific number of linguistic features and is situated within an SLA theoretical framework. However, since the study included the practice of asking the students to revise their previous tasks and try to improve them in a second draft before starting the new tasks, the study also draws on writing theories and revision models in which CF is viewed as an aid helping to improve the overall quality of writing in terms of both content and grammatical accuracy. Revision allows students the opportunity to address the quality of their writing. The current study also draws on methodology that is characteristic of L2 writing studies in that it was conducted in a writing classroom and did not include a control group. These features (i.e. setting the study in a writing classroom and not including an empirical control group) are usually what distinguish writing studies from SLA studies as they focus on improving the quality of writing and developing the students' abilities as autonomous editors with CF provided as a helping tool to achieve these goals (Ferris, 2010).

Empirical Research

Prior to reviewing empirical studies, an overview of the terminology used in written CF research is required to understand the related research. CF has been operationalized in different ways in written CF research. One way classifies the different types of CF according to the degree of its explicitness into direct and indirect CF: direct CF is more explicit, and indirect CF is less explicit. Direct CF is “the provision of the correct linguistic form by the teacher to the student” (Ferris, 2006, p. 83). Indirect CF, on the other hand, occurs “when the teacher indicates in some way that an error has been made... but does not provide the correct form” (Ferris, 2006, p. 83).

Another way of classifying the different types of CF is by dividing them according to how many linguistic features/errors are targeted in the given CF. According to this criterion, CF can be focused, targeting one or few features, or unfocused, targeting several features. Focused

CF is to “select one or two specific types of errors to correct” in students’ writing (Ellis, 2009b, p. 98), whereas unfocused CF is to “correct all (or most) of the students’ errors” in writing (Ellis, 2009b, p. 98).

Within these criteria, there exist secondary types of CF that may fit into one or more of these categories. One of them is metalinguistic CF which has been investigated extensively in the literature. Metalinguistic CF is providing the students with “some kind of metalinguistic clues as to the nature of errors” (Ellis, 2009b, p. 98). It is usually provided at the bottom of students’ writing, in the margins, or in a separate handout. Metalinguistic CF can be considered a type of direct CF if it explains the related grammatical rule for the error, why it is considered an error, how to fix it, and/or what the correct form is. Conversely, it can be considered a type of indirect CF if it only provides metalinguistic information on the target feature but does not provide the correct form, for instance, the teacher marks the error and provides a code to the error type, such as “prep” for preposition errors, without giving the correct form. Further, if the teacher provides metalinguistic codes for all the errors in a text, it can be considered unfocused metalinguistic CF; but if the teacher provides metalinguistic information on one or two problematic linguistic features in, for example, a handout and provides it for the whole class, it can be considered focused metalinguistic CF. These examples show how complex the typology of CF can be.

The present study investigates the differential effect of focused, direct and indirect CF on different linguistic features and also examines students’ preferences of CF types and whether these preferences impacted their learning outcomes. As such, the literature review of empirical studies is divided into five areas: research on direct CF, research on indirect CF, research on the differential effect of direct and indirect CF, research on how different types of linguistic features

are effected by different types of CF, and research on students' preferences for different types of CF.

Research on Direct CF

Direct CF can be carried out in many ways, such as inserting a missing word or crossing out the error and providing the correct form (Ferris, 2006). Direct CF is explicit in two ways: it identifies the error in some way, such as underlining or circling, and it provides the correct form making the linguistic feature explicit to be noticed. Sheen (2007) explained the efficacy of the direct approach through Schmidt's Noticing Hypothesis (1995; 2001). Schmidt argues that learning does not happen without noticing. Students should first notice the target features before they can learn them. Triggering noticing can easier be achieved through direct CF than indirect CF due to its explicitness.

Much of the research on direct CF compares the performance of experimental groups that received one or more than one type of direct CF with the performance of a control group that did not receive any form of CF. There is, however, some of the early studies that did not use a "real" control group. Instead, they compared CF groups to content comments' groups. For the purpose of this literature review, the content comments' groups are interpreted as control groups.

Direct CF studies have had varying results in regard to its efficacy for L2 acquisition. The findings of early research were not positive. They reported a lack of improvement in students' writing and/or students' second language acquisition (e.g. Kepner, 1991; Polio et al., 1998). Kepner (1991) divided 60 students of Spanish, in a US university, into two groups. One group received direct CF on all language errors whereas the control group received content comments only. The results showed that the performance of both groups after the intervention was similar. The direct CF group did not make fewer errors by the end of the study which lasted for 12

weeks. Based on the results, Kepner concluded that CF has little value in the L2 writing class. Similarly, Polio et al.'s study (1998) showed no differences in linguistic accuracy between the control group who received no CF and the experimental group who received error correction (i.e. direct CF), grammar instruction, and editing exercises.

However, the findings of early research on the effect of direct CF on language development may have been influenced by specific study design features. The most important one may be the approach used to provide CF, which was the unfocused approach. The unfocused approach has been critiqued in the literature for the possibility of overloading students' attentional capacity as it requires students to attend to different types of errors at one time (Ellis, 2009b; Sheen et al., 2009). In Kepner (1991) and Polio et al. (1998), although the given CF was direct and explicit, which are theorized to be important conditions needed to trigger noticing and, consequently, lead to L2 improvement, the given CF was comprehensive, i.e. unfocused, covering all language errors in students' writing. Such a comprehensive approach may have overburdened students' limited learning capacities leading to no or little learning.

More recent research has attempted to avoid the confounding effect of treatment types and suggested a positive effect of direct CF on language improvement. In a study by Ellis, Sheen, Murakami, and Takashima (2008), forty-nine EFL students were divided into two treatment groups receiving direct CF and a control group receiving no CF. One treatment group received focused CF on articles only and the other received unfocused CF targeting a variety of errors. The results indicated that both experimental groups outperformed the control group. The findings did not show any difference in the performance between the two experimental groups suggesting that both types of direct CF, focused and unfocused, were equally more effective than no CF.

Sheen et al. (2009) investigated a similar question to the one examined in Ellis et al., (2008). However, the results of Sheen et al. (2009) revealed a different effect of the two approaches indicating that focused, direct CF is more effective than unfocused, direct CF. The difference between the results of the two studies can be interpreted by looking at the target features of the two studies. In Ellis et al. (2008), only four errors were corrected in both, focused and unfocused, groups with the difference being that in the focused group all the four errors were article errors whereas in the unfocused group two errors were always articles and the other two were of a different type of linguistic features, such as prepositions and past tense; thus, both groups received the same amount of CF on their writing. On the other hand, in Sheen et al. (2009), the focused group received correction on articles, between 2 and 8 errors were corrected, while the unfocused group received CF on articles, past tense, copula Be, and prepositions. A maximum of two errors per linguistic feature, and up to 8 errors for the whole writing, were corrected in the unfocused group. There exists a difference in the number of errors targeted in the given correction between the two studies. Since the given direct CF to the unfocused group in Sheen et al. (2009) was more comprehensive than that of Ellis et al. (2008), it may have overwhelmed the students' learning memory making the unfocused CF less effective than focused CF. However, both studies found that focused, direct CF is effective in producing more accurate subsequent writing.

The current research on direct CF has not examined only one type of direct CF, but many different options of direct CF, such as direct CF accompanied by metalinguistic explanations versus direct corrections only. This body of research suggests that all types of direct CF are effective to some degree. Yet, the results are inconclusive in regard to which type of direct CF is more beneficial for improving students' writing/second language. Bitchener (2008) compared the

effect of three types of direct CF on students' accuracy improvement: direct CF accompanied by written metalinguistic explanations (provided in a handout) and oral metalinguistic explanations (provided in a mini lesson for the whole group), direct CF accompanied by written metalinguistic explanations, and direct CF only. The results showed that the three experimental groups outperformed the control group on the accurate use of articles. However, the study did not provide conclusive results in regard to which type of direct CF is more effective.

These inconclusive results in terms of the most effective type of direct CF are seen again in Bitchener and Knoch's study (2010a) and have raised the impact of yet another aspect of the study design: the length of the investigation. Bitchener and Knoch (2010a) investigated the effect of the same three direct CF types investigated in Bitchener's study (2008), but over a longer period of time, ten months. In all of the posttests (the immediate posttest and three delayed posttests distributed at different times after the treatment), the experimental groups outperformed the control group on the accurate use of the target features (articles). Yet, no differences were found in the effectiveness of the three different treatments provided in the study.

Despite the inconclusive results of Bitchener and Knoch (2010a), the study has a particular value in research because of its 10-month duration. It is the only recent study that investigated the effectiveness of CF over such an extensive period of time and provided evidence that CF can have a positive long-term effect on L2 learning. One limitation of the study though, as the researchers explained, was the target features. The study targeted only two functions of the article system: 'a' for first mention and 'the' for subsequent mention. These functions are not only relatively easy but also partially known to the participants. Future longitudinal research on CF should consider targeting more challenging linguistic features. Only when such research is

conducted and the results are considered, can we have a better understanding of whether written CF has a long-term effect on the acquisition of complex linguistic features.

One study that investigated two types of direct CF and suggested a differential effect is Sheen (2007). The study included three groups receiving one of the following treatment/control: direct CF, direct CF accompanied by metalinguistic comments that explain the correct form (provided at the bottom of students' writing), or no CF. The results showed that both CF groups outperformed the control group. Also, the two direct CF types had a differential effect: the group who received direct CF with metalinguistic comments outperformed the group who received direct CF only. Sheen attributed the results to the type of awareness triggered by each type of direct CF. Sheen argues that direct CF can trigger noticing only whereas direct CF accompanied by metalinguistic comments can promote both noticing, the first level of awareness, and understanding, which is a deeper level of awareness, as suggested by Schmidt (1995, 2001). Thus, Sheen (2007), unlike Bitchener (2008) and Bitchener and Knoch (2010a), provides more conclusive results in regard to the more effective types of direct CF.

The findings of Sheen (2007) are contradicted by a similar study, Shintani and Ellis (2013). The study investigated two types of direct CF similar to the ones investigated in Sheen (2007) with the difference that the group which received metalinguistic CF did not receive any direct corrections on the target features but only metalinguistic explanations for the target features provided in a handout. The results indicated that the metalinguistic explanation group improved in both the revised text and the immediate posttest (a new piece of writing) but not in the delayed posttest. As for the direct CF group who received direct corrections only, the students did not improve in either the revised text or the new pieces of writing (i.e. both the

immediate and delayed posttest). These findings suggest no short or long-term effect for direct CF and no longitudinal effect for metalinguistic CF.

The negative findings of Shintani and Ellis (2013) can be attributed to many reasons. First, the metalinguistic explanation group did not receive any error correction, or even error identification, on their writing. Instead, they all received the same metalinguistic explanation in the form of a handout. This type of treatment may have affected the degree of explicitness that usually accompanies metalinguistic CF and made the students unclear about where the errors were in their writing and whether their self-correction was in fact accurate. Second, the given treatment included only one session of CF. One session of either direct CF or metalinguistic explanation without direct CF on indefinite articles might not provide enough of the negative evidence (Gass, 2003) needed for the students to notice (Schmidt, 1995) the mismatch between their production and the target-like production. Another limitation of the study is the approach used. It is highly focused targeting only indefinite articles, a practice that is usually not possible in real classes due to time constraints and teachers' desire to improve as many linguistic features as the class time and students' level allow.

In conclusion, unlike the early research, the more recent research shows that focused, direct CF is effective in promoting the acquisition of certain grammatical features (e.g. Bitchener, 2008; Bitchener & Knoch, 2010a; Sheen, 2007). The effectiveness of direct CF can be attributed to its explicitness (Ellis, 2009b) and immediacy in providing the correct form to the learner. Another suggestion provided by the current research is that direct CF accompanied by metalinguistic information on the target features might be more effective than direct corrections alone (e.g. Sheen, 2007). However, because there is more research suggesting no differences between the different types of direct CF (e.g. Bitchener, 2008; Bitchener & Knoch, 2010a),

further research is certainly needed. Also, the current research suggests a long-term effect of direct CF on the learning of articles. Future research need to investigate the possible long-term effect of direct CF on other types of linguistic features that are more challenging and more complex than articles.

Research on Indirect CF

Unlike direct CF, indirect CF does not provide the learner with the correct form; it leaves the correction to be made by the learner. This type of CF can take several forms, such as circling, underlining, and/or coding (Ferris, 2006). The following review presents research on indirect CF that compares the performance of students who received indirect CF on their writing with the performance of a control group who received no CF. Some of the early research, however, included a comparison group rather than an empirical control group, such as Sheppard (1992) in which the performance of an error correction group was compared to that of a content comments group.

It has been argued that since indirect CF encourages the students to self-correct their errors using either their previous knowledge or any external resources such as writing and/or grammar books, it can promote guided learning and problem-solving on the part of the learner (Lalande, 1982; Chandler, 2003). For this reason, indirect CF is more likely to foster long-term learning than direct CF (Bitchener & Knoch, 2008; Lalande, 1982). Research, however, has not always lent support to these arguments, as will be seen in this review.

In the earlier years, research on indirect CF has suggested either no or little positive effect on language learning. Sheppard (1992), for example, compared two groups, a treatment group who received indirect CF in the form of error coding and a control group who received content comments in the form of requests for meaning clarification. The indirect CF included, but was

not limited to, verb tense, punctuation, and subordination. The results showed that the content-comments group outperformed the error-correction group in both content and accuracy. The error-correction group made some improvement on accuracy, but the improvement was achieved in only one linguistic feature, verb use. Moreover, Sheppard found that the complexity of the writing produced by the error-correction group decreased over time as a result of the students avoiding the linguistic features on which they received error correction.

There are other studies that suggested the same limited effect for indirect CF (e.g. Truscott & Hsu, 2008). This limited or negative effect can be attributed to many reasons; one of them is the nature of indirect CF. Indirect CF provides the learner with little information in regard to how to correct the error and what the correct form is (Bitchener, 2008). It requires the learners to carry out the corrections by themselves and then wait to confirm, or disconfirm, these corrections from the teacher through another draft that may, or may not, be required. The additional time spent on waiting for the corrections may counterbalance the cognitive effort spent to figure out the corrections, which then leads to limited learning (Chandler, 2003). Indirect CF is, therefore, disadvantaged by both the little information it provides to the learner and the time it takes to confirm the accuracy of the corrections made by the student.

A more recent study that also showed no accuracy improvement in the production of new writing as a result of indirect CF is Truscott and Hsu (2008). In the study, students were divided into two groups, one experimental group receiving unfocused, indirect CF and one control group receiving no CF. Both groups were asked to complete a writing task in the first session, revise it in the second session, and then write a new piece in the third session. The results showed that the CF group outperformed the control group significantly in the accuracy of revised texts. However, a week later, when both groups were asked to write a new text, it was found that the accuracy

levels of both groups were identical. Truscott and Hsu concluded that CF is only effective in producing more accurate revised work but ineffective in producing accurate new writing in the near or far future.

The negative results that Truscott and Hsu reported in their study can be ascribed to many reasons. First, the indirect CF provided in the study was in the form of underlining only. No information was given on the error type, why it is considered an error, and/or how to correct it. Consequently, what is learned from CF in the second draft might not be stored in the student's long-term memory as it did not trigger anything but probably noticing. Also, the given CF was unfocused, which means that it targeted all errors in students' writing. As mentioned previously, unfocused CF can overload the students' attentional capacity and, thus, might not be stored in student's long-term memory. In addition, only one session of CF was provided in the study. Providing students with as little as one instance of unfocused, indirect CF and then expecting them to carry out what they learned from this one session into a new context is probably not going to work. For learning to take place and be transferable to new contexts, consistent and longer treatment is imperative in the L2 environment.

Although some research suggested no or only a limited effect of indirect CF on language learning, there are other studies that suggested otherwise. Fathman and Whalley (1990) divided 72 ESL students into one control group and three experimental groups receiving one of the following treatments: indirect CF in the form of underlining, content comments only, or a combination of indirect CF and content comments. The results indicated that both of the indirect CF groups outperformed the control group and the content comments group, suggesting a value in the practice of providing students with indirect CF on their writing. However, the study lasted for only a few days and the improvement in accuracy was measured through revisions only (i.e.

second drafts). These conditions can make the results less valuable. To suggest that a certain practice is useful for second language acquisition, it has to lead to learning that is transferable to new contexts, and the learning has to last for a period of time that is longer than a few days. There are other studies in which accuracy was measured through revisions only and reported positive effect of indirect CF (e.g. Ashwell, 2000). Although these studies showed that CF can be used as a tool to improve a particular piece of writing, they do not provide evidence for language learning (Truscott & Hsu, 2008). What is meant by learning here is internalizing the linguistic feature(s) and being able to use it accurately, and probably unconsciously, in not only revised work but also in new contexts.

There are other studies that lasted for semesters and measured accuracy in new pieces of writing and still suggested a positive value for indirect CF. These positive results were achieved under certain conditions which include, but are not limited to, providing more than one instance/session of CF. Lizotte (2001), for instance, provided his students with continuous indirect written CF over the course of one semester. By the end of the semester, students' accuracy improved significantly without any negative effect on fluency. These findings were echoed in Chandler (2000) who reported similar positive effects of indirect CF on students' accuracy in revised and subsequent writing over one semester without affecting fluency or quality.

Another condition under which indirect CF can lead to writing development is making the given indirect CF more explicit by coding the students' error types or providing metalinguistic explanations to the error(s). In a study by Ferris, Liu, Sinha, and Senna (2013), the researchers provided the participants with focused, explicit, indirect CF on their writing and asked the participants to revise their written work. The given CF took the form of marking the

error and coding the error types. The participants' revised work indicated writing improvement suggesting a positive effect for explicit, indirect CF. By underlining the error, coding the error, and/or providing metalinguistic information on the error as rule reminders, students have a better chance to self-edit their errors by accessing their prior grammatical knowledge gained from formal grammar classes (Ferris et al., 2013).

As research shows, indirect CF can be effective for writing development and/or second language acquisition when provided under certain conditions (e.g. Ashwell, 2000; Chandler, 2000; Lizotte, 2001). These conditions include making the given CF more explicit and providing indirect CF for as much as needed, usually for more than one instance. The constant and consistent treatment can counterbalance the little information provided through indirect CF, thus, possibly leading to long-term learning. Future research should investigate whether the proficiency level of students affects how much they benefit from this type of CF. Also, future research needs to investigate if indirect CF is more useful to acquire certain linguistic features that are different from the type of linguistic features acquired through direct CF. The nature of indirect CF, that is based on self-correction, may promote the learning of different types of linguistic features. When research provides answers to these questions, practitioners would gain better knowledge of how to use indirect CF, with what type of students, and with what type of linguistic features.

Research on the Differential Effect of Direct and Indirect CF

All the studies mentioned previously investigated the efficacy of either one approach, direct or indirect CF, or different options of one approach, for example direct CF with metalinguistic explanations versus direct CF only. These studies cannot be used to suggest the superiority of direct or indirect CF over each other. Only those studies that investigate the

differential effect of the two approaches, by providing direct CF and indirect CF to different groups of participants, can be used to suggest which approach is more facilitative for writing improvement and second language acquisition.

A limited number of studies have compared direct and indirect CF. Similar to most of the early research on written CF, the early research on the differential effect of direct and indirect CF suggested no or limited effect for the two approaches. Semke (1984) divided 141 students of German, in a US university, into four groups receiving one of the following treatments on their writing: content comments, direct CF, combination of content comments and direct CF, or indirect coded CF. The given CF was unfocused covering all language errors. The results showed that there was no positive effect for any of the four feedback types. Semke concluded that CF does not improve accuracy or fluency but, instead, affects them both negatively.

Similarly, Robb, Ross, and Shortreed (1986) compared four types of CF that differ in terms of their degree of salience. One of the four groups was provided with direct CF while the remaining three groups were provided with one of the following indirect CF types: error-type coding, error highlighting, and marginal error totals. Researchers reported that the writing of all groups improved by the end of the academic year regardless of the given type of CF. Because the differences in performance between the treatment groups were negligible, researchers concluded that the writing improvement was due to writing practice, not the given CF. They also suggested that the less time-consuming CF methods (i.e. indirect CF) might be enough and that the time spent on direct CF should be better spent on other productive writing activities since it does not lead to any more accuracy than indirect CF.

Other early studies found more improvement with one type of CF than the other, but the difference was not statistically significant. Lalande (1982) divided 60 students of German in a

US university into two groups, direct CF and indirect CF groups. Both groups received unfocused CF covering all language errors. The researcher found that the students who received indirect CF and then self-corrected their errors made more improvement in accuracy than the direct CF group. Lalande concluded that indirect CF was more effective than direct CF because it engaged students in guided learning and reflection that can lead to long-term learning. However, the improvement achieved by the indirect CF group was not statistically significant. Furthermore, accuracy was measured through revisions only, not through new pieces of writing. Hence, the results cannot be used to make a case for long-term effect of indirect CF as Lalande (1982) suggested.

Most early research, including the three studies mentioned above, did not include a control group that did not receive any type of CF. All groups in the above studies received some type of feedback on their writing. In this case, even if the study indicated improvement in students' writing after intervention, the positive effect could be the result of other factors, such as writing practice (e.g. Robb et al., 1986). Another limitation in early research is targeting all errors in the given CF. Robb et al. (1986), for instance, targeted all categories of lexical, syntactic, and stylistic errors. Providing the students with corrections that cover all their errors may be one of the reasons why the students did not improve in accuracy in early research.

More recent research on the differential effect of direct and indirect CF has attempted to avoid some of the limitations mentioned above and has produced different results. The suggestion made by the recent research is that direct CF is more effective than indirect CF in at least one of the following writing tasks/tests: revised text, new text given in an immediate posttest, and/or new text given in a delayed posttest. One study that compared the effectiveness of direct and indirect CF is Chandler (2003). This study investigated four types of CF, direct CF

and three indirect options: underlining with error type description, description of error type only given in the margins, and underlining only. The results suggested that direct CF was superior to the three options of indirect CF in producing more accurate revised writing. The researcher attributed the results to the immediacy of the direct approach compared to the indirect approach. As for subsequent writing, both direct corrections and underlining only led to writing improvement, and there was no statistically significant difference between them.

The efficacy of direct CF has also been found to be sustainable over time. Van Beuningen, Jong, & Kuiken (2008) investigated the effect of direct and indirect CF and the effect of writing practice on students' revised and new texts. Sixty-two high school learners of Dutch, in The Netherlands, were divided into two experimental groups, direct and indirect CF, and two control groups, writing practice and self-correction. In the first session, all groups were given a writing task to complete. In the following week, the experimental groups received their writing back with CF, either direct or indirect, and they were asked to rewrite their texts revising all errors. The writing practice group was only given a new writing task to complete while the self-correction group was asked to rewrite the first task correcting any errors they can detect by themselves. In the last session, all groups were given a new writing task to complete. The results showed that both experimental groups outperformed both control groups in producing accurate text revisions. However, only the direct CF group produced more accurate subsequent (new) writing. The researchers concluded that direct CF can have long-term effect on accuracy improvement because it is faster in providing students with the needed corrections and, thus, it makes it easier for the students to internalize the given CF. Indirect CF, on the other hand, is disadvantaged by the additional delay for the students in knowing whether their self-corrections are accurate.

The Van Beuningen et al. study (2008), however, had clear limitations. First, the study investigated the effect of CF using the unfocused approach targeting a wide variety of grammatical and non-grammatical linguistic features. Although the study suggested that unfocused CF can lead to language improvement, there is no evidence that the students will retain this level of improvement months or a year after the treatment. The unfocused approach does not allow in depth reflection on the identified errors, an aspect that can contribute to retention. Secondly, the length of the study was short, three weeks only. This is too short to allow the investigation of long-term effect on language learning. Measuring the accuracy of students' writing over a longer period of time would strengthen the study's conclusions.

Bitchener and Knoch's study (2010b) provided a stronger argument for the long-term effect of direct CF. The study investigated the efficacy of three types of CF within the focused approach targeting two functions of the article system: 'a' for first mention and 'the' for subsequent mention. Advanced ESL students were divided into four groups receiving one of the following treatment/control: written metalinguistic CF, indirect CF (circling), written metalinguistic CF accompanied by oral form-focused instruction, or no CF. All groups were asked to write a new piece of writing on three occasions: a pretest, an immediate posttest, and a delayed posttest. The results revealed that all treatment groups outperformed the control group in the immediate posttest; however, only the direct CF groups were able to retain their level of improvement in the delayed posttest (i.e. the group of written metalinguistic CF and the group of written metalinguistic CF accompanied by oral form-focused instruction) suggesting that direct CF is more effective than indirect CF for long-term learning.

Bitchener and Knoch's study (2010b) lasted for ten weeks. The delayed posttest was administered eight weeks after the immediate posttest. Although the period of two months is

fairly short when making a suggestion of long-term learning, it still shows that the students attended to the given CF in that one single session and then were able to retain the same level of accuracy two months later. What made this achievement possible for the direct CF groups was not direct correction, which was not provided for any group, but the metalinguistic information on the target features, whether written only or accompanied by form-focused instruction. Sheen (2007, 2010) argued that, based on the Noticing Hypothesis, direct CF can promote noticing, whereas metalinguistic CF can promote understanding, which contributes to a deeper level of learning. Thus, the more explicit and direct the CF is, the more learning should take place; and that is exactly what the results of Bitchener and Knoch (2010b) indicate. Both groups who received metalinguistic CF outperformed the indirect CF group in the delayed posttest.

Research on how Different Types of Linguistic Features Respond to Different Types of CF

Linguistic errors have been classified in the literature as treatable and untreatable errors (Ferris, 1999). Whether an error is treatable depends on the nature of the linguistic feature, specifically whether the feature is “rule-governed” (Bitchener et al., 2005) or, using another term, “rule-based” (Bitchener, 2012). Treatable errors are errors that “occur in a patterned, rule-governed way” (Ferris, 1999, p. 6). They include, but are not limited to, subject-verb agreement, verb tense or form, articles, pronouns, and spelling (Ferris, 2006). Untreatable errors, on the other hand, are idiosyncratic by nature and, thus, cannot be treated by consulting a certain set of rules (Ferris, 1999; Bitchener et al., 2005). They are also called “item-based features” (e.g. Bitchener, 2012). Untreatable errors, or item-based features, can belong to one of the following categories: word choice, idioms, and sentence structure (Ferris, 2006).

The knowledge of rule-based and item-based features represents separate domains of knowledge (Ferris, 1999, 2002). It has been suggested that syntactic knowledge and lexical

knowledge are learned in different manners (Schwartz, 1993). One way of learning the rule-based features is through the explicit teaching of grammar rules, whereas the item-based features can be learned through practice, exposure to L2, and other means. For this reason, some researchers, such as Ferris (1999, 2002) and Bitchener et al. (2005), recommend using different types of CF with different types of linguistic features. However, there is not much research on the relationship between the type of the linguistic features being treated and the type of the given treatment (i.e. CF).

One of the few studies that looked at this relationship is Bitchener et al. (2005). The study examined the effect of focused, direct CF on the acquisition of three linguistic features: past simple tense and the definite article (treatable features) and prepositions (less treatable features). The participants received one of the following treatments: direct, explicit written CF accompanied by student-researcher five-minute conferences; direct, explicit written CF only; or no CF. The results showed a significant improvement in students' use of the past simple tense and the definite article when combining oral and written CF for students. However, no overall effect was found for any type of the given CF on the use of prepositions. These results suggested that direct CF is effective with only a certain type of linguistic features. It is more effective with treatable errors than with untreatable or less treatable errors.

The practices of L2 writing teachers seem to reflect this distinction as well. One of the research questions Ferris (2006) asked is, "Are different types and categories of errors affected differently by error treatment?" (p. 94). Surveying hundreds of texts produced by 92 ESL students over a 15-week semester, it was found that teachers marked nearly 59% of treatable errors using indirect CF. The students, however, did not make progress in the use of these treatable categories in their revisions (i.e. second drafts), but towards the end of the semester,

they improved in the use of those treatable categories. Conversely, it was found that teachers marked more than 65% of untreatable errors using direct CF. In their revised texts, students were successful in editing these errors based on the given direct CF but made no progress in the use of those untreatable errors over time. These results suggest that teachers usually provide indirect CF on treatable errors with the justification that since the errors are treatable, students will be able to self-correct them. On the other hand, teachers usually provide direct CF on untreatable errors with the justification that students will not be able to fix them since these errors are idiomatic in nature.

The question remains: should different CF treatments be used for different error types? Ferris (2006) cannot be used to suggest an answer to this question as the study reported teachers' CF practices only and did not investigate their effectiveness. Although it was found that the treatable errors were eliminated over time in students' writing, these findings cannot be used to draw conclusions in regard to the reason(s) behind students' improvement in the use of those treatable features. Many reasons can be suggested to interpret this improvement, such as the fact that the errors are treatable and, thus, learnable with time, writing practice, and/or the given indirect CF. Moreover, many studies suggest that direct CF, not indirect CF, is effective in improving students' use of treatable features (e.g. Bitchener et al., 2005; Bitchener, 2008; Ellis et al., 2008). These different and conflicting results on how certain linguistic categories respond to a certain type of CF call for more research.

In a recent study, Van Beuningen, Jong, and Kuiken (2012) addressed how different types of linguistic features respond to direct and indirect CF. The main goal of the study was to investigate the effect of direct CF, indirect CF, writing practice, and self-correction on language improvement, which are the same types of treatment/control investigated in Van Beuningen et al.

(2008). What distinguishes Van Beuningen et al. (2012) from Van Beuningen et al. (2008) is the addition of a delayed posttest, carried out three weeks after the first posttest. The findings showed that both experimental groups, direct CF and indirect CF, were able to retain their level of improvement gained from CF in the delayed posttest, unlike the two control groups.

Van Beuningen et al. (2012) also conducted a separate analysis to determine how grammatical and non-grammatical errors respond to different types of CF. The analysis showed that grammatical errors benefited more from direct CF whereas non-grammatical errors, such as spelling and capitalization errors, benefited from indirect CF. Still, the study did not address how untreatable or less treatable grammatical features respond to CF. Moreover, because the delayed posttest was administered only three weeks after the treatment, the study cannot be used to address the long-term effect of direct or indirect CF on the acquisition of different types of linguistic features. Based on these limitations and the limited body of research on the differential effect of direct and indirect CF on the acquisition of treatable and untreatable linguistic features, further research is clearly needed on this important topic to inform theory, research, and the writing classroom.

Research on Students' Preferences for Different CF Types

Why some learners benefit from CF while others fail to do so is another question raised in research. Many variables may impact the effectiveness of CF including individual learner differences. One individual difference may be learners' preferences for how CF should be provided on written errors. Students have different preferences in regard to the most effective types of CF, the amount of CF to be given, and the type of errors to be addressed. There are studies suggesting that learners' preferences of CF may affect the students' ability to notice and attend to the given CF (e.g. Kartchava and Ammar, 2013). Other studies also suggested that

learners' beliefs about language learning may affect learning outcomes (e.g. Mori, 1999; Peacock, 1999).

One question investigated in regard to learners' preferences of CF practices is how much CF should be provided on writing. A number of studies have suggested a similar finding; the majority of students prefer that all of their errors are addressed in CF provided on their writing (e.g. Amrhein & Nassaji, 2010; Lee, 2005). Amrhein and Nassaji (2010) investigated the CF preferences of 33 adult ESL learners in two English schools in Canada. The results indicated that the majority of students (93.9%) believed that for written CF to be most effective, it has to be provided on all errors. On the other hand, all the students in this study rejected the questionnaire options that indicated that CF should target major errors only or the errors that interfere with communication only. Also, all the students rejected the option that stated that feedback should respond to content only and not address any errors. Similarly, Lee (2005) examined the perceptions and attitudes of L2 secondary students towards written CF. The students' responses to a questionnaire and follow up interviews revealed that the majority of students (82.9%) preferred to receive comprehensive CF targeting all of their errors. These results show how much language learners value CF, believing that CF is a useful tool for language learning, and, thus, it should be provided on all errors.

Another question investigated in research on students' CF preferences is what the most effective types of CF are from the students' perspective. For written CF, some studies suggest that students prefer their errors to be corrected explicitly rather than implicitly, because, as the students explained, it allows them to understand the error and how to fix it (Amrhein & Nassaji, 2010). In Chandler (2003), two-thirds of the participants thought that the "correction" response (i.e. direct CF) was the most useful method that can help them to correct their errors easily, while

nearly half the students thought that “underlining with description” (i.e. indirect CF) was the method from which they learned the most to write more accurately in the future. Amrhein and Nassaji (2010) reported similar preferences: the participants, ESL students, preferred their errors to be “marked and corrected”, while they thought that “a clue with no correction” is not useful as it does not provide specific and clear advice on their errors. Lee (2005) reported that 75.7% preferred that they receive direct corrections to all of their errors; one student explained this preference by stating, “sometimes I don’t know why I made the errors. I think it’s clearer if the teacher writes it down [the correct form] for me” (p. 7). Further, Chen, Nassaji, and Liu (2016) investigated the preferences of 64 EFL learners using a questionnaire. The results indicated that students preferred explicit types of CF, such as “correcting the error and then providing an explanation for the error”, to implicit types of CF, such as “simply indicating that you have an error... without locating or correcting the error”. Still, many students in the same study provided positive responses to self-correction techniques. The researchers in Chen et al. (2016) explained that the more advanced the students were, the less they required explicit feedback on their errors.

Studies investigating oral CF provide an interesting perspective. Many studies suggested that students have a positive attitude towards self-correction (e.g. Kartchava & Ammar, 2013; Yoshida, 2008), and also a positive attitude towards explicit CF that provides information, explanations, and examples on the errors to help them self-correct them (Agudo, 2012; Park, 2010). Thus, while students prefer direct CF to indirect CF when provided with written CF, they prefer self-correction when it comes to oral CF. The differences between students’ preferences of oral CF types and their preferences of written CF types may be related to the nature of oral CF being carried out in front of other students in comparison to written CF being given privately on individual papers.

As for what type of errors should be addressed in CF, research suggests that language learners are concerned about the accuracy of their writing, and, thus, they prefer that their teachers address their grammatical errors in writing. For example, Amrhein and Nassaji (2010) reported that the students had positive opinions on CF addressing focus-on-form errors, such as grammatical, spelling, and vocabulary errors while they had either neutral or negative attitudes towards CF addressing the content or ideas of their writing. Diab (2005) reported a case study in which both of the two ESL students participating in the study expressed their desire to receive error correction, with both students using the phrase, “I want my/all the grammar mistakes corrected because I want to become better/it’s better” (p. 39). However, other studies suggested other types of feedback demanded by advanced L2 learners. In Chen et al. (2016), in which participants came from three English proficiency levels: intermediate, advanced-intermediate, and advanced, the most favored error types on which the students wanted to receive written CF were “organizational errors” followed by “grammatical errors” and then “vocabulary errors”. Chen et al. (2016) also reported that the participants regarded “extended comments” to be essential for writing improvement and that the students’ most favored types of comments were “comments on the overall quality of writing” while “comments on grammar” were their least preferred types of comments.

Overall, research show that L2 students have strong opinions on how CF should be implemented. Whether these preferences had an effect on language learning outcomes is also a question addressed in research. The answer to this question is still not clear as few studies were conducted to address this possible relationship. Peacock (1999) examined the beliefs of 202 EFL students about language learning to understand if particular beliefs had any impact on language learning. The results indicated that some students’ beliefs were associated with lower EFL

proficiency, suggesting that these beliefs were detrimental to L2 learning. However, Peacock also acknowledged that it is difficult to decide the direction of causality in his study, i.e. to decide if the beliefs caused the negative learning outcomes or vice versa. Mori (1999) investigated the beliefs of 187 students in a US university enrolled in a Japanese-as-a-foreign-language class and examined the relationship between their learning beliefs and L2 achievement. The results suggested an association between certain beliefs and high or low levels of achievement in L2, for instance, learners who believed that language learning was easy achieved higher levels of L2 proficiency. Moreover, as mentioned previously, there are some studies suggesting that advanced EFL learners require feedback on content and organization of their writing more than feedback addressing grammatical errors (e.g. Chen et al., 2016) which may indicate there is a relationship between the students' English proficiency level and their preferences for certain types of feedback on writing.

Other studies suggested a weak link or no link between students' beliefs/preferences and their performance in L2. Tanaka and Ellis (2003) examined the relationship between Japanese university students' beliefs about language learning and their English proficiency before and after their 15-week study-abroad program. Correlations indicated a weak relationship between the students' language learning beliefs and their TOEFL scores both before and after their study abroad period. In Kartchava and Ammar (2013), four groups (high-beginner college students) received one instruction with different types or no oral CF on two specific spoken grammatical errors and completed a beliefs' questionnaire. To measure learning, all groups completed two spoken tasks, one for each target feature, before and immediately after the intervention. To measure noticing, the participants completed lesson reflection sheets and immediate recall sessions in which they were prompted to write down their thoughts about the lesson during class

activities. The results indicated that learners' beliefs about CF affected their ability to notice the given corrections but did not affect their learning outcomes. These mixed and inconclusive results on the relationship between learners' preferences and learning outcomes or the effectiveness of CF call for further research.

Implications for the Present Study

Current research suggests that using written direct CF and also indirect CF can have positive effects on L2 acquisition and writing development in the presence of certain conditions. These conditions include targeting a limited number of linguistic features in each piece of writing, targeting linguistic features that are amenable to CF (articles, for example), and providing written CF on several pieces of students' writing, with the rationale that one instance of CF might not be enough to lead to L2 acquisition (e.g. Ashwell, 2000; Bitchener, 2008; Bitchener & Knoch, 2010a; Lizotte, 2001; Sheen, 2007).

However, other questions related to the practices of effective CF remain inconclusive. One question is: Is there a differential effect between direct and indirect CF on the grammatical accuracy of different linguistic features? The available research suggests that direct CF might be more effective than indirect CF for the acquisition of certain linguistic features (Bitchener & Knoch, 2010b; Van Beuningen et al., 2008). However, much of the current research has targeted one type of linguistic feature, rule-based features. If direct CF is useful for the acquisition of certain linguistic features, indirect CF might be useful for other types of linguistic features. Research should examine the potential benefit of indirect CF compared to direct CF in terms of facilitating learning of different types of linguistic features. This is one of the current study's goals.

The long-term effect of written CF is another important question. The answer to this question can suggest to what extent CF is useful for L2 acquisition. One interpretation for long-term learning is for the learning to be transferable into a new context. The effect of CF should be evident not only in revised texts but also in new pieces of writing. Yet, much of the previous research on the differential effect of direct and indirect CF has either used revised texts or new texts to measure students' accuracy in writing. The ability to revise does not prove that CF has a long-term effect on learning (Storch, 2010). Also, asking the students to only read the given corrections then start writing a new piece of writing may not give the students enough time to think about and process the given corrections (Ferris, 2010). For these reasons, the current study included both revisions and new pieces of writing in the design of the study.

Another line of inquiry that we need to investigate in order to understand how CF works (or may not work in certain cases) is the impact of variables related to individual differences. One of these variables is students' preferences of CF types. Few studies have addressed the relationship between preferences of CF types and learning outcomes, and findings have been inconclusive, calling for more research. One of the current studies' goals is to examine students' preferences of direct and indirect CF and discover if these preferences have an impact on the effect of written CF on accuracy in writing.

Based on these gaps in research, the three goals/questions guiding this study are: What is the effect of written CF on the grammatical accuracy of EFL learners' use of the target features over an eight-week period? What is the differential effect of direct CF and indirect CF on the accurate use of different types of linguistic features, rule-based and item-based linguistic features? And, what is the impact of individual differences, specifically preferences of CF types, on the effectiveness of written CF on the grammatical accuracy of students' writing?

Chapter Three: Methodology

Using a quasi-experimental design, the present study investigated the effect of direct and indirect CF on the acquisition of two different types of linguistic features. The study also examined the impact of students' preferences of CF types on the effectiveness of CF on grammatical accuracy of writing. A total of 50 students participated in the study. They were divided into four experimental groups receiving three sessions of direct or indirect CF and completing two pretests, an immediate posttest, and a delayed posttest over the course of twelve weeks. They also completed two questionnaires. The study was carried out in a college writing class taught by the researcher in an EFL context.

This chapter presents the method used to investigate the research questions. It starts with a description of the context of the study and the research participants. It then explains the study design, procedure for data collection, instructional treatments, data collection tools, and the procedure for data coding and data analysis. Further, the chapter reports on the findings of the pilot study and the impact of these findings on the main study.

Context

The study was conducted in a bachelor of nursing program in a large public university in Saudi Arabia. The nursing program was four years long preceded by a preparatory year in which the students were enrolled in an English intensive program. In the nursing program, the language of instruction was English except for courses outside the nursing program, which were taught in Arabic. In the first year of the program, students took the required nursing courses plus one English writing course called Expository Writing. This writing course was divided into two parts, Expository Writing 1 taught in the first semester and Expository Writing 2 taught in the second semester of the first year.

This study was carried out in the second semester, Expository Writing 2. The course was offered once a week for two hours. The goal of the course was to prepare students to write academic essays for their college classes. The course focused on accuracy in many of its topics, such as grammar basics, ESL writers' challenges, clarity, and writing papers using APA style. Because accuracy in writing represented a key aspect in this course, the course was chosen to be the context of the present study.

The researcher acted as the class instructor in the course during the period of the study. The researcher is a faculty member in the university where the study took place and was given access to the participants on condition that she teach the course for the entire semester, including the study period. It was therefore necessary to take on this dual role. To mitigate any possible conflict of interest, a colleague in the same institution who was not associated with the research project or with the participants as students acted as a research assistant on behalf of the researcher.

The research assistant was invited into class to carry out the following tasks: describe the project and distribute, review, and collect the consent letters (Appendix A) and the questionnaires (Appendix B and Appendix C). These tasks were carried out with the researcher absent from the classroom. Before the study started and to help the research assistant performed her tasks effectively, I explained to the assistant the nature of the project, what the students were required to do as part of the project, and also what is meant by every question in both questionnaires, in case students had questions during the process.

The consent forms and questionnaires were not provided to the researcher until the course ended and all marks were submitted. As for the writing tasks and tests, all students participated in those tasks as part of their instruction. During the course and the duration of the study, the

students were treated as students rather than research participants. The researcher did not know who did or did not consent to participate in the study until the course ended. This was intended to allow the participants to choose not to participate without concern for prejudice on behalf of the researcher as the instructor of the course.

Participants

Fifty students participated in the study. The participants were nursing students in their first year of their undergraduate program. The program admits male and female students as separate cohorts. I taught the female section only. All the participants were Saudi citizens because the school is a public university providing free education for citizens only. The age range for the participants was between 19 and 21 years old.

All of the participants had passed the intensive English language program in their first year in the university before being admitted to the nursing program. The English program was forty weeks long, designed to enhance proficiency in the four language skills: listening, speaking, reading, and writing. Upon passing all the English classes with a minimum of a C grade and achieving a cumulative GPA of at least 3.5 of 5 in the preparatory year, the participants were admitted to the Applied Medical Sciences College which houses the Nursing Department. The students also had the option to skip the preparatory year (i.e. the intensive English program) if they achieved a score of 61 in the TOEFL IBT (Internet Based Test) or a score of 5.5 in the IELTS. Based on these requirements/scores, it was assumed that the students' level in English was intermediate.

To know more about the participants' English language learning backgrounds, the participants completed a questionnaire (Appendix B) at the beginning of the study asking them about their first language, any foreign languages they know, and how they learned them. The

questionnaire revealed the following information. All the participants reported speaking Arabic as their first language and English as a foreign language and completing the English program in the university preparatory year. Forty-five participants (90%) stated that they had taken one English course in each of their previous seven years of education (i.e. one course in every year of their intermediate and secondary/high public schools). Three students stated that they had studied one English course in each of their previous eight years while the remaining two students stated that they had studied one English course in every year of their previous education years (i.e. twelve years, the total number of study years in a private school where English instruction starts in the first grade). The majority of the participants (74%) stated that school was their only source of learning English while some participants listed studying in a private institute as an additional source for their English education, and one student stated that she had spent one year in the US to study English. Further, the majority of the participants had never visited an English speaking country with only three students who had done so for tourism. Seven participants mentioned that they had visited a foreign country (e.g. Turkey) where they had to speak English as a lingua franca for a period of time that did not exceed a few months.

As expected, the questionnaire revealed that the participants had very similar background in terms of their English learning. They comprised a fairly homogenous group with few differences among them. They all used English as a foreign language, and most of them had received the same amount of English instruction. This similar backgrounds suggests that the participants were at a similar level in terms of their English language proficiency and English writing skills at the start of the study.

Target Structures

Two target features were selected based on preliminary data collected in a visit to the university the year before the study began. During this visit, I collected and reviewed 60 samples of students' English writing taken from a pool of students similar to the target participants; some of them were students of English in their first year in the Arts College while the rest were students in the intensive English program in the preparatory year. Further, seventeen instructors in both the English bachelor program and the intensive English program completed a survey (see Appendix D) that asked them about the most common types of errors in their students' writing and the most common types of CF they use when providing feedback on their students' writing.

The writing samples and the surveys revealed that the most common writing errors in this EFL context were verb tenses, copula Be, articles, prepositions, subject-verb agreement, relative clauses, passive construction, word choice, and mechanical errors, such as spelling and punctuation. For the purpose of the study, the errors were categorized as either rule-based linguistic features or item-based linguistic features. The errors that can be treated according to a clear set of grammatical rules are categorized as rule-based features whereas the errors that are idiosyncratic and do not follow clear fixed rules are categorized as item-based features (see Bitchener et al., 2005; Ferris, 1999).

From the rule-based features, the present simple tense was selected, and the use of prepositions was selected as an item-based feature. These choices were made based on several reasons. First, both structures were addressed in the participants' previous years of learning English. It is argued that CF is more effective when targeting linguistic features that the learners began to acquire than when targeting entirely new linguistic features which the learners may not be ready to acquire (Ellis et al., 2008). Students are expected to know these structures partially

and draw on their familiarity with these structures when provided with CF on them. Second, although the simple present tense and prepositions are partially known to the students and seem simple to describe, they continue to constitute problems for EFL learners. Research has shown that copula Be, third person singular present tense –s, and prepositions cause serious problems for ESL/EFL learners even at a high proficiency level in English (Celce-Murcia & Larsen-Freeman, 1999). Thus, it is beneficial for the students to receive CF on them to further draw their attention to these problematic features and, in turn, learn more about them and hopefully acquire them. Third, errors made within these categories can be easily identified in students' writing, which provides the opportunity for more reliable data coding. Fourth, obligatory uses of these features appear in many types of writing, which provides the researcher the opportunity to examine the students' use of these features in their writing (Ellis et al., 2008).

What Constitutes an Accurate Use of the Simple Present Tense?

When providing CF on students' writing, the following criteria were used to determine when and how to provide CF on the simple present tense. The simple present is used for “general statements of facts... [and] to express habitual or everyday activity” (Azar, 1986, p. 81). Thus, whenever a sentence in students' writing referred to a habit or fact, it was decided that the present simple has to be used in this sentence. The given CF targeted the form of the main verb in those sentences. Examples include verbs in the base form or –s form, such as water *freezes* at 0 degrees centigrade; copula Be verbs, such as the world *is* round; and the verbs “has” and “have” to express possession, such as I *have* a meeting (Azar, 1986; Celce-Murcia & Larsen-Freeman, 1999; Hacker & Sommers, 2011).

What Constitutes an Accurate Use of Prepositions?

Explaining how each preposition is used in the English language can be difficult because “the differences among them are subtle and idiomatic” (Hacker & Sommers, 2011, p. 286).

While most are idiomatic, a small number are less so and can be used according to certain grammatical rules, such as the prepositions of time and place (e.g. in, at, and on). Nevertheless, even this type of prepositions can be troublesome, for example, “a person rides *in* a car but *on* a bus, plane, train, or subway” (Hacker & Sommers, 2011, p. 286). Therefore, prepositions, even the less idiomatic, do not work according to clear, fixed rules. CF was provided on all the prepositions that appeared in students’ writing on the basis that all of them are item-based features to a certain degree.

CF Types

In the pre-study visit to the university where the study was to take place, seventeen English instructors were asked in a survey (Appendix D) about the types of CF that they usually provide to their students. Sixteen instructors chose the option “I try to correct all linguistic errors” (i.e. unfocused CF) while one instructor only chose “In every assignment, I correct a limited number of linguistic features” (i.e. focused CF). The surveys also indicated that the teachers alternate between direct and indirect CF approaches when providing CF. Thirteen instructors chose “I underline the errors and provide the correct form” (i.e. direct approach) and eight teachers, some of them were the same teachers who also chose the direct approach, chose “I underline the errors without correcting them” (i.e. indirect approach). Moreover, a high percentage of the instructors indicated that they sometimes provide their students with metalinguistic CF, written and/or oral. Fifteen instructors stated that they sometimes provide their students with oral metalinguistic CF in the form of a mini-lesson provided to the whole

class or private conferences provided to single students. Eight instructors stated that they sometimes provide written metalinguistic CF in the margins of students' writing. To sum up, the survey findings suggested that students in the target context receive different types of CF. They mostly receive unfocused CF in the form of direct CF, indirect CF, written metalinguistic CF, and/or oral metalinguistic CF.

In the present study, the given treatment was in the form of focused, direct or indirect CF. Focused CF was chosen over unfocused CF for several reasons. First, focused CF is easier for learners to grasp because it requires the learner to attend to several corrections to only one or two types of linguistic features. In other words, focused CF provides the learners with “the rich evidence they need to both understand why what they wrote was erroneous and to acquire the correct form” (Ellis, 2009b). Conversely, unfocused CF requires the learner to attend to many corrections to several types of different errors, which may overwhelm the learner; for this reason, unfocused CF may not provide the learner with the opportunity to reflect in depth on every error (Ellis, 2009b). Second, research has shown that when providing focused CF to one group of learners and unfocused CF to another group, the focused group outperformed the unfocused group suggesting that focused CF is more effective in improving accuracy (e.g., Sheen et al. 2009). Further, most of the CF research that has yielded positive results for the effect of written CF had targeted one or two linguistic features only (e.g. Bitchener, 2008; Bitchener & Knoch, 2010a & b). On the contrary, most of the studies that suggested a limited or negative effect of written CF had provided their participants with mass corrections covering all errors (e.g. Polio et al, 1998; Truscott & Hsu, 2008).

As for the use of direct and indirect CF, students in the target context are accustomed to both types of CF. However, a decision was made to use direct CF with two treatment groups and

indirect CF with the other two treatment group in order to examine the potential benefit of each type of CF separately and see if one type of CF is more/less appropriate with certain linguistic features than the other type. The following section states how each type of CF was operationalized in the instruction.

Operationalizations

Direct written CF. This type of CF was operationalized as the identification of the errors through underlining and the provision of the correct form above the erroneous one.

Example: Ali speak Arabic → speaks

Indirect written CF. This type of CF was operationalized as the identification of the errors through underlining or circling. The correct forms were not provided, leaving the errors for the student to self-correct in a second draft.

Example: Ali speak Arabic.

Both types of CF, direct and indirect, were used within the focused approach. This means that for every treatment group, only one type of linguistic features was addressed through one type of CF.

Design

The present study used a quasi-experimental design to collect and analyze both quantitative and qualitative data. The first two research questions were investigated through pretest, immediate, and delayed posttest tools. The third research question was investigated through a questionnaire completed by the participants at the end of the study.

Since the writing tasks and tests (Appendix E) which were used in the study were also part of the class work, it would have been unethical to include a control group that did not

receive any type of CF. All the students in this context expect, and are accustomed to, receiving some type of feedback on their writing. To compensate for the lack of an empirical control group, the participants acted as a control group for themselves by completing two pretests with time interval between them at the beginning of the study. The purpose of the two pretests was to determine if the students' accurate use of the target features improved in pretest 2 when compared to pretest 1 before any treatment was provided.

The study was divided into two stages, a control stage and a treatment stage. In the first stage, weeks 1 to 5, the participants acted as the control group receiving no treatment and completing two pretests, the first in week 1 and the second in week 5. In weeks 5 to 12, the participants acted as experimental groups receiving three CF sessions, in weeks 6, 7, and 8, and completing two posttests, an immediate posttest in week 8 and a delayed posttest in week 12.

In the CF sessions, the participants were divided randomly into four experimental groups: A, B, C, and D. Group A received direct written CF on the simple present tense. Group B received indirect written CF on the simple present tense. Group C received direct written CF on prepositions. Group D received indirect written CF on prepositions. Table 1 describes the groups according to the given treatment and the target linguistic features.

Table 1

Study Groups

Group	Number of Students	Treatment	Target structures
Group A	13	Direct CF	Simple present
Group B	12	Indirect CF	Simple present
Group C	13	Direct CF	Prepositions
Group D	12	Indirect CF	Prepositions

Procedure for Data Collection

The study took place in one intact class and lasted for twelve weeks. Before data collection started, permission was granted by the program coordinator to invite the students to participate in the study. Since the researcher was the class instructor, a research assistant acted on behalf of the researcher in the process of explaining the study to the students and inviting them to participate in the study. All the students in the writing class gave consent to participate in the study. When explaining the project by the research assistant to the students, the students were not given any more information about the study than what was mentioned in the consent letter (Appendix A). For example, they were not told what the target features were or what types of treatment/CF they would receive.

In the first stage of the study, the participants acted as a control group. In week 1, they completed pretest 1. They also completed Questionnaire 1 (Appendix B) which was about their English learning background. The results of this questionnaire are presented earlier in this chapter. During this stage, from week 1 to week 5, the students continued their writing classes as usual, not including any writing practice, written CF on the target features, or instruction on the target features. In week 5, the students completed pretest 2. Pretest 2 was used to reveal if there was any improvement in students' writing when compared to pretest 1 before any treatment was provided.

In the second stage of the study, from week 5 to week 12, the participants acted as experimental groups. This stage started where the first stage ended. Pretest 2 was used to determine the accuracy level of students' use of the target features at the beginning of the second stage (i.e. the treatment stage). Pretest 2 was also used as the first writing task on which the

students received CF in the following week. In this stage, students were assigned randomly to one of the four experimental groups explained previously.

In the following week, week 6, all groups received back pretest 2 (i.e. task 1) with the assigned type of CF for each group. Both groups A and B received written CF on present simple tense only with the difference being that group A received direct CF whereas group B received indirect CF. Both groups C and D received written CF on prepositions only with the difference being that group C received direct CF whereas group D received indirect CF. All groups were given 10 minutes to review the given CF on their writing and think about why some words were marked as errors and how to correct them if the correct forms were not given. Then, they were given 20 minutes to rewrite task 1 in class based on the given CF. When rewriting the task, the students were instructed to copy the corrections if the corrections were provided (as in the direct CF groups) or to self-correct the underlined errors if no corrections were provided (as in the indirect CF groups); they were also told that not all errors were marked or corrected by the instructor and that they should try to improve the overall quality of their writing by considering both content and accuracy. When the students finished rewriting task 1, they were given task 2, a new piece of writing, to complete in 40-45 minutes.

In week 7, all groups received a second session of CF with the same steps which were done in the previous week. They received back task 2 with the assigned CF types, were given 10 minutes to go over the given CF, and were given 20 minutes to rewrite task 2 in class. Then, they were given task 3, a new piece of writing, to complete in 40-45 minutes. In week 8, a third session of CF was conducted, with the same procedure, followed by the immediate posttest, a new piece of writing, to complete.

From week 9 to week 12, the students did not do any writing tasks and, in turn, they did not receive any written CF or instruction on the target features. However, it is highly possible that during the experimental stage (i.e. from week 5 to week 12), the students completed and submitted written assignments for their other nursing classes, on which they received content feedback only. Receiving content feedback that addresses the main goals of the assignments is the usual practice when evaluating any assignment in this college based on what their instructors said when asked about the type of feedback they provide on students' assignments. In week 12, all groups completed the delayed posttest, a new piece of writing, in 40 to 45 minutes. They also completed Questionnaire 2 (Appendix C) asking them about their preferences of CF practices and their experience with the given CF in the study.

All the tasks, tests, and questionnaires were completed by all the students (i.e. the four treatment groups) in the same room at once. The classroom size was large, and the number of the chairs available was more than the number of the students. As such, the students were instructed to leave one chair/desk empty between every two students. They were also instructed not to talk to each other while completing the writing tasks. This was done to ensure that the students did not help each other with the given CF and to ensure that the treatment did not leak from one group another (i.e. the students who received direct corrections on a specific feature should not be able to help the students who received underlining only on the same feature). If the students needed any help or had questions, they directed their questions to the researcher/instructor. Table 2 explains the design and timeline of the study, specifically the times of the writing tests, writing tasks, and CF sessions.

Table 2***Design and Timeline***

Stage	Week	Activity
Stage 1: Control Stage	Week 1	Pretest 1 + Questionnaire 1
Beginning of stage 2	Week 5	Pretest 2 (also considered as task 1)
Stage 2: Treatment Stage	Week 6	1. CF on pretest 2
		2. Rewriting pretest 2
		3. Task 2
	Week 7	1. CF on task 2
		2. Rewriting task 2
		3. Task 3
	Week 8	1. CF on task 3
		2. Rewriting task 3
		3. Immediate posttest
	Week 12	Delayed posttest + Questionnaire 2

Tools and Instruments

The study used the following instruments: two questionnaires and six written texts. Depending on their function, four texts were called “writing tests” while the remaining two were called “writing tasks”. The writing tests, which include pretest 1, pretest 2, immediate posttest, and delayed posttest, were used to measure the students’ accurate uses of the target features, and thus they were scored/coded for analysis purposes. The writing tasks, on the other hand, were used for treatment purposes only and, thus, were not coded for analysis later; they were

completed by the students in the CF sessions, corrected by the researcher, and returned to the students in the following class meeting to examine the given CF and rewrite them.

The topics of the six written texts are given in Appendix E. These topics were adapted from two writing textbooks. The first was *Interactions 2* (Pavlik & Segal, 2007), and the second was *First Steps in Academic Writing* (Hogue, 2008). These books were chosen because they are used in the target context for students with a similar English proficiency level as the participants in the study. The first book is used for the classes of Composition 1, given to students of English in their first year in their undergraduate program. The second textbook is used in the classes of Expository Writing 1, given to nursing students in their first semester of their undergraduate program. To achieve reliability, the topics of the texts were designed to be similar, and they were administered under the same conditions (i.e. 40-45 minutes in duration for in-class writing).

As for the questionnaires used in the study, Questionnaire 1 (Appendix B) was designed to understand the students' English learning background. Although the participants came from a homogenous group, there might have been differences in their English-learning background that might be captured by Questionnaire 1. It was completed by the participants at the beginning of the study. On the other hand, Questionnaire 2 (Appendix C) was designed to elicit the students' preferences of different dimensions of written CF. It was completed by the participants at the end of the study.

The students' responses to Questionnaire 2 were used to answer the third research question. This questionnaire was composed of six closed-ended questions where the first five questions were multiple choice questions and the last question was a yes-no question followed by an open-ended question asking about the reason for the given yes/no response. In the first three questions, the students were asked about their preferences of CF types, with the first question

concerning the most effective type of CF from the students' perspective (direct CF or indirect CF) while the following two questions were concerned with the students' reactions when provided with direct CF and indirect CF on their writing and also suggested possible reasons for these reactions. The fourth and fifth questions were concerned with the types of linguistic errors on which the students want to receive either direct CF or indirect CF. The last question examined whether the students thought their writing improved as a result of the given treatment in the study and why. When designing these questions (i.e. Questionnaire 2), the following points were considered: 1. the language of the questions should be easy for the students to understand and, thus, should contain no jargon, such as direct, indirect, item-based, or rule-based features; 2. the questions should address the main goal of the study, which is treating different types of errors using different types of CF; for this reason, half the questions were about the direct/indirect contrast while the other half was about types of errors to be treated, which were mainly grammatical errors versus vocabulary errors; 3. the questions were also based on what previous research on CF preferences has investigated, such as the amount of CF, the types of CF, and the types of errors to be treated from students' perspectives (e.g. Amrhein & Nassaji, 2010; Chandler, 2003).

Pilot Study

The design of the study and the instruments were pilot tested before the beginning of the main study. In the semester that preceded data collection for the study, the pilot study was conducted in the English undergraduate program in the same institution. Permission was obtained from the chair of the English department to invite the class instructors of Composition 1 and their students to participate in the pilot study. The instructors of four intact classes of Composition 1 agreed to participate in the pilot study.

Participants

Seventy-two students registered in the class of Composition 1 gave consent to participate in the pilot study. They were students in their first year of the English bachelor program. They were all female Saudi citizens. Their level of English was mostly low-intermediate in the context of the program. All the participants had completed one English course in every year of their previous intermediate and secondary schools, a total of seven years. In addition, in their first year of the college, they had completed three English courses: General English, Reading, and Writing. Upon passing the admission English test and completing the university's first year requirements, they were admitted to the English language bachelor program.

These participants (the pilot participants) were considered similar to the participants in the main study in several ways. Both groups of participants (i.e. the pilot and actual) were students in their second year in the university and first year in their undergraduate program; therefore, the age range for both groups was the same. All of the participants in both groups were female and Saudi citizens. Also, both the pilot and the main studies were conducted in an English writing class, which would likely have made the given treatment/CF relevant, needed, and expected by the participants. Further, based on the information drawn from Questionnaire 1, all the participants had a similar English learning background in terms of their reported sources of English study. There was, however, one difference between the two groups, the English proficiency level. The pilot students had a lower level in English writing than the actual participants. Because the nursing students had spent their first university year in an intensive English language program, their level in English was between intermediate and high-intermediate. While this can be considered a limitation of the pilot sample, it was not expected to affect the results of the pilot study. The goal of the pilot study was to examine the validity of the

study design and the tasks. As such, the students' writing in the pilot study was not scored for accuracy. It was only examined to ensure that the tasks generated the use of the target features.

The class instructors were also key participants in the pilot study. Three instructors gave consent to participate in the study. They all had an MA degree in TESOL or applied linguistics. Their expertise of teaching EFL ranged from three to ten years. They were all Saudis, nonnative speakers of English. They participated in the study by giving their permission to visit their classes for data collection and by providing feedback on the writing tasks and tests in terms of their suitability to the target participants. Also, one instructor gave permission to the researcher to attend one of her composition classes for the purpose of observing and documenting class instruction and teacher-students' interaction.

Procedure for Data Collection

The procedure for data collection in the pilot study was similar to the main study in the design, length, and timeline. The pilot study lasted for twelve weeks in which the participants acted as a control group between weeks 1 and 5 and as experimental groups between weeks 5 and 12. Every class of the four participating classes represented one experimental group. Both classes 1 and 2 received CF on simple present while classes 3 and 4 received CF on prepositions. As for the type of given treatment, both classes 1 and 3 received direct CF while classes 2 and 4 received indirect CF on the target features. Further, similar to the main study, the participants completed two questionnaires; the first was about their English learning background, and the second was about their preferences of CF types and their perception of the given treatment. They also completed two pretests, two writing tasks, and two posttests. They received three CF sessions. In each session, any previous writing task was returned to the students with the

assigned treatment, the students rewrote the previous task based on the given CF, and then they completed a new task.

In addition, the pilot study included data collected from class observations. I was given permission from one instructor to attend one of her composition classes to observe and document teacher-student interactions and any behaviors related to CF, such as oral error correction. Some of the students' writing samples were also collected from the four participating classes to document the types of CF used in the target context and the common language errors in students' writing.

Results of Piloting the Design and the Instruments

The goal of the pilot study was to discover if the design had any flaws and if the instruments were suitable to the target participants and served the goal of the study. Based on the participants' responses, some features of the design and instruments were modified or changed. The modifications/changes included the content of the writing tasks, tests, and questionnaires, and the time given to complete them.

Instruments. Based on the students' responses to the writing tasks and tests and the instructors' responses to an evaluation sheet for the tasks (see Appendix F), some tasks/tests were replaced while others were only modified. The students' responses to the tasks and tests revealed many things about them. First, it was found that the majority of the students were able to generate ideas and write compositions of a reasonable length, about two thirds to one page, in response to the tasks/tests. This shows that the students were familiar with the suggested topics. What confirms their familiarity is the fact that the tasks were adapted from writing textbooks which the students either used or will use in their writing classes. Still, a small number of students did not produce much writing in response to the suggested topics, which could be

attributed to many reasons; one of them is the students' level in English. To avoid this problem in the main study, I set a minimum for the students' writing. I told the students in the main study that they had to write at least two thirds of a page in response to the writing tasks/tests.

Some students in the pilot study asked about the meaning of some words, which indicated that some of the vocabulary was above the students' level in English. Since the main study was conducted with students whose English proficiency level was higher than that of the participants in the pilot study, the vocabulary should not be a problem. However, to avoid any possible problems with vocabulary in the main study, some of the vocabulary in the prompts was changed, and the participants in the main study were allowed to use their dictionaries whenever needed. Using dictionaries, traditional or electronic, is a common practice in EFL writing classes in Saudi Arabia, making this decision ecologically acceptable. In addition, since the study did not target vocabulary use, but the use of two grammatical structures, helping the students with the vocabulary by allowing dictionary use was not expected to affect the results of the study.

The time given to complete the writing tasks constituted another challenge in the pilot study. The participants were given 30 minutes to complete each new writing task. Some students complained about the time not being enough to generate ideas and write a more developed composition. Some stated that 45 minutes to an hour was more reasonable to complete the given tasks. It was also noticed that when the participants complained about a topic being unfamiliar, they usually needed more time to complete the task. Based on this observation, 40-45 minutes was assigned to complete each new task in the main study, and the topics which were unfamiliar to the students were either replaced or modified.

The class instructors also provided feedback on the writing tasks and tests. They ranked the topics according to their suitability/difficulty for the participants. They stated that most of the

topics were suitable to the students' level except for two topics, "characteristics of good teachers" and "characteristics of good nurses". They suggested that either the two topics be changed or that the students be told in advance about the topics, a week before the class, to allow them to prepare their ideas ahead of time. However, since asking the students to prepare for writing before class can affect the results of the study significantly, those two tasks were replaced instead.

Further, the students' responses to the tasks showed that some tasks/tests were not comparable. For example, pretest 1 was about the advantages and the disadvantages of the Arts College while pretest 2 was about the advantages and disadvantages of living with parents. If the tasks and tests generated different content, the results/scores will not be comparable; thus they cannot be used to measure the students' development in writing from time 1 to time 2. As a result, pretest 1 and pretest 2 were changed to "the advantages and disadvantages of small colleges" and "the advantages and disadvantages of large colleges/universities" in the main study.

Other tasks that were changed based on the pilot study were "the characteristics of good teachers" and "the characteristics of good nurses". They were changed because of the instructors' feedback, mentioned previously, and because of the students' responses. The students' responses showed that these two particular tasks did not generate the use of one of the target features, the present simple tense. The students used more modals than simple present in these tasks. They used modals, such as "must" and "should", in almost every main sentence in their writing. For example, they wrote "Good teachers should be knowledgeable in their field" and "Good nurses must be able to act fast in emergencies". Hence, examining the students' use of the simple present and providing feedback in these tasks were not possible. The two tasks were then

changed in the main study to “the advantages and disadvantages of living with parents” and “the advantages and disadvantages of living in a student dorm”.

Furthermore, both the immediate and the delayed posttests were modified for similar reasons. The immediate posttest was about “the perfect destination for tourism”. When examining the students’ responses to this test, it was noticed that a good number of students used the past simple tense to talk about a trip they had made in the past. As for the delayed posttest which was about “the benefits of living in a foreign country”, it was found that some students used more modals than the present simple tense in their response to this topic. They wrote sentences like “if I live in a foreign country, I would/will/could...”. The two posttests were then changed in the main study to “the advantages and disadvantages of studying abroad” and “the advantages and disadvantages of studying in your country”.

As for the questionnaires, the student’s responses to Questionnaire 1 revealed that the questions were clear and the vocabulary used was easy to understand. Thus, no changes were made to the content of Questionnaire 1. However, some students asked about the difference between a question that asked if they had visited a foreign country and another that asked if they had visited an English speaking country. Based on this observation, for the main study, the research assistant was instructed to explain to the students the difference between these two questions by providing examples. As for the time given, the students were able to complete Questionnaire 1 within ten minutes; hence, the students in the main study were given the same amount of time to complete the questionnaire.

Conversely, Questionnaire 2 received some modifications. The students’ responses to an evaluation sheet for the tool (see Appendix G) and also their questions about some questions led to the change of the wording of those questions to make them clear and easy to understand. For

example, the question “From your perspective, what is the most effective way to...” was changed to “In your opinion, what is the best way to...”. Also, based on the time needed to complete this questionnaire, the students in the main study were given fifteen minutes to complete it.

After all these changes were incorporated into the instruments, the new writing tests/tasks were also pilot tested before the beginning of the study. Four students were contacted via email. They were students of English in their first year in the College of Arts, different from the ones who participated in the pilot study. They agreed to respond to the six tasks/tests and fill out an evaluation sheet (see Appendix H) for each task. Their responses to the six tasks were examined. It was noticed that the new tasks led to the use of the target features more than the old tasks/tests did, and that the new tasks generated more comparable content and use of grammatical structures than the old tasks/tests did. The evaluation sheet filled out by the four students also revealed that the suggested topics were familiar and interesting, the vocabulary used in the prompts was easy to understand, and the time given (40 minutes) was enough to complete each task. Based on these results, no further changes to the instruments were made.

Care was taken to pilot as many of the procedures and tools in the study, including the length of the writing and the time given to complete the tasks (i.e. new tasks, second drafts, and questionnaires). The final design and tools were all determined based on the pilot study. Along the same lines, it was considered important that the students produce writing in the same genre for all the CF sessions and tests, so that the writing was comparable and the scores represented their improvement based on the given treatment with similar writing.

Design. The pilot study included data collected from class observations. This was done to find out if there were any variables in instruction or teacher-student interaction that could impact

the effectiveness of direct and indirect CF on the accurate use of the target features. As explained earlier, only one teacher gave me her permission to attend and document one of her composition classes. In this one class, it was found that when the teacher provided oral CF, the focus was more on the content than on the form of the language. For example, when the students were discussing their answers to one of the exercises concerning topic sentences, the teacher provided feedback on whether the topic sentences suggested by the students were suitable. The teacher's feedback was first given to each student separately when they were answering the exercise individually. Then the teachers' feedback was given to the whole class at once when the class started discussing the exercise together as one group. The teacher also discussed the students' ideas in these topic sentences by commenting on the content being too detailed, too general, or suitable to be included in a topic sentence. Therefore, the given CF in this particular writing class was mostly related to the content and the task in question and did not address form. One possible reason is that in any oral communication in L2 classes, the attention is mostly given to the content as the available time is too short to focus on many aspects of one sentence, while in written CF the time available allows the attention to be given to both of the content and form, either at the same time or at separate times where the attention can be given to content first and then to form at another time. In the main study, however, class observations were not possible because the researcher was also the instructor which made taking notes not possible. Audio-recording the classes was also not an option because it was impossible to not record students in the class if they did not agree to participate in the study. For this reason, the main study did not include any data taken from class observations. Consequently, the third research question which asked about the possible variables affecting the students' performance in writing as a response to written CF was modified to address the impact of students' preferences as one of the possible

variables and was investigated through the students' responses to Questionnaire 2 only with no class observation data to examine.

Procedure for Data Scoring and Coding

The four writing tests, pretest 1, pretest 2, immediate posttest, and delayed posttest, were scored using the "obligatory occasion analysis" method (Ellis & Barkhuizen, 2005). The obligatory occasion analysis "constitutes a method for examining how accurately learners use specific linguistic features... it involves a comparison between the forms used by the learners and target language norms" (Ellis & Barkhuizen, 2005, p. 73). In this analysis method, all obligatory occasions of the target features are first identified in the student's text. Then the student's text is examined to determine whether or not the student used the target features in all the occasions. The score is then calculated by dividing the student's accurate uses of the target feature by the total number of the obligatory occasions. The score is expressed as a percentage.

However, since this analysis method does not account for the overuse of features, another analysis method was needed for scoring papers that contained overuse of morphemes. One of these methods is the 'target-like use analysis' method proposed by Pica (1984) (as cited in Ellis & Barkhuizen, 2005). In the target-like use analysis method, the number of the student's accurate uses of the feature is divided by the number of the obligatory occasions *plus the number of the non-obligatory occasions* (i.e. the number of the overuse of the feature).

Using the obligatory occasion method with its modification, the target-like use method, here's the procedure followed to score the four writing tests:

1. The students' writing was examined, and all the places where the target feature should be used were identified. These occasions were counted. The resulted number was the number of the 'obligatory occasions'.

2. The composition was examined again to see if a feature had been supplied for each obligatory occasion. Any errors related to the target features were circled. The student's accurate uses of the target feature were counted. This number was the number of the 'accurate uses of the target features'.

3. The composition was examined one more time to identify when the student used the target feature where it was not required (i.e. the student overused it). The overuses were counted. This was the number for 'overuse'.

4. The number resulted from step 2 (i.e. number of student's accurate uses of the target feature) was divided by the total of numbers from steps 1 and 3 (i.e. number of obligatory occasions plus number of overuse). Then the score was converted into a percentage by multiplying the resulted number by 100.

Example: the composition has 9 obligatory occasions; the student used the target feature 6 times correctly, and overused the target feature one time. In this case, the score would be: $6 \div (9+1) = .6$, and the percentage is 60%.

Scoring Present Simple Tense Tests

The steps mentioned above were followed to score papers in which the target feature was the simple present. However, since the present simple has a certain structure in which the morpheme added to the verb differs according to the subject, it was important to differentiate between failure to use any morpheme (i.e. the failure to realize that the simple present tense is

required in a certain context) and the usage of an incorrect morpheme. Dulay and Burt (1980 as cited in Ellis & Barkhuizen, 2005) proposed a scoring method for similar cases. In their method, the use of the correct morpheme was given 2 points, the use of an inaccurate morpheme received 1 point, while the failure to use any morpheme was given 0 points.

In the present study, I gave .5 point for identifying the tense accurately (i.e. understanding when to use the present simple tense) and .5 point for using the correct morpheme (e.g. using the 3rd person singular –s when required). Thus, if the student knows *when* to use the tense and *how* to use the tense accurately, she will get a full point (i.e. 1 point). For example, if the composition had 10 obligatory occasions, the student was able to identify 6 occasions but in one occasion, the student used the wrong morpheme; in this case the score would be: $5.5/10 = .55 * 100 = 55\%$. This scoring method reflects the student's possible interlanguage development as it differentiates between the student's awareness of the context that requires the use of the simple present and the student's ability to use the tense accurately with the correct morphemes.

Another decision made when scoring for the present simple tense was to ignore any errors related to the subject if the subject was clear and could be understood from context. If the student missed the subject in a sentence in which the present simple tense was used, this error was ignored, and, in turn, the score was not affected. That is because using the subject is not germane to the use of the present simple in that it does not make a distinction between tenses. What is essential to the use of the present simple in this study was the form of the verb (i.e. verb agreement). However, in sentences where the missing subject was open for interpretation and required a decision about the form of the verb (e.g. whether or not a 3rd person singular –s was needed), the form of the verb was considered an error.

Along these lines, the use of modals in the present tense, which occurs occasionally in students' writing, was also ignored. For example, "may not be able" and "can get" were both ignored when counting the obligatory occasions and when counting the accurate uses later, because modals behave differently. To avoid this complexity, the impact of modals was discounted. A sample of how present simple tests were scored is given in Appendix I.

Scoring Prepositions Tests

The same steps mentioned previously were also used for scoring writing in which the target feature is preposition use. However, unlike present simple, prepositions were scored as either correct or incorrect. This is because prepositions are morphologically simple, and, therefore, there is no distinction between knowing when to use prepositions and how to use them. If a preposition is not used in the right place, it is wrong. Also, by definition, item-based features, such as prepositions, do not always follow clear rules that explain when and how to use them. They do not require an understanding of complex meaning that would guide the learner on how to use them. The learner has either learned the use of prepositions in the context or not. As an example of how preposition use was scored, if a paper had 10 obligatory occasions, and the student used the *correct* prepositions in 7 occasions, the score would be $7/10 = .7 * 100 = 70\%$. A sample of how prepositions' tests were scored is given in Appendix J.

Scoring Reliability

To evaluate inter-rater reliability, a colleague from the MA applied linguistic program at York University re-scored 20% of the writing tests. This colleague was a native speaker of English, teacher certified, and had taught English for seventeen years. Before this second marker started the coding process, she was familiarized with the coding procedure, how and when to score, what errors to include, and what errors to ignore. She was also given writing samples that

had been scored by the researcher as illustrations. She then independently coded 20% of the writing tests. The scores given by the researcher and the scores given by the second marker were then compared to calculate the agreement percentage. Further, to establish intra-rater reliability of the scoring process, 20 texts were selected randomly from pretest 2 and rescored five months after the initial scoring. The initial scores were compared with the ones given later, and the agreement percentage was calculated.

For inter-rater reliability, the overall agreement between the final scores given by the researcher and the scores given by the second marker was 82.5%. Further, all the texts were examined to compare the obligatory occasions and the accurate uses which were identified by the researcher to the ones identified by the second marker to find out where the differences between the two scoring processes lie. The overall agreement between the obligatory occasions identified by the two markers was 87.5%, while the overall agreement between the accurate uses identified by both markers was 82.5%.

Beside the fact that scoring writing can differ from one instructor to another even if they follow the same scoring criteria and guidelines, there were other reasons that led to the different scores in this study. First, it was noticed that most of the differences in scores came from the prepositions papers. There were some preposition uses that were accepted by one marker but not by the other. For example, the preposition “for” in the sentence, “you will be more responsible for your time” was marked as an error by the second marker while the researcher accepted this use. The differences in rating prepositions come from the nature of prepositions. Since they are item-based features, they are more susceptible to be rated differently because some of the uses of prepositions are idiomatic and, thus, whatever is accepted by one marker might be not accepted by the other.

Another factor that caused some of the differences in the final scores is how often the target feature was used in a certain text. When the obligatory occasions were few in a text, a difference of one number between the two markers, either in the obligatory occasions or the accurate uses, can lead to a big difference in the final score. For example, in one text, the researcher identified ten obligatory occasions and nine accurate uses leading to the score of 90%. The second marker, on the other hand, identified nine obligatory occasions and nine accurate uses leading to the score of 100%. In that case, the additional obligatory occasion that was identified by the researcher came from this sentence, “they can show what Saudi Arabia to their friends’; the second marker ignored the missing verb and thus gave a perfect score. The numbers of the obligatory occasions were close, 9 versus 10, and the number of accurate uses were the same, yet the final scores were significantly different.

Another reason for some of the different final scores was related to differences in the calculation of overuse. For some participants, the researcher and the second marker identified the same obligatory occasions and the same accurate uses, but they identified a different number of overuse which led to different final scores. For example, in the sentence “... which could lead them improve their knowledge”, the researcher ignored the missing “to” before “improve” as the error was more related to the use of infinitive than to the use of simple present whereas the second marker marked “improve” as overuse on the ground that the verb was used in the base form, which is similar to the verb form in simple present tense. This difference led to a different final score despite agreeing on the other components (i.e. obligatory occasions and accurate uses).

The differences in scores were also related to whether or not the student’s intended meaning was clear. It was noticed that the clearer the text was in meaning and the more accurate

in grammatical structure, the more often both markers agreed on the final score and its components. On the contrary, if the text was full of errors and the meaning was ambiguous, the final score and the components were usually different between the two markers. For example, in the sentence, “many students who study faraway are more respected than other students about the culture of the other country”, the preposition “about” was considered an overuse by the first marker. Because the meaning of the sentence was not clear, the meaning or the use of “about” was not clear, and thus, was marked as overuse by the researcher whereas the second marker marked this preposition as accurate. Therefore, if the meaning of the text was not clear, it became open to different interpretations and then different scores.

Further, whether to accept conversational or informal uses of the target features in academic writing was one of the decisions that contributed to the different scores. For example, in the sentence “if they are studying in a foreign country, they will learn a new language”, “are studying” was considered as an obligatory occasion by the researcher because it refers to a general statement/habit, and thus “are studying” was marked as an error. The second marker, however, did not count this sentence in either the obligatory occasion or the accurate uses/errors on the ground that present continuous can be used in informal English to refer to general statements. This is one of the differences that may have stemmed from contrast between scoring done by a native speaker versus that of a speaker of English as a foreign language. Whereas the native speaker (i.e. the second rater) considered “are studying” as an accepted informal register, the non-native speaker of English (i.e. the researcher) considered it as progressive tense used in a context where it is not needed and, thus, an error.

For intra-rater reliability, the researcher selected 20 texts randomly from pretest 2 after five months of the initial scoring and rescored them. The overall agreement between the initial

scores and the ones given later was 90%. The 10% difference came from errors that were either missed or considered as an error in the first time but not in the second time. The percentages achieved for intra-rater and inter-rater reliability were reasonable considering the fact that scoring writing, unlike mathematics for example, can differ from one marker to another and from one time to another due to the different interpretations that might be suggested by the marker in the scoring process at different times.

When the scores were reviewed prior to analysis, whenever a difference was found between the researcher's scores and the second marker's scores or between the initial scores and the ones given later, the text was examined again to see if the score given by the second marker or given by the researcher in the second scoring process was justifiable or if the difference was the result of a mistake. If it was justifiable, then the new score, whether given by the second marker or the researcher, was used instead of the initial score.

Procedure for Data Analysis

Since the current study required both between-groups and within-groups comparisons to measure the effect of the given treatment, a factorial mixed ANOVA was chosen to address the research questions. To answer the first and second research questions, there were three independent variables, two of them were between-subjects' factors (CF type and linguistic features) and one factor was within-subjects' factor (time). Since there were three independent variables, CF type, linguistic feature, and time, a three-way mixed ANOVA was performed. In this three-way mixed ANOVA, the test scores were entered as a dependent variable while time (three levels: pretest 2, immediate posttest, and delayed posttest), CF type (two levels: direct CF and indirect CF), and linguistic feature (two levels: simple present and prepositions) were entered as independent variables.

The students' responses to Questionnaire 2 were used to answer the third research question. Quantitative data were collected from the students' responses to the close-ended questions, and qualitative data were collected from the students' comments given as a response to the last open-ended question. To analyze the quantitative data, descriptive statistics were computed for each question to find out the learners' preferences of certain CF practices; then, two-way mixed ANOVA was chosen to examine if there was a relationship between learners' preferences of CF types and their accuracy scores. In this two-way mixed ANOVA, time (three levels: pretest 2, immediate posttest, and delayed posttest) and learners' preferences of CF types (two levels: direct CF and indirect CF) were entered as the independent variables whereas students' scores in the three writing tests were entered as the dependent variable. To analyze the qualitative data, the participant's comments given as a rationale for their response to the last yes/no question were compiled and grouped according to their common themes.

Chapter Four: Data Analysis and Results

The current study investigated three questions. The first question asked about the effect of written CF on the grammatical accuracy of EFL learners' use of the target features over an eight-week period. The second question asked about the differential effect of direct CF and indirect CF on the accurate use of different types of linguistic features, rule-based and item-based linguistic features. The third and last question asked about the impact of individual differences, specifically preferences of CF types, on the effectiveness of written CF on the grammatical accuracy of students' writing.

This chapter presents the results of data analyses. All statistical analyses were carried out using SPSS 23. The chapter begins by reporting the results of initial data analysis done to determine if the levels of the four treatment groups were similar at baseline and also if the students' performances in the two pretests were stable before treatment started. Following the results of these initial tests, the chapter presents the results of the analyses carried out to investigate the three research questions.

In order to address the first and second research questions, the scores of pretest 2, immediate posttest, and delayed posttest, were subjected to three-way mixed ANOVA. To address the third research question, data collected from Questionnaire 2 were subjected to analysis that involves three steps: 1. the students' responses to the close-ended questions were subjected to descriptive statistics; 2. the students' responses (comments) to the last open-ended question were grouped according to common themes; 3. two-way mixed ANOVA was performed to see if there was a relationship between students' preferences of CF types and their accurate use of the target features in the three writing tests. Descriptive statistics followed by inferential statistics for all the analyses done are reported in this chapter.

Groups' Performance at Baseline

Before carrying out the main analyses to investigate the research questions, analysis of the two pretests was computed to check certain requirements. A one-way independent ANOVA (i.e. between-groups one-way ANOVA) was used to determine if there were any differences between the four groups at baseline (i.e. before any treatment was provided). A paired-samples t-test was used to determine if there were any statistical differences between pretest 1 and pretest 2 scores.

Group Differences at Baseline

The scores of pretest 1 and pretest 2 were subjected to analysis to determine if there were any group differences at both points of time before treatment started.

Pretest 1. Pretest 1 data were normally distributed, with skewness of $-.894$ and kurtosis of $.179$. However, there were two outliers in the data, as assessed by inspection of a boxplot for values greater than 1.5 box-lengths from the edge of the box. Both cases came from group B. They were outliers because they represented low scores compared to the rest of the group. In this case, there were certain choices to consider: remove the outliers, modify the outliers by replacing them with a value that is less extreme (e.g. the next smallest value instead), or include the outliers in the analysis anyway. Because the sample size was small, removing the data might affect the results significantly. Thus, it was decided to run the test with the outliers as they were and then run the test again with the outliers modified to a less extreme value and then compare and report both results.

When running the test with the outliers as they were, there was homogeneity of variances, as assessed by Levene's test of homogeneity of variances ($p = .19$). A one-way ANOVA revealed

that there were no statistically significant differences in pretest 1 scores between the four treatment groups at baseline, $F = 1.41$, $p = .25$. Then, another one-way ANOVA was conducted after modifying the two outliers into a less extreme value. When modifying the outliers, the output indicated that the assumption of homogeneity of variance was violated ($p = .02$). Thus, Welch ANOVA, a modified version of the one-way ANOVA, was used. It indicated that the p value is more than .05 ($p = .12$), thus leading to the same conclusion given by the one-way ANOVA when the outliers were kept as they were, which was, there were no significant differences in pretest 1 scores between the four groups at baseline.

Since both tests of the one-way ANOVA suggested the same result, it was decided to keep the outliers as valid data. Both of the following tables report the results of the one-way ANOVA when conducted with the outliers kept without changes. Table 3 reports the descriptive statistics for pretest 1, and Table 4 reports the results of one-way ANOVA analysis.

Table 3

Descriptive Statistics for Pretest 1

Treatment Groups	N	M	SD
Group A	13	86.67	8.00
Group B	12	88.45	8.84
Group C	13	92.00	7.67
Group D	12	84.63	12.42
Total	50	87.99	9.48

Table 4***One-way ANOVA analysis for Pretest 1***

	<i>df</i>	F	<i>P</i>
Between Groups	3	1.41	.25
Within Groups	46		
Total	49		

Pretest 2. Similar to pretest 1, a one-way ANOVA was performed on the scores of pretest 2 to determine if the performance of the four groups was similar or different before any treatment was provided. Data from pretest 2 was normally distributed for each group, as assessed by Shapiro-Wilk test ($p > .05$). There was homogeneity of variance, as assessed by Levene's test of homogeneity of variance ($p = .49$). The one-way ANOVA indicated that group means are not statistically significantly different ($F = 1.02$, $p = .39$). Table 5 shows the descriptive statistics for pretest 2, and Table 6 reports the results of the one-way ANOVA analysis.

Table 5***Descriptive Statistics for Pretest 2***

Treatment Groups	N	M	SD
Group A	13	88.71	8.39
Group B	12	87.91	8.38
Group C	13	82.79	11.42
Group D	12	84.42	11.22
Total	50	85.95	9.97

Table 6***One-way ANOVA Analysis for Pretest 2***

	<i>df</i>	F	<i>p</i>
Between Groups	3	1.02	.39
Within Groups	46		
Total	49		

However, because there was one outlier in pretest 2, case 14 in group B, as assessed by inspection of boxplots, a second one-way ANOVA was conducted after modifying the outlier into the next smallest value. When running the test again, the same conclusion was drawn. There was homogeneity of variance, as assessed by Levene's test ($p = .32$); and the one-way ANOVA showed that there were no statistically significant group differences in pretest 2 between the four groups ($p = .33$).

In conclusion, the two one-way ANOVAs which were conducted to compare the four treatment groups in pretest 1 and in pretest 2 suggested the same conclusion: the four treatment groups (i.e. groups A, B, C, and D) started at a similar level in the use of the target features before CF was provided.

Pretests' Differences

The students completed two pretests with a four-week time interval between them. The goal of completing two pretests was to establish a baseline and to ensure that the estimate of the baseline in terms of the students' use of the target features is accurate and consistent. Pretest 2 was used to confirm the level of the students' use of the target features and also to see if any

changes in performance took place in pretest 2 when CF was not provided between the two pretests. For this purpose, the scores of both pretests were subjected to a paired-samples t-test. As mentioned previously, pretest 1 and pretest 2 data were normally distributed. There were three outliers, two outliers in pretest 1 and one outlier in pretest 2. It was decided to run the t-test twice, one time with the outliers as they were and the second time with the outliers modified to the next smallest value. The results of both tests were then compared to see if the outliers had any effect on the results.

When running the t-test with the outliers as they were, the test indicated that there were no statistical significant differences in scores between pretest 1 and pretest 2 ($p = .101$). When running the t-test again after modifying the outliers into the next smallest value, the same conclusion was drawn as the p value was also more than .05 ($p = .66$). Since the results were not materially affected by the outliers as revealed by comparing the results of the two t-tests, the outliers were kept as valid data. In the following tables, Table 7 and Table 8, the numbers given are the results of the t-test when the outliers were kept as they were without a change.

Table 7

Paired-Samples T-Test Descriptive Statistics

	N	M	SD
Pretest 1	50	87.99	9.48
Pretest 2	50	85.95	9.97

Table 8

Paired Samples T-Test Analysis

	M	SD	T	df	p
Pretest 1- Pretest 2	2.05	8.64	1.67	49	.10

Table 7 reports the descriptive statistics for pretest 1 and pretest 2, and Table 8 reports the results of the paired samples t-test analysis. The descriptive statistics in Table 7 show that the mean score in pretest 1 was 87.99 while the mean in pretest 2 was 85.95. This small difference suggests that the baseline was stable and that the pretests' measures were reliable because they led to similar performance.

Since the t-test did not reveal statistical differences between pretest 1 and pretest 2, it was decided that pretest 2 is the pretest that will be used when carrying out the main analyses to investigate the research questions. That is because pretest 2 proceeded the treatment period immediately, and, thus, it revealed more accurately the students' ability to use the target features immediately before any treatment was provided.

Groups' Performance after Treatment

To answer the first and second research question, a three-way mixed ANOVA and pairwise comparisons using Bonferroni adjustment were computed. This section reports the descriptive statistics of the mixed ANOVA, whether and how the data met the assumptions of the mixed ANOVA and the decisions taken accordingly, and the inferential statistics of the mixed ANOVA and the pairwise comparisons.

Descriptive Statistics

Table 9 shows the descriptive statistics (the mean scores and standard deviations) for the accuracy scores for each treatment group on each of the three test occasions, pretest 2, immediate posttest, and delayed posttest.

Table 9***Descriptive Statistics for Accuracy Scores by Treatment Groups***

	CF Type	Linguistic Feature	M	SD	N
Pretest 2	Direct CF	Simple Present	88.71	8.39	13
		Prepositions	82.79	11.42	13
		Total	85.75	10.27	26
	Indirect CF	Simple Present	87.91	8.38	12
		Prepositions	84.42	11.22	12
		Total	86.16	9.84	24
	Total	Simple Present	88.32	8.22	25
		Prepositions	83.57	11.12	25
		Total	85.95	9.97	50
Immediate Posttest	Direct CF	Simple Present	89.27	6.82	13
		Prepositions	91.14	6.82	13
		Total	90.20	6.75	26
	Indirect CF	Simple Present	95.94	4.37	12
		Prepositions	84.98	12.07	12
		Total	90.46	10.49	24
	Total	Simple Present	92.47	6.60	25
		Prepositions	88.18	9.99	25
		Total	90.33	8.66	50
Delayed Posttest	Direct CF	Simple Present	91.91	6.05	13
		Prepositions	91.86	6.04	13
		Total	91.89	5.92	26
	Indirect CF	Simple Present	91.73	8.69	12
		Prepositions	87.55	7.47	12
		Total	89.64	8.21	24
	Total	Simple Present	91.82	7.27	25
		Prepositions	89.79	6.98	25
		Total	90.81	7.13	50

The descriptive statistics in Table 9 show that the total mean scores for the four treatment groups increased from pretest 2 to the posttests. It was 85.95 in pretest 2, and then increased to 90.33 in the immediate posttest and 90.81 in the delayed posttest. This suggests that the given treatment had some effect on the students' performance in the posttests. Whether this effect was statistically significant is a question answered later using inferential statistics.

One example of how the groups' performances changed throughout time is the performance of group B (shown as "indirect CF/simple present" in the Table 9). It started with a mean score of 87.91 in pretest 2; then the mean score increased to 95.94 in the immediate posttest and then dropped to 91.73 in the delayed posttest. On the other hand, group D (shown as "indirect CF/prepositions" in table 9) started with 84.42 in pretest 2; after treatment, the mean score for group D did not change much, with a mean score of 84.98 in the immediate posttest and 87.55 in the delayed posttest.

Statistical Assumptions

For a mixed ANOVA to provide valid results, the data needed to meet five assumptions: normality, no outliers, homogeneity of variances, homogeneity of covariances, and sphericity. For the normality assumption, the data had to be approximately normally distributed. This assumption was checked through skewness and kurtosis. The values of skewness and kurtosis were all between +1 and -1 at all the time points: pretest 2, immediate posttest, and delayed posttest. This indicates that test scores were normally distributed as the test was not significant.

As for whether the data from the three tests included any outliers, there were no outliers as assessed by inspection of a boxplot in pretest 2, the immediate posttest, or the delayed posttests. In the previous tests which were conducted to determine group differences at baseline and to determine pretests differences, there was one outlier in pretest 2. However, in the three-

way mixed ANOVA, the data were divided twice into two groups. First, the test scores were divided according to the target linguistic features into simple present group and prepositions group, in which no outliers were detected. Second, the test scores were divided according to the type of CF into direct CF group and indirect CF group, in which also no outliers were detected. Therefore, according to the variables used in the three-way mixed ANOVA, there were no outliers in any of the three tests.

The third assumption of mixed ANOVA is the homogeneity of variances. Levene's test of equality of error variances tests this assumption. When the test was performed, the results indicated that there was homogeneity of variances, except for one time point which is the immediate posttest. The results of Leven's test are presented in Table 10.

Table 10

Levene's Test of Equality of Error Variances

	F	<i>df1</i>	<i>df2</i>	<i>p</i>
Pretest 2	.825	3	46	.487
Immediate Posttest	5.013	3	46	.004*
Delayed Posttest	.635	3	46	.596

Table 10 shows that Levene's test was significant for the immediate posttest whereas the significance levels for pretest 2 and the delayed posttest were greater than .05. When examining the descriptive statistics in Table 9, it was evident that the indirect CF/prepositions group (i.e. group D) in the immediate posttest had more variations compared to the other groups, with a standard deviation of 12.07. This means that the homogeneity of variances assumption was violated because of one group in one of the variables, group D in the immediate posttest. One

way to deal with the data when one of the assumptions is violated is to transform the data. However, this solution is not always successful. That is because we cannot transform the data for one particular group at one specific time without transforming the data of all groups at all times. While transformations may fix the one variable that did not meet the homogeneity assumption, it may damage the other data that is already normally distributed and meet the homogeneity of variances assumption. To solve this problem, it was decided to run the mixed ANOVA anyway for two reasons. First, since the group sample size, whether divided according to CF type, linguistic feature, or both, were approximately equal, we can run the three-way mixed ANOVA anyway because it is somewhat robust to heterogeneity of variance in these circumstances. Second, after running the mixed ANOVA, we can look at the results of the Box's test of equality of covariance matrices to check its significance level. If it is not significant, this means that the results of the mixed ANOVA are valid and can be interpreted despite the fact that one of the variables did not meet the homogeneity of variances assumption.

After running the mixed ANOVA, there were two more assumptions to be checked before being able to interpret the results of the ANOVA, homogeneity of covariances and sphericity. As mentioned above, homogeneity of covariances can be tested with Box's test of equality of covariance. The results showed that the test was not significant ($p = .462$), which means that the assumption had been met and the results of the ANOVA are valid and can be accepted. As for sphericity, it was tested with Mauchly's test for sphericity. The results of this test showed that sphericity had not been violated ($p = .935$).

Inferential Statistics

Table 11 presents the results of the tests of within-subjects effects. The three-way mixed ANOVA showed that there was a statistically significant interaction between the three variables:

time, CF type, and linguistic feature on the grammatical accuracy of students' use of the target features, $F(2,92) = 5.334$, $p = .006$, partial $\eta^2 = .104$. The three-way interaction of time, CF type, and linguistic feature means that the simple two way interactions of CF type and linguistic feature were different at the different levels of time.

Table 11

Three-way Mixed ANOVA: Treatment and Accuracy [Tests of Within-Subjects Effects]

	<i>df</i>	F	<i>p</i>	η^2
Time	2	10.31	.000*	.183
Time * CF Type	2	.81	.447	.017
Time * Linguistic Feature	2	.77	.466	.016
Time * CF Type * Linguistic Feature	2	5.33	.006*	.104
Error(Time)	92			

To determine where the differences lie between the treatment groups and at what times, a post hoc pairwise comparisons using Bonferroni adjustment was computed. Table 12 presents the results of the pairwise comparisons that compare the four treatment groups to each other at every time point (i.e. in every writing test).

Table 12***Pairwise Comparisons for the Treatment Groups***

CF Type	Linguistic Feature	Time (I)	Time (J)	Mean Difference (I-J)	SD	<i>p</i>
Direct CF	Simple Present	2	1	.56	2.35	1.000
		3	1	3.20	2.25	.483
			2	2.64	2.29	.763
	Prepositions	2	1	8.35	2.35	.003*
		3	1	9.08	2.25	.001*
			2	.72	2.29	1.000
Indirect CF	Simple Present	2	1	8.03	2.45	.006*
		3	1	3.82	2.34	.329
			2	-4.22	2.38	.249
	Prepositions	2	1	.57	2.45	1.000
		3	1	3.13	2.34	.561
			2	2.57	2.38	.859

The analyses in Table 12 reveal that groups B and C (i.e. indirect CF/simple present group and direct CF/prepositions group in Table 12) were the only groups which achieved statistically significant improvement in the accurate use of the target features over time. For group B, the increase in accuracy was statistically significant between the pretest and the immediate posttest (*Mean difference* = 8.03, *SE* = 2.45, *p* = .006). However, the performance of group B in the delayed posttest did not reach the significant value when compared to the pretest,

$p = .329$, or the immediate posttest, $p = .249$. Group C, on other hand, achieved a statistically significant increase in accuracy between the pretest and immediate posttest (*Mean difference* = 8.35, $SE = 2.35$, $p = .003$) and also between the pretest and delayed posttest (*Mean difference* = 9.08, $SE = 2.25$, $p = .001$). As for the difference between the immediate posttest and the delayed posttest for group C, the pairwise comparisons revealed no significant difference between the two posttests ($p > .05$) which means that the students' accuracy levels gained in the immediate posttest were retained in the delayed posttest.

Conversely, groups A and D (i.e. direct CF/simple present group and indirect CF/prepositions group in Table 12) did not achieve a statistically significant improvement in their use of the target features between pre-intervention and post-intervention ($p > .05$). However, when examining the profile plots, given in Figures 1, we can see some improvement for both groups between pre-intervention and post-intervention. Also, the profile plots showed that the performance lines for groups A and D are almost parallel. This suggests that they performed similarly from the pretest to the immediate and delayed posttests. They both did not improve much in the immediate posttest, but then they achieved higher accuracy scores in the delayed posttest; these higher scores, however, did not reach the significance level.

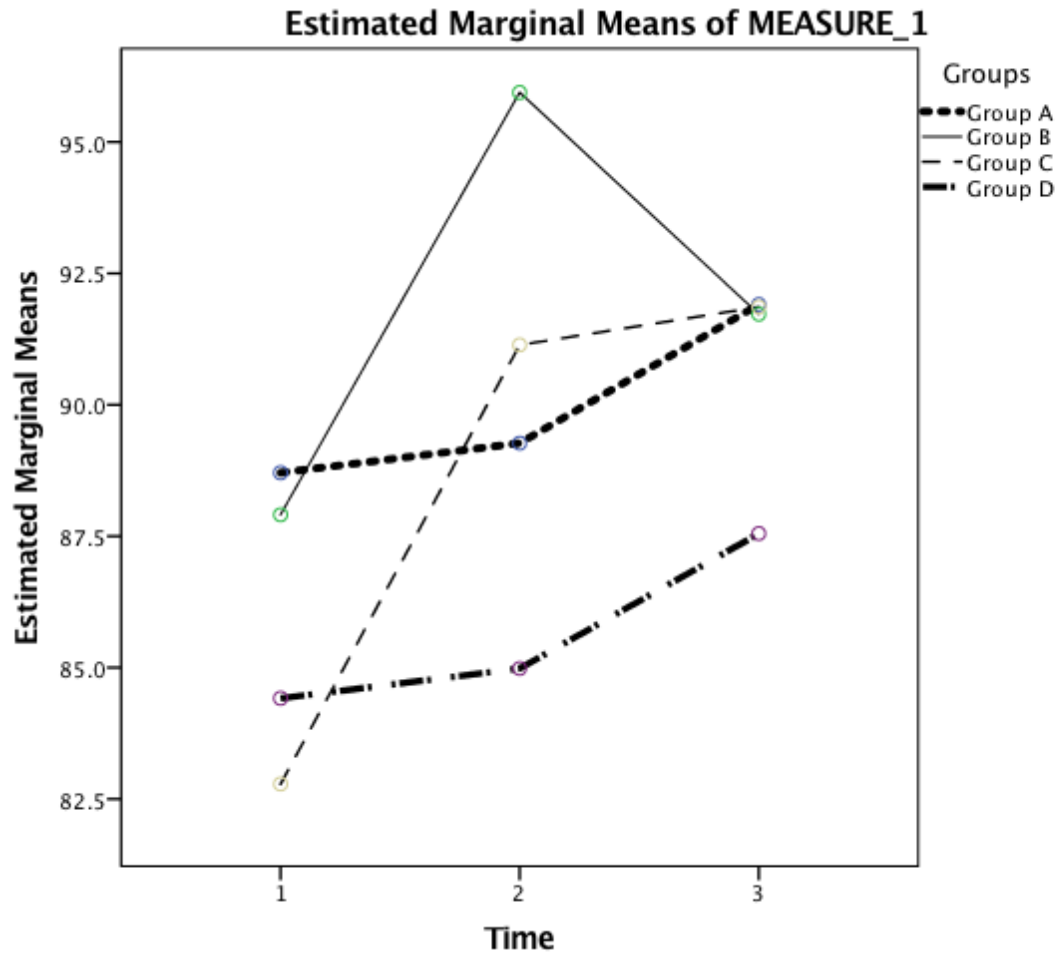


Figure 1. Profile Plots for the Four Treatment Groups

Since there is an interaction effect of CF type, linguistic feature, and time, the main effects of the single variables are compromised. However, there were some interesting effects that need to be examined further. First, the main effect of time showed a statistically significant difference in accuracy of the use of the target features across the three time points for the participants when all groups are combined in one group, $F(2, 92) = 10.309, p < .0005$, partial $\eta^2 = .183$. To know where these differences in time lie, the pairwise comparisons table for time, Table 13, was consulted. Post hoc analysis with a Bonferroni adjustment revealed that accuracy

increased statistically significantly from pre-intervention to immediate posttest (4.38 (95% CI, 1.39 to 7.36), $p = .002$), and from pre-intervention to delayed posttest (4.81 (95% CI, 1.96 to 7.66), $p < .0005$), but not from immediate posttest to delayed posttest (0.43 (95% CI, -2.47 to 3.33), $p = 1.000$). This indicates that accuracy scores improved from pretest 2 to the immediate posttest. Then, the level of improvement in accuracy stayed roughly the same from the immediate posttest to the delayed posttest.

Table 13

Pairwise Comparisons for Three Time Points: Pretest 2, Immediate, and Delayed Posttests

(I) Time	(J) Time	Mean Difference		p	95% Confidence Interval for Difference	
		(I-J)	Std. Error		Lower Bound	Upper Bound
2	1	4.38	1.20	.002*	1.39	7.36
3	1	4.81	1.15	.000*	1.96	7.66
	2	.43	1.17	1.000	-2.47	3.33

Another main effect worth mentioning is the effect of the linguistic feature, see Table 14: tests of between-subjects effects. The table shows a low p value for the linguistic feature, $F(1,46) = 3.702$, $p = .061$, partial $\eta^2 = .074$. While the p value is not significant, its low value may indicate that the target features may have affected the accuracy scores at least at one time point. When consulting the pairwise comparisons for the linguistic feature, it was found that in the immediate posttest, there was a statistically significant difference ($p = .002$) between the simple present and the prepositions groups who received indirect CF (i.e. groups B and D) with group B (i.e. indirect CF/simple present) outperforming group D (i.e. indirect CF/prepositions).

Table 14***Three-way Mixed ANOVA: Treatment and Accuracy [Tests of Between-Subjects Effects]***

	<i>df</i>	F	<i>p</i>	η^2
CF Type	1	.07	.791	.002
Linguistic Feature	1	3.70	.061	.074
CF Type * Linguistic Feature	1	1.51	.225	.032
Error	46			

Students' Preferences for Different CF Practices

To answer the third research question, the data of Questionnaire 2 were subjected to analysis. First, descriptive statistics of the quantitative data (frequencies, percentages, and bar charts) were computed for each close-ended question to form a profile of students' preferences of certain CF practices related to the amount of CF, the type of CF, and the type of errors to be treated in written CF. Second, a two-way mixed ANOVA was computed to see if there was a relationship between students' preferences of CF types and their accuracy scores in the writing tests. Third, the comments provided by students as a response to the last open-ended question were compiled and analyzed to understand the students' preferences of CF types and their reasons for these preferences.

Students' Preferences

Descriptive statistics was computed to examine the student's responses to each question in Questionnaire 2. Table 15 reports the descriptive statistics (frequencies and percentages) for the responses of each question.

Table 15***Descriptive Statistics for the Students' Responses to Questionnaire 2***

Questions	Answers	Frequency	%
1. In your opinion, what is the best way of providing corrective feedback on your writing?	a. Underline and correct all errors	20	40%
	b. Underline all errors without correcting them	11	22%
	c. Underline and correct only the most serious (difficult) errors	17	34%
	d. Underline only the most serious (difficult) errors without correcting them.	2	4%
2. What is your reaction to the teacher's feedback when she underlines your error(s) and also provides the correct form?	a. You like it because it provides you with the correct form immediately.	31	62%
	b. You don't like it because you want to figure out the correct form by yourself.	19	38%
3. What is your reaction to the teacher's feedback when she underlines your error(s) but does not provide the correct form?	a. You like it and you try to figure out the correct form by yourself.	21	42%
	b. You don't like it but you still try to figure out the correct form.	28	56%
	c. You don't like it and you don't try to figure out the correct form.	1	2%
4. What are the error types that you want your teacher to underline and correct?	a. All errors.	16	32%
	b. All grammatical errors.	14	28%
	c. Only the grammatical errors that are difficult for students at my English level.	18	36%
	d. Vocabulary errors.	2	4%
5. What are the error types that you want your teacher to underline only without providing the correct form?	a. All errors.	11	22%
	b. All grammatical errors.	11	22%
	c. Only the grammatical errors that are difficult for students at my English level.	4	8%
	d. Vocabulary errors.	24	48%
6. Do you feel that your writing has improved as a result of the given corrective feedback?	a. Yes	43	86%
	b. No	7	14%

The first question in Questionnaire 2 asked about the most effective type of CF from students' perception and provided four options for the students to choose from. Before examining the results, I will explain how each option in the first question was coded as a specific CF type. The first option "underline and correct all errors" was coded as unfocused, direct CF while the second option "underline all errors without correcting them" was coded as unfocused, indirect CF. The third option "underline and correct only the most serious errors" was coded as focused, direct CF while the fourth option "underline only the most serious errors without correcting them" was coded as focused, indirect CF. While the treatment provided in the study was only focused, the given responses to the first question in Questionnaire 2 addressed both of the focused and unfocused CF approaches. This was done because previous research showed that the majority of the students in different contexts, EFL and ESL, expressed a strong preference for CF addressing all their errors (i.e. unfocused CF). Consequently, by not including the unfocused CF options in the question, the students might be forced to choose a type of CF (i.e. focused CF) that they may not believe is effective.

The descriptive statistics in Table 15 show that for the first question, which was about the types of CF the students prefer to receive on their writing, 40% of the participants selected the unfocused, direct CF approach, and 34% selected the focused, direct CF, while only 22% preferred the unfocused, indirect CF approach, and 4% preferred the focused, indirect CF approach. To examine the relationship between students' preferences of direct/indirect CF and their accuracy scores in the writing tests, it was decided to categorize the responses to the first question into two groups. Those who selected one of the following options, a) underline and correct all errors, or c) underline and correct only the most serious errors, were included in the group of direct CF preference, while those who selected one of the following options, b)

underline all errors without correcting them, or d) underline only the most serious errors without correcting them, were included in the group of indirect CF preference. According to this new categorization, 74% preferred the direct CF approach while only 26% preferred the indirect CF approach.

The second question in the questionnaire was about the students' reaction to direct CF. As expected from their responses to the first question, the majority of the students (62%) stated that they like it, while the remaining 38% stated that they do not like it. Although the response to this question (62%) does not entirely match the response to the first question (74%), it still shows a positive attitude towards direct CF. As for why some students gave a different response to the second question which then led to the different percentages, a possible reason is the way the second question was formulated. The second question did not only ask whether or not the student preferred direct CF, but it also suggested possible reasons for this preference. It is probable that while some students had a positive attitude towards direct CF, they still chose the second option which stated, "I don't like [direct CF] because I want to figure out the correct form by myself", because of the reason it suggested, self-learning, which might seem more appealing to some students than the reason presented by the first (direct CF) option, which was "because it provides the correct form immediately".

The third question was about the students' reaction to indirect CF. Their responses showed that most of them (58%) did not have a positive attitude towards indirect CF. Yet, they chose the second option which stated that despite their negative attitude, they still try to respond to indirect CF by trying to figure out the correct form, except for one student who selected the last option which states "I don't like it and I don't try to figure out the correct form" reflecting a very negative attitude. These responses reflect a negative attitude towards indirect CF in this

EFL context, but they also show the students' readiness to self-correct their errors if provided with indirect CF despite their dislike to this type of CF.

The fourth question concerned the type of errors that the students want their teacher to underline *and* correct (i.e. receive direct CF on them). The students' responses to this question were distributed almost equally on the first three options. Eighteen students (i.e. 36%) selected "only the serious grammatical errors" option, sixteen students (i.e. 32%) selected "all error types", while fourteen students (i.e. 28%) selected "all grammatical errors". The remaining two students preferred to receive direct CF on vocabulary errors only. When the responses to the second option "all grammatical errors" and the third option "only the serious grammatical errors" were combined, it was found that 64% of the participants preferred to receive direct corrections on their grammatical errors. The remaining students preferred direct CF on all types of errors, except for two who preferred direct corrections on vocabulary errors only.

The fifth question was similar to the previous one with the difference that it concerned the type of errors that the students want their teacher to only underline but not correct (i.e. receive indirect CF on them). Most of the students (48%) chose vocabulary errors. The remaining students were divided between "all error types" (22%) and "all grammatical errors" (22%), while only four students chose "only the serious grammatical errors". The students' responses to this question stand in contrast to the previous one in one option, "vocabulary errors". Unlike the previous question, which showed that the majority of students prefer to receive direct CF on their grammatical errors, this question show that the majority of students prefer to receive indirect CF on their vocabulary errors.

The last question investigated whether the students thought their writing had improved as a result of the given treatment and why. In response to this question, 86% of the students, (n=43),

responded yes, whereas the remaining seven students responded no. All of the students provided comments to explain their yes/no response and what they thought of the given treatment. The majority of the students addressed at least one of the following aspects in their comments: the amount of the given treatment (focused CF addressing not many errors), the type of the given treatment (direct CF or indirect CF), and/or the types of errors targeted.

When reading the comments given by the seven students who thought their writing did not improve and the treatment they received was not useful, all the given reasons were related to the approach used. Two students, from the seven, received direct CF on simple present while they appeared to prefer indirect CF. Thus, one of them wrote, “I want to figure out the correct form by myself...”, and the other wrote, “underline[ing] all errors without [giving the] correct form is [a] useful way... but some grammar mistakes, it’s better to give the correct answer”. Another student received indirect CF on simple present, but because she preferred to receive direct corrections instead, she wrote, “underlining the errors without correction is not useful; sometimes I don’t know what is wrong... it doesn’t help me to improve my English writing”.

Further, three students from the group who thought the given treatment was not useful came from the prepositions groups, C and D. Two of them commented on the approach being too focused, correcting only prepositions, while they preferred the unfocused approach. One student wrote, “not a lot of errors mentioned by the teacher”, and another wrote, “it’s not useful for me because I got some corrections of my errors and the other is not”. One student who was from group C, direct CF/prepositions, wanted more than direct corrections, she seems to ask for metalinguistic CF in her following comment, “because I don’t know *why* it is wrong”.

The rest of the participants who responded to the last question with “yes” also presented reasons related to the approach used being effective or not. Some of the comments from the

direct CF groups were, “it’s very helpful because it helps me to understand why it’s wrong and know how to correct it”, “it is useful because I don’t spend much time to figure the correct answer, and may be my correction is wrong if I correct it”, and, “the teacher gave me corrective feedback on some errors which I can’t figure out the correction; so when she gave me the correction, I became able to know how to fix it”.

Some students from the indirect CF groups were also appreciative of the approach they experienced, especially those who came from group B, indirect CF/simple present. One student wrote, “it helps me to think more about the grammar and my mistakes when I’m doing another writing”. Another student wrote, “without [providing] the correct form, I was encouraged to figure out my mistakes and try to solve them. It makes me independent. Also when I figure out the correction, I will not forget them and they stick in my mind”. These reasons suggest that students think that the cognitive effort spent to figure out the corrections can lead to long-term learning.

However, one student from group D, indirect CF/prepositions, did not think that the approach she experienced was the most effective way, she wrote, “I learned more about ‘in’, ‘on’, ‘of’, and the correct place, [but] the next time I wrote the same mistake and even the second time I still don’t know the right answer”. Another student wrote in her comment, “it [indirect CF] is not good only if I can’t figure out [the corrections of] some errors”, which suggests that some students do not prefer indirect CF because it frustrates them when they do not know what’s wrong in their writing or how to correct their errors.

Some students commented on how the given CF made them focus more on the grammaticality of their writing. One student commented, “I learn to focus on some details of grammar ... and sometimes think about grammar”, while another wrote, “... it makes me focus

on my grammatical errors”. On the other hand, there were other students who wanted more than focusing on grammar, such as the following student who wrote, “it was effective for grammar... but I don’t think that I improved in terms of vocabulary... vocabulary errors should be underlined and given the correction immediately.”

Other students commented on the focused approach used in the study. Some thought that focusing on one error type only is not the most effective method to provide CF. One student wrote, “the teacher is always focusing on prepositions. I don’t like that because I want to know about all my errors, and I want all the correction.” The students’ negative attitude towards the focused approach might be justifiable for the reasons some students suggested. There were several students who suggested that the focused approach used in the study was confusing. It focused on one type of errors and left the remaining errors in the text without any marking, which made the students not sure if their writing was accurate or if they had more mistakes left for them to discover. One student complained, “I didn’t receive much underlines..., so I’m wondering if my writing was good or bad”, and another commented, “... Sometimes the teacher doesn’t correct all the errors. So I think my writing is alright, but maybe it’s full of mistakes.” One student suggested a solution to avoid confusing the students by correcting some errors and leaving the rest, she wrote, “... It will be much better to put a final mark, so when I know my mark, I will know there are mistakes I have to discover”. Conversely, there were some students who appreciated the focused approach. A student from group A, direct CF/simple present, wrote, “... the teacher corrected the basic things to me, but not all, I like this way because at the same time she let me try to figure out the other mistakes.”

Relationship between Students' Preferences of Direct CF or Indirect CF and Accuracy

The accuracy scores were subjected to analysis to see if they had been affected by the students' preferences of direct and indirect CF. For this purpose, the participants' responses to the first question in Questionnaire 2, which states "In your opinion, what is the best way of providing corrective feedback on your writing?", were subjected to two-way mixed ANOVA. As mentioned previously, those who selected "a" or "c", both describe the direct approach, were put into the group of direct CF preference, while those who selected "b" or "d", both describe the indirect approach, were put into the group of indirect CF preference.

Descriptive statistics. Table 16 shows the descriptive statistics for the two groups, the group of direct CF preference and the group of indirect CF preference. The descriptive statistics do not show big differences between the two groups in the three writing tests. For example, in the immediate posttest, the direct CF preference group had a mean score of 90.05 while the indirect CF preference group had a mean score of 91.11.

Table 16***Descriptive Statistics for Accuracy Scores by Preference Group***

	Preference of CF Types	Mean	SD	N
Pretest 2	Direct CF	85.78	9.75	37
	Indirect CF	86.42	10.97	13
	Total	85.95	9.97	50
Immediate Posttest	Direct CF	90.05	8.22	37
	Indirect CF	91.11	10.13	13
	Total	90.33	8.66	50
Delayed Posttest	Direct CF	91.37	7.00	37
	Indirect CF	89.21	7.83	13
	Total	90.81	7.13	50

Statistical assumptions. Data were normally distributed. The values of skewness and kurtosis were all between +1 and -1 at all the time points/tests. There was homogeneity of variances at the three time points, as assessed by Levene's test of homogeneity of variance ($p > .05$). There was homogeneity of covariances, as assessed by Box's test of equality of covariance matrices ($p = .838$). Mauchly's test indicated that the assumption of sphericity was met for the two-way interaction ($p = .650$). These tests indicate that the assumptions of mixed ANOVA were met, and thus, the mixed ANOVA was appropriate for the data in question.

Inferential statistics. Table 17 shows the results of the two-way mixed ANOVA analysis, the tests of within-subjects effects.

Table 17***Two-way Mixed ANOVA: CF Preferences and Accuracy [Tests of Within-Subjects Effects]***

	<i>df</i>	F	<i>p</i>
Time	2	6.50	.002*
Time * Preferences	2	.79	.456
Error(Time)	96		

The mixed ANOVA, Table 17, revealed that there was no statistically significant interaction between students' preferences of CF types and time on accuracy scores in pretest 2, immediate posttest, and delayed posttest, $F(2, 96) = .79, p = .456$. However, the main effect of time showed a statistically significant difference in accuracy at the different time points, $F(2, 96) = 6.50, p = .002$, partial $\eta^2 = .119$. To know where these differences lie, the pairwise comparisons table for time was consulted. Post hoc analysis with a Bonferroni adjustment revealed that accuracy scores increased statistically significantly from pretest to immediate posttest (4.48 (95% CI, .81 to 8.15), $p = .012$), and from pre-test to delayed posttest (4.19 (95% CI, .89 to 7.48), $p = .008$), but not from immediate posttest to delayed posttest (-0.29 (95% CI, -3.66 to 3.08), $p = 1.00$). This indicates the students improved in their use of the target features in both posttests in comparison to where they started from (i.e. pretest 2). They improved from the pretest to the immediate posttest and then retained the same level of improvement in the delayed posttest.

On the other hand, the main effect of preference of CF types, which is the results of the tests of between-subjects effects in Table 18, showed that there was no statistically significant difference in accuracy between the group of direct CF preference and the group of indirect CF

preference, $F(1, 48) = .004, p = .947$, which could suggest that students' preferences of direct and indirect CF did not affect the accuracy of their use of the target features. In other words, it appeared according to these results that whether the students preferred direct or indirect CF had no relationship with how they performed in writing in response to the given CF.

Table 18

Two-way Mixed ANOVA: CF Preferences and Accuracy [Tests of Between-Subjects Effects]

	<i>df</i>	F	<i>p</i>
Preferences	1	.004	.947
Error	48		

Summary of the Results

The following table, Table 19, presents a summary of the results in relation to the three research questions. It summarizes the test(s) used to investigate each of the research questions and the results of those tests.

Table 19***Summary of the Results***

Research Question	Test	Results
RQ 1: What is the effect of written CF on the grammatical accuracy of EFL learners' use of the target features over an eight-week period?	One-way ANOVA (two pretests)	No statistically significant differences between the four groups in both pretests.
	T-test (two pretests)	No statistically significant differences between the two pretests.
	Three-way mixed ANOVA (three writing tests)	Statistically significant interaction between the three variables (time, CF type, and linguistic feature type) on the accuracy of the target features. Statistically significant effect for time.
RQ 2: What is the differential effect of direct CF and indirect CF on the accurate use of different types of linguistic features, rule-based and item-based linguistic features?	Post hoc pairwise comparisons using Bonferroni adjustment	Groups B (indirect CF on simple present) and C (direct CF on prepositions) achieved a significant improvement in accuracy over time. As for groups A (direct CF on simple present) and D (indirect CF on prepositions), there were no significant differences between pre-intervention and post-intervention.
RQ 3: What is the impact of individual differences, specifically preferences of CF types, on the effectiveness of written CF on the grammatical accuracy of students' writing?	Descriptive statistics and qualitative analysis (responses to Questionnaire 2)	Students have different preferences of various CF practices. While some of them seem to be based on their limited learning experiences, others reveal how students view certain CF practices and should be addressed and known by their teachers.
	Two-way mixed ANOVA (three writing tests)	No statistically significant interaction between students' preferences of CF types and time on accuracy. Statistically significant effect for time. No statistically significant effect for CF preference.

Chapter Five: Discussion and Conclusions

The present study investigated three research questions. The first question concerned the effect of written CF on the grammatical accuracy of EFL learners' use of the target features over an eight-week period. Analysis of the accuracy scores achieved in three writing tests suggested that written CF had a positive effect on the grammatical accuracy of the use of the target features. The second question concerned the differential effect of direct CF and indirect CF on the accurate use of different types of linguistic features, rule-based and item-based linguistic features. Statistical analysis showed that direct CF led to a statistically significant improvement in the grammatical accuracy of the use of prepositions in both immediate and delayed posttests, whereas indirect CF led to a statistically significant improvement in the grammatical accuracy of the use of simple present in the immediate posttest only. There was also improvement in the use of the simple present in the delayed posttest for the group provided with indirect CF, but the improvement was not statistically significant. The last research question concerned the impact of individual differences, specifically students' preferences of CF practices, on the effectiveness of the given CF on the grammatical accuracy of students' writing. The results of Questionnaire 2 suggested that students have different preferences regarding types of CF and types of linguistic errors to be treated; however, analysis results indicated no relationship between students' preferences for different CF types and their accurate use of the target features.

This chapter presents the findings, the possible conclusions that may be drawn from them, and the factors that might have affected and led to these findings. The results are discussed in light of previous CF research and related theories. The chapter concludes with the limitations of the study, teaching implications, and recommendations for future research.

Effect of CF on Accuracy

The first research question concerned the effect of written CF, regardless of its specific type, on EFL learners' accurate use of the target features over time. Analysis of the students' scores in pretest 2, the immediate posttest, and the delayed posttest indicated that written CF had a positive effect on the grammatical accuracy of the target features across the three time points. Accuracy improved significantly from the pretest to the immediate posttest, and the gains in accuracy were retained a month later in the delayed posttest. These results suggest that written CF has a positive effect on the accurate use of simple present tense and prepositions

The positive results of the current study are in line with several previous CF studies (Bitchener, 2008; Bitchener & Knoch, 2010a; Ellis et al., 2008; Ferris et al., 2013; Sheen, 2007; Sheen, et al., 2009; Van Beuningen et al., 2012). These studies examined the effect of different types of CF, direct, indirect, focused, and unfocused, and the findings indicated the same result: written CF had a positive effect on the accurate use of the target features, not only in revised writing, but also in new pieces of writing, which can suggest long-term benefits. Since the target features in the current study were different from those in the above mentioned studies, the present study adds to the current evidence that supports the effectiveness of CF.

The study suggests that CF can affect linguistic features other than articles, the most commonly addressed target feature in recent CF studies. Written CF can positively affect the accurate use of simple present and prepositions. These two types of features are different from articles in many ways. First, the use of simple present involves many rules, such as the use of copula Be, the use of third person singular –s or no –s for first and second person singular and plural, and the use of “have” to express possession. On the other hand, only two functions of the article system have been targeted in many recent studies, “a” for first mention and “the” for

subsequent mention (e.g. Ellis et al., 2008). Second, articles are rule-based items, whereas prepositions, one of the target features in the present study, are item-based. These comparisons suggest that CF can lead to improvement in the use of not only rule-based features when CF targets two functions only, such as the two main functions of the article system, but also rule-based features when CF targets many functions, such as all the uses and forms of simple present, as well as item-based features, such as prepositions, which do not function according to clear grammatical rules.

The improvement in the use of the target features in the study may suggest a support for the noticing hypothesis (Schmidt, 1995; 2001), although noticing was not in itself investigated. Schmidt (1995; 2001) in his hypothesis argued that it is only through conscious attention that language learning can take place; thus noticing is essential for language learning. Based on this hypothesis, focus-on-form interventions, such as CF, are also important and facilitative for second language acquisition (DeKeyser, 1994). This is because they can help to direct the learner's attention to certain aspects in L2 that are needed to be learned. The fact that the participants showed gains in the use of the target features suggests that the given CF, whether a combination of underlining and direct correction or only underlining, may have triggered the students' noticing, leading the students to compare their current language to the target language and then to use these features more accurately in their new pieces of writing.

However, when the participants are divided into groups that considered both treatment and feature, the results indicated that although all the groups improved in the use of the target features from the pretest to the two posttests, not all of them had achieved statistically significant gains. In the use of simple present, group B, which received indirect CF, outperformed group A, which received direct CF. In the use of prepositions, group C, which received direct CF,

outperformed group D, which received indirect CF. There are a number of possible reasons that might explain why the four treatment groups achieved different levels of accuracy; these reasons are discussed in the coming section.

Another interesting observation that emerged from this study is that, written CF can lead to accuracy improvement that cannot be achieved through writing practice alone. This result came when comparing the performance of the experimental groups after treatment (i.e. in the posttests) to their performance before treatment (i.e. in the pretests). The students completed two pretests one month apart to compensate for the lack of a control group. The purpose of the two pretests was to investigate if the students' use of the target features improved in pretest 2 compared to pretest 1 as a result of writing practice alone. In the month separating the two pretests, although the students did not complete any writing tasks in the experimental class, they continued to participate in writing activities in their usual writing classes and they also continued to attend their other nursing classes in which they probably submitted a number of written assignments. It is suggested in the literature that writing practice alone, without CF, might be enough to improve both writing and second language acquisition (e.g. Robb et al., 1986; Truscott, 1996). To know if the writing practice which took place during the month separating the two pretests had any impact on the students' grammatical use of the target features, the scores of the two pretests were compared and analyzed. Analysis of the scores showed no significant differences between the pretests. This result indicates that writing practice did not appear to lead to any improvement in the use of the target features. On the contrary, most of the students achieved lower scores in pretest 2 compared to pretest 1. One possible reason for this might be that the students did not receive pretest 1 back with CF, which may have given the students the impression that they only needed to complete their writing tasks and not worry about

the grammatical accuracy of their writing because there would be no follow up or any coming CF addressing their errors. Therefore, when comparing how the students performed in the posttests after CF was provided to how they performed in pretest 2 when no CF was provided on pretest 1 or any writing tasks during the month separating the two pretests, the results suggest that CF can lead to improvement in the use of the target features whereas writing practice failed to do so. This result echoed the results of both Van Beuningen et al. (2008) and Van Beuningen et al. (2012) in which the writing practice groups did not achieve accuracy improvement in the immediate or delayed posttests unlike the two experimental groups which received either direct CF or indirect CF.

Effect of Different Types of CF on the Acquisition of Different Types of Linguistic Features

The focus of the second research question was the differential effect of direct CF and indirect CF on the accurate use of rule-based and item-based linguistic features. Analysis of the scores of the writing tests showed that there was an interaction effect between time, type of linguistic feature, and type of CF. The interaction of the three variables indicates that accuracy was affected differently by the different combinations of CF type and linguistic feature over time. This means that when a group made an improvement in the use of the target feature, the reason was not the CF type alone or the type of target linguistic feature alone, but a combination of both.

The results of additional analysis showed that group B (indirect CF on simple present) and group C (direct CF on prepositions) achieved statistically significant gains in the use of the target features while group A (direct CF on simple present) and group D (indirect CF on prepositions) failed to do so. The interaction effect among the study variables suggests that the reason for the different results achieved by the four groups was the given treatment: those who

achieved significant gains received the appropriate type of CF for the type of the target feature while those who failed to do so received a type of CF that might be not appropriate for the target features. Therefore, to respond to the second research question, the findings suggest that direct and indirect CF have differential effects: in this study, indirect CF was more effective than direct CF when targeting a rule-based feature, simple present, while direct CF was more effective than indirect CF when targeting an item-based feature, prepositions.

Immediate posttest scores for Group B, which received indirect CF on simple present, indicated statistically significant gains while those of group A, which received direct CF on the same target features, were not statistically significant. Simple present tense is a rule-based feature; it works according to a clear set of grammatical rules. Receiving indirect CF on something that is governed by clear rules appears to be sufficient and more effective than receiving direct CF. Moreover, because that simple present had been taught to the students several times in previous years of education, it was familiar to them and as such had been partially acquired. In this case, the students may have needed only a reminder or a trigger to activate their knowledge of the target feature, which was done through underlining (i.e. indirect CF). Further, asking the students to rewrite their tasks in a second draft in each CF session may have helped them to attend to the given CF by comparing their errors, in their current language, to the target language, a concept called ‘noticing the gap’ by Schmidt (1995; 2001). The students completed three CF sessions which means they completed three second drafts, one for each task. It is possible that by rewriting their tasks based on the received CF, group B were engaged in a thinking process in which the previous knowledge was activated in order to self-correct their errors. Because of this cognitive processing, the students were able to use the target features more accurately in the new piece of writing (i.e., the immediate posttest). This result supports

what Lalande (1982) suggested as the benefit of indirect CF, which is promoting problem-solving and guided learning. In Lalande's research (1982), the indirect CF group outperformed the direct CF group leading to the suggestion that indirect CF is more effective because of the thinking processes encouraged through self-corrections.

In the delayed posttest, however, the results showed that although group B's scores showed gains from pretest to the delayed posttest, these gains were not statistically significant. This means that the performance of group B, indirect CF/simple present, dropped in the delayed posttest compared to their performance in the immediate posttest. One possible reason was the complex nature of the target feature. Present simple tense involves the use of many rules: copula Be, "have" to express possession, third person singular –s, and no –s for first and second person singular and plural, which all depend on the subject; so if the student was not able to identify the accurate subject in a complicated subject clause, such as, 'studying in a large college with many students is/are', the accurate form of the verb might be not used. All of these rules may have needed a more explicit and detailed type of CF than only marking the error to achieve long-term learning effect. If indirect CF, which is hypothesized to promote long-term learning, was accompanied by a simple explanation of the simple present tense rules in a handout distributed to the students of group B, it may have led to long-term learning. Metalinguistic information, if added to indirect CF, may lead to another level of awareness which is understanding that can lead to long-term learning (Sheen, 2007; Sheen, 2010).

On the other hand, group A, which received direct CF on simple present, did not achieve statistically significant gains in either the immediate or the delayed posttests. These results differ from the current research which mostly reports positive effect for the use of direct CF with rule-based features (e.g. Bitchener, 2008; Bitchener & Knoch, 2010a; Sheen, 2007). These different

results could be due to a number of factors. The first factor might be the complexity of the target feature, as mentioned previously. The rules of simple present are more complicated than the rules of articles examined in previous research, leading to the different results between the current study and the previous research in regard to the effect of direct CF on rule-based features. As for why group A did not improve in the use of simple present when group B achieved significant improvement in the use of the same target features, one possible factor was that group A received direct CF while group B received indirect CF. Since the students in group A were provided with direct corrections, they were not required to self-correct their errors in the second drafts. It is possible that the students in group A copied the given corrections in the second draft blindly without paying much attention to the errors and how they were corrected. This means that the students did not process the given feedback or attend to it seriously because they were not required to do anything with it in their second drafts, unlike group B who were asked to figure out the corrections by themselves and therefore experienced more thinking processes, leading them to use the target feature more accurately in the immediate posttest.

Another possible factor affecting group A's results, which did not indicate statistically significant gains in any of the posttests, may have been the impact of the students' advanced level in the use of the target features. In the pretest, group A achieved higher scores in the use of the target feature than what the other groups achieved. Seven students in group A (i.e. half the group) achieved a high score in pretest 2 (i.e. a score between 90% and 100%) compared to five, three, and four students achieving a score in the same range in groups B, C, and D respectively. Although these scores did not lead to significant differences between the four treatment groups at baseline, they still indicate that group A started at a higher level than that of the other groups, which was also obvious in the profile plots given in Figure 1 in the previous chapter.

Consequently, it is possible that because of a ceiling effect, group A was not able to achieve further gains in the posttests to exceed their already high performance in the pretest.

For the preposition groups, group C, which received direct CF, achieved statistically significant gains in the immediate posttest, and they maintained this level of improvement one month later in the delayed posttest. On the contrary, group D, which received indirect CF, did not achieve a statistically significant improvement on any of the two posttests. Because of the nature of prepositions as item-based, idiosyncratic, and difficult to master even at a high proficiency level, it is highly probable that the students were not able to correct their preposition' errors by themselves or by consulting a grammar book. Receiving direct corrections in this case was imperative as the errors could not be fixed easily or perhaps at all by the student. Through direct CF, the students in group C were able to know/understand why some of their prepositions were marked as errors and what the correct forms were.

Conversely, group D (i.e. indirect CF on prepositions) received what was probably the least effective type of CF that can be used with item-based features. This was because indirect CF provided very little information, if any, on a linguistic feature that did not function according to clear grammatical rules. If the given CF did not provide much information on the type of error, the related grammatical rule, or the correct form, and the error could not be fixed easily by consulting a grammar handbook, how would the student be able to self-correct it? Thus, the study results suggest that if the linguistic feature is clear (i.e. rule-based), indirect CF might be enough and can save teacher time, whereas if the feature is not clear in terms of its usage (i.e. item-based), it is important to provide a clear type of CF, such as direct CF, which could be the only source of the right answer for some students, and warrants teacher intervention.

The results of the study on the differential effect of direct and indirect CF differ from some of the previous research done to investigate the same question. For example, Bitchener and Knoch (2010b) suggests that direct CF is more effective than indirect CF for the acquisition of rule-based features, unlike the current study which suggests the opposite. Bitchener and Knoch (2010b) compared direct and indirect CF on the acquisition of articles, a rule-based feature, and suggested a longitudinal effect of direct CF on the use of rule-based features. The given direct CF in Bitchener and Knoch's study (2010b) was in the form of metalinguistic information, not direct corrections. The participants may have improved in the use of articles because of the metalinguistic information given in a handout that explained the usage rules and provided examples of the accurate use of articles. This type of detailed information might be the reason why the direct CF groups in Bitchener and Knoch (2010b) were able to achieve long-term learning in addition to outperforming the indirect CF group. The current study, on the other hand, provided direct CF in the form of direct corrections only, making the given direct CF less explicit than direct CF provided in the form of metalinguistic information on the target rule-based feature.

Furthermore, Chandler (2003) compared direct CF with three forms of indirect CF (underlining with description of error type, underlining only, and description only) targeting twenty-three linguistic error types. Based on the results, Chandler suggested that direct CF is more beneficial than indirect CF in producing accurate revised and subsequent writing. However, the difference between the effect of direct and indirect CF (i.e. between direct correction and underlining only) on subsequent writing was not statistically significant in Chandler (2003) suggesting that direct and indirect CF were probably equally effective. As for the difference between the effects of direct and indirect CF on revisions in Chandler's study, it cannot be used

to suggest that direct CF is more effective than indirect CF because all the students needed to do to produce more accurate revised work was to copy the given direct corrections. The current study suggests that both approaches, not only the direct approach as in Chandler (2003), can be effective depending on the linguistic feature targeted.

Van Beuningen et al. (2008) also suggested that direct CF is more helpful than indirect CF in producing more accurate new writing. The given CF in Van Beuningen et al. (2008) covered a wide variety of errors: nine error categories including verb form, word choice, spelling, capitalization, and punctuation. Therefore, Van Beuningen et al. (2008) suggested that direct CF is more effective than indirect CF for the learning of almost all types of linguistic features, whether rule-based or item-based. However, what made the indirect approach ineffective for the learning of any type of linguistic features in Van Beuningen et al. (2008) was probably providing CF on all errors in students' writing. Since indirect CF does not provide much information on the errors but instead requires the student to discover what's wrong with the underlined errors and self-correct them, providing CF on only one or two types of errors can help the student focus more on the marked errors, discover the pattern or the link between them, and then use his/her previous knowledge or refer to any source, such as a grammar book, to self-correct the errors. All of this cognitive effort was not possible in Van Beuningen et al. (2008) due to the mass correction given on all errors. Conversely, the current study provided indirect CF on one type of errors limiting the number of errors in each task and allowing the student to focus on the marked linguistic forms and think about why they were marked as errors and how to correct them. Therefore, one of the reasons for the positive effect of indirect CF in this study might have been the use of the focused approach, unlike the approach used in Van Beuningen et

al. (2008). While the focused approach was not the aim of this study, it was retained as a constant feature of the instruction for all groups and may have had an impact.

One explanation for why different types of linguistic features respond differently to different types of CF, as the current study suggests, is that different types of linguistic features may represent different domains of knowledge (Ferris, 2002). It has been suggested that syntactic knowledge and lexical knowledge are learned in different manners (Schwartz, 1993). For this reason, they should be treated/taught differently. Ferris (1999; 2002) suggested using different types of CF with different types of linguistic features, underlining for treatable errors (rule-based) and a combination of direct correction and strategy training for untreatable errors (item-based). Ferris' suggestions are in accordance with the suggestion made by the current study: using indirect CF for rule-based features and direct CF for item-based features.

However, there are studies that provide different suggestions. Van Beuningen et al., (2012) suggested that grammatical features benefit more from direct CF whereas non-grammatical linguistic features, such as spelling and capitalization, benefit more from indirect CF. This means that all grammatical features, whether rule-based or item-based, are affected more positively by direct CF. Further, Bitchener et al. (2005) suggested that direct CF does not have a positive effect on the use of prepositions. The different results between the current study and these studies may be attributed to differences in the design of the studies. Van Beuningen et al. (2012) provided unfocused CF on Dutch writing in the Netherlands, while the current study provided focused CF on English writing in an EFL context. Bitchener et al. (2005) provided CF on three linguistic features: the definite article, the past simple tense, and prepositions, in each piece of writing, unlike the present study which was highly focused and targeted only prepositions for group C. The approach used in Bitchener et al. (2005) was less focused and thus

may have distracted the students' attention from prepositions to the other target features which might seem more important to the learner. This might be why the participants in Bitchener et al. (2005) were able to achieve short- and long-term gains in the learning of the rule-based features while no improvement was achieved in the use of prepositions. The current study, on the other hand, suggested that learning of prepositions can respond to certain types of CF. Based on the results of the present study, when asked about the most effective type of written CF and the type of linguistic features which respond better to CF, the answer can be "it depends". Whatever type of CF is appropriate for one type of linguistic feature can be ineffective for another type of linguistic feature.

Students' Preferences for Different CF Types

The third question investigated the impact of individual differences, specifically the students' preferences of CF types, on the effectiveness of the given treatment on the grammatical accuracy of the target features. The accuracy scores in the three writing tests and the students' responses to the first question in Questionnaire 2, which asked whether the students preferred direct or indirect CF, were subjected to analysis. The analysis results suggested no relationship between students' preferences of CF types and the effect of the given treatment on the accurate use of the target features. In other words, the students' attitudes towards direct and indirect CF did not appear to impact the effect of CF on the grammatical accuracy of the target features. As for the students' responses to the rest of Questionnaire 2, they revealed that the students have strong preferences for certain types of CF and for certain types of linguistic errors to target when providing CF.

In their responses to Questionnaire 2, the students indicated a greater appreciation for direct CF than for indirect CF. Their responses to the first question suggested that the majority of

the participants (74%) preferred direct CF. The students' comments also reflected this preference. Some students explained in their comments that CF that marks and corrects their errors (i.e. direct CF) gives them the opportunity to understand the error and how to fix it, saves their time by providing the correct form immediately, and is the safest way to make sure the correction is accurate. Similar comments given by students on the direct approach can be found in other studies, such as Chandler (2003) in which most of the participants considered the "correction method" (i.e. direct CF) as the easiest method to correct their errors, and some commented on this method as the method consuming less time because it provides the correct forms immediately. This suggests that learners in different contexts, EFL and ESL, prefer the direct approach for its immediacy.

Responses on Questionnaire 2 also showed that the majority of students had a negative attitude towards indirect CF. Some of the comments suggested that it frustrated students not knowing the correct form immediately when receiving CF that only marks the error. This frustration was spelled out in some of the comments made on indirect CF, such as "I don't know what is wrong", "I don't know why it is wrong", and "I still don't know the right answer". However, there was a small group of participants who expressed a positive attitude towards indirect CF in their responses to the questionnaire. Some students explained in their comments that they think they learn more from indirect CF than from direct CF because the indirect approach encourages them to think for themselves to self-correct their errors. The students also suggested that the effort spent to figure out the corrections is helpful. One student stated, "when my teacher underlines my mistakes without giving correction, that gives me chance to think more about my mistakes and not repeat it again". This logic/reason for the students' preference for indirect CF has also been found in other studies, such as and Chandler (2003), Lalande

(1982), and Ferris and Robert (2001). In Chandler (2003), for example, one of the students who thought that the “underlining method” (i.e. indirect CF) is the response from which she had learned the most commented on the method by saying, “... because I can look up for the correct answer by myself, and this makes easier to remember the mistakes I made so I won’t do it again” (p. 288-289).

In addition to the findings regarding preferences for direct and indirect CF, students’ responses to the questionnaire also revealed strong preferences for other features of CF and for certain types of linguistic errors to be targeted when providing CF. First, Questionnaire 2 showed that the students have a strong preference for unfocused CF over focused CF. This preference emerged in their responses to a number of questions in Questionnaire 2 and also in some of their comments which showed a dislike for the focused approach used in the study, such as the following student who commented, “[the given CF] is useful because I learned to avoid my errors in another writing, but I don’t like it because it doesn’t provide all errors”. This result is not surprising as many studies suggested the same preference expressed by ESL and EFL learners. In Lee (2005), the majority of the participants, who were 320 EFL students in Hong Kong, said they prefer to receive comprehensive CF addressing all their errors. In Amrhein & Nassaji (2010), the majority of the participants, who were 33 ESL students in Canada, also expressed the same preference; the results also indicated a significant difference in opinions between teachers and students in terms of how much CF should be provided on writing with the students demanding larger quantities of CF than what their teachers believed they should or would provide.

Students’ preference for unfocused CF likely reflects a belief that this is what they need to improve their writing. They may think that feedback that addresses all of their errors in each

piece of writing will convert easily into language intake, which is what research has suggested to be untrue. Research has suggested that unfocused CF, whether direct or indirect, leads to a limited effect on the development of students' writing because it can overwhelm their limited learning capacities (Ellis, 2009b; Sheen et al., 2009). Another reason for the students' negative attitude towards focused CF is shown in some of their comments. Some students explained that the focused approach can be confusing because it targets some errors and leaves other errors unspecified in terms of how many there are and of what type. When targeting a few errors in students' writing, the students may be left wondering if the rest of their writing is accurate, given that there are no more errors marked, or if it includes more errors left for them to find and correct. This is an interesting and valid point raised by students and should be considered by future research to suggest a solution on how to provide CF that does not target all errors and overwhelm the students' limited learning capacity without correcting only one type of errors and leaving students frustrated and unsure of the accuracy of the rest of their texts.

As for the types of linguistic errors on which the students would like to receive direct CF, the students' responses were inclusive, covering almost all of the given options: all errors, all grammatical errors, and only the serious grammatical errors. However, "only the serious grammatical errors" received the highest percentage followed by "all errors". These responses suggested two possible conclusions. First, they indicated that a good number of the students wanted to receive direct CF on all errors, supporting again the conclusion made earlier that students prefer unfocused CF to focused CF. Second, these responses, which showed that most of the students (n=32) wanted to receive direct CF on their grammatical errors whether all grammatical errors or only the serious ones, suggested that students were mostly concerned about the grammatical accuracy of their writing. These results are consistent with those of

previous studies which reported a similar concern expressed by students for grammatical accuracy in writing (e.g. Amrhein and Nassaji, 2010; Diab, 2005). The results of Chen et al (2016), however, suggested that advanced L2 learners are less concerned about the focus-on-form errors, such as grammatical and vocabulary errors, and more concerned about writing organization and the overall quality of their writing and, thus, require feedback that addresses these aspects instead of feedback addressing form only.

On the other hand, the students expressed a different preference when asked about the type of errors on which they would like to receive indirect CF. The majority of the students expressed their preference to receive indirect CF on vocabulary errors only. Because the students showed a negative attitude towards indirect CF in their responses to several questions in the questionnaire, their preference to receive it on vocabulary errors may indicate that they thought that vocabulary errors are easy to be self-corrected or that vocabulary errors are infrequent errors in their writing, and thus receiving indirect CF on them would be acceptable. This conclusion, if valid, adds to the evidence that suggests that the students have a negative attitude towards indirect CF in that they only want to see it on less frequent errors. As for the grammatical errors, which the students are mostly concerned about, the students prefer to receive direct corrections on them, as revealed by their responses to another question.

We can notice that students' preferences and ideas about the most effective CF methods do not necessarily match what research, including the current study, supports as effective. Students preferences for certain writing practices and for specific CF types might be based on their previous learning experiences which may not necessarily be useful for writing improvement (Cohen & Cavalcanti, 1990). For example, research recommends against unfocused CF, which was highly preferred by students in this study. Research suggests that unfocused CF, whether

direct or indirect, is less effective than focused CF (e.g. Sheen et al., 2009) and leads to no or limited effect on the development of students' writing and the acquisition of the target features (e.g. Kepner, 1991; Polio et al., 1998; Sheppard, 1992; Truscott & Hsu, 2008).

Whether these preferences affect learning and writing outcomes is another question examined in research. Although the students' responses to Questionnaire 2 showed a strong preference for direct CF over indirect CF, these preferences did not appear to affect their performance in writing. The results of the present study suggested a lack of relationship between students' preferences of CF types and their accurate use of the target features in response to written CF. The available few studies that investigated this relationship produced different and inconclusive results. The early research suggested a possible link between language learning beliefs and L2 performance (Mori, 1999; Peacock, 1999). However, those early studies suggested an association only and did not provide strong evidence that learner's beliefs are the cause of the language outcomes. Peacock (1999) stated, "it is very difficult to settle the question of causality" (p. 259). That is, it is difficult to confirm if those beliefs are the cause of the learners' achievement in L2 or if the learners' achievement in L2, low or high, is the cause of the learners' attitudes towards language learning.

The results of the current study corroborate those of more recent research that suggested no relationship between beliefs and language outcomes. Tanaka and Ellis (2003), for example, found a weak, insignificant correlation between learners' beliefs and their TOEFL scores. Kartchava and Ammar (2013) found that learners' beliefs affected how much they noticed the given CF but did not affect their language outcomes. Perhaps these results are not surprising; beliefs do not always translate into actions as students will not always act based on their beliefs of how language should be learned (Ellis, 2008). Therefore, although assuming the existence of a

relationship between preferences and learning outcomes seems to be logical to a certain degree, research does not show a direct link between them.

Limitations

The current study has a number of limitations that should be considered when interpreting the findings. The first one was the lack of a control group. The writing tasks and tests were part of the class work, which means that all students participated in them. Students in this context, and probably in other English language learning contexts, expect to receive some type of feedback on their writing. As such, including a control group that does not receive any type of meaningful CF for research purposes only could be considered unethical and ecological invalid in this context. Thus, it was decided not to include a control group.

Not including a control group in any study can invalidate the results because the effect achieved by the end of the study could be the result of factors other than the given treatment. In this study, two features, however, support the validity of the results. First, the significance levels achieved were high indicating low probability of error and suggesting that the effect created was not a coincidence but the result of the given treatment. Second, the study was designed to include two pretests. By including two pretests with time interval between them, it was possible to examine how the students perform without treatment and then compare this pre-treatment performance to their performance after treatment. Including two pretests may have had an advantage as well in that the same participants were compared to themselves in two periods of time, one with treatment and one without treatment, whereas when including a control group that does not receive any type of treatment, two different groups of people with many individual differences, that could affect the results, are compared.

Another possible limitation was the researcher acting as the class instructor. I was given access to the class by the department on the condition that I teach the class for the entire semester. The department reasoning was to save the class time by assigning one teacher/researcher to manage all the work related to this class including the writing tasks. Playing this dual role can affect different dimensions of the study including the students' voluntariness and anonymity and the effect of the power relationship between the instructor and students. To lessen these effects, a research assistant was assigned to manage the tasks that were related to the research project, but not related to the class work, such as inviting the students to participate in the study, signing the consent letters, and completing the questionnaires. However, there may have been a positive side to this dual role as it gave me a better understanding of the context and the participants.

The sample size can also be considered a limitation. CF studies often include a sample similar in size to the one in the current study because providing written CF can be time- and energy- consuming. However, the participants in the current study were divided into four groups making the size of each group small. Having a small sample size can affect the statistical tests' ability to detect small effects of the variables and their interaction if one truly exists. Increasing the sample size in the current study, however, was not possible because the target context was a small college with only one writing class.

Another limitation is the length of the study. In the present study, data collection lasted for twelve weeks from the first pretest to the delayed posttest. Between the immediate posttest and the delayed posttest, there were four weeks without treatment. Since the delayed posttest is used as an indication of how much learning is retained from the given treatment, having measured the effect over a longer period of time would strengthen the study's suggestion of a

long-term effect of CF. However, if the treatment was to be extended, many factors could interfere with the treatment and affect the results, such as the possibility that students would have extra writing practice. Although the study suggested no effect for writing practice compared to CF on the grammatical accuracy of students' writing, these results might have been different if the study had lasted for a longer period of time that included submitting more written assignments for other nursing classes, completing more writing exercises, and practicing the language in other classes because English is the medium of instruction in all nursing classes.

One aspect of how the study conditions may have affected the results is the fact that all the participants (i.e. the four treatment groups) were placed in the same room at once when completing all the tasks. It was planned that the students should not talk to each other in the CF sessions to avoid helping each other with the given CF. Still, it remains possible that some students may have consulted each other on why some words were marked as errors and how to correct them. However, if this happened, it should not have affected the results significantly because the students were clearly instructed not to talk to each other and were also asked to leave one empty chair between every two students, and all of this was done under the supervision of the class instructor/researcher.

The language used in formulating the statements in Questionnaire 2 may have affected the students' responses to some questions. The first and second questions in Questionnaire 2 were compounded in that they addressed two aspects at the same time in the given options. For example, the second question asked about the student's attitude towards direct CF, and the given options suggested possible reasons for each response. As revealed by the students' responses to the first and second questions, mentioned in detail in the previous chapter, it seems that the language used affected their responses. Whereas 74% of the students chose the direct CF method

as a response to the first question asking about the most effective CF method in their opinion, only 62% responded with “I like [direct CF] because it provides the correct form immediately” as a response to the second question which asked about their reaction to direct CF. Based on this limitation, a more developed questionnaire is needed to break down these compound questions into smaller ones and, consequently, produce more accurate and precise answers to the posed questions. The current results of Questionnaire 2 can be used as a starting point on which further research is established to investigate the question of the impact of CF preferences on how the students respond to written CF.

Conclusions and Implications

The goal of the current study was to contribute to the body of research on the effect of CF on writing and L2 development by investigating the effect of direct and indirect CF on the learning of different types of linguistic features, rule-based (treatable) features and item-based (less treatable) features. The results of the study suggested that different types of linguistic features should be treated differently when providing CF. Treatable errors, those that work according to clear grammatical rules and thus can be self-corrected, may benefit more from indirect CF. Less treatable or untreatable errors, those that are idiomatic in nature and do not function according to clear rules, may benefit more from direct CF.

Direct CF provides students with explicit guidance on how to correct the target feature (Ellis, 2009b). This explicit guidance is much needed on item-based features because most probably these features cannot be self-corrected by the student, and, thus, they need to be corrected by an expert in the language like the teacher. Therefore, teachers are encouraged to provide their students with direct corrections on the less treatable errors to help them know what the error is and how to fix it. On the other hand, rule-based features seem to be better treated

with indirect CF. Since they are treatable, they can be self-corrected by the student making indirect CF sufficient. Also, indirect CF promotes more reflection and guided learning which should foster long-term learning (Ferris, 2002). However, since the present study did not result in statistically significant improvement in the delayed posttest for the indirect CF/treatable errors' group, it might be useful to combine indirect CF with metalinguistic information on the target features provided in a handout. The combination of metalinguistic information and indirect CF may lead to more improvement and possibly long-term learning because it does not only mark the error, but it also explains and reminds the students of the rules that control the target features, and in turn, it can help the learner with self-correction. Further, indirect CF and a handout that is printed for the whole class consume less time and energy than direct CF and/or metalinguistic CF provided in the margins of each text, thus, this method is suitable to be used in large classes and for teachers with a heavy workload.

On the last research question which investigated students' preferences of CF types and if they are related to writing improvement, the statistical analysis suggested no relationship between students' preferences of CF types and the effectiveness of the given CF on accuracy in writing. As for students' preferences of CF practices, the students showed more interest in unfocused, direct CF targeting grammatical errors than focused, indirect CF targeting other types of errors. The teachers' responsibility is to be aware of these preferences and to try to meet the students' expectations to bridge the gap between their students' demands and their own practices in class (Diab, 2005). If, however, meeting these preferences is not possible for different reasons, it is recommended that the teacher explain to students why feedback is provided to them in this particular way and how it is intended to help the development of their writing (Diab, 2005). If the students understand the benefit of, for example, focused CF or indirect CF, they might be

more capable of accepting these types of feedback on their writing and may change their negative attitude towards them when they experience their effectiveness on their writing over time. Teachers will always need to make an informed decision on how to handle their students' errors, depending on students' needs and level in writing and sometimes regardless of students' preferences for certain types of CF that might be unrealistic or unsupported by research.

An emergent aspect observed in the study, although was not one of the goals of the study, is the used focused approach. The positive results of the study in addition to the positive results of previous research in which CF targeted one (or a few) type of error may suggest that the used focused approach was useful. Although the study did not compare focused with unfocused CF, it was interesting to see that almost all the studies that employed the focused approach reported positive effect of CF (e.g. Bitchener, 2008; Bitchener & Knoch, 2010a; Ellis et al., 2008) while almost all the studies that employed the unfocused approach reported either limited or negative effect of CF (e.g. Polio et al., 1998; Sheppard, 1992; Truscott & Hsu, 2008). Unfocused CF targeting different types of linguistic features at a time can overload a student's limited memory and exhaust teachers (Ferris, 2002). Conversely, providing many corrections for one single feature (i.e. focused CF) gives the student rich evidence of how to use this feature in subsequent writing (Ellis, 2009b). Also, some researchers argue that the focused approach is more effective in achieving long-term learning; Ferris (2010) stated, "students would utilize written CF more effectively for long-term language acquisition and writing development when there are fewer, clearer error types on which to focus attention" (p. 192). Further, there are some studies that compared the effect of focused CF to unfocused CF and suggested that focused CF is more effective than unfocused CF (e.g. Sheen et al., 2009).

Based on these results and suggestions, teachers may want to avoid addressing all types of errors in students' writing and select one to three types of linguistic features to correct in every writing task. Teachers can be selective in two ways. First, they can choose to focus on form or content at different times (Sheen, 2007). Second, if they choose to focus on form, they can select a limited number of grammatical features to focus on instead of targeting all the grammatical features in every assignment (Sheen, 2007). Choosing the target feature(s) for each assignment can be done based on what the teacher sees as the most problematic features in students' writing that need to be brought to their attention, or it can be done through negotiation with students.

Directions for Future Research

The findings of the current study add to the body of research on written CF by providing suggestions on how different types of CF could have different uses and purposes. These findings also raise many questions in need of further investigation. The first one is, what variables have an impact on the effectiveness of the given treatment? In the present study, only one possible factor was examined, students' preferences. In future research, there is a need to investigate other factors to discover how some students benefit from CF while other students in other situations fail to do so. These factors could include individual differences, such as motivation, aptitude, learning strategies, and attitude towards language learning in general and CF in specific. This can be done through qualitative studies, such as case studies in which the writing of a small number of learners is examined and analyzed closely to discover how some learners gain or fail to gain any improvement from the given CF.

Moreover, there is a clear need for more studies investigating the treatable/untreatable contrast with different linguistic features. Most of the previous research examined the effect of

CF on fairly easy features which were also partially known to the participants. Future research needs to examine other linguistic features which are either new to the students or more complex than the ones examined so far. Without this type of research, some of Truscott's arguments against CF (1996; 2008) might be true, in that CF may be ineffective for the acquisition of complex grammatical features (Ellis et al., 2008).

Another question in need of further investigation is the focused/unfocused contrast. "A mass correction directed at a diverse set of linguistic phenomena... is hardly likely to foster the noticing that might be needed for CF to work for acquisition" (Ellis et al., 2008, p. 368). Nonetheless, the unfocused approach is the most authentic CF approach. It is the approach that is commonly used by L2 instructors in their writing classes as most of the teachers would like to improve the overall accuracy of their students' writing, not just several linguistic features by the end of the course. Also, as suggested by research, unfocused CF is the type of CF preferred the most by students. However, since research suggests a limited value in unfocused CF, another way to meet teachers' and students' desire of improving as much as possible of students' writing is to examine the effect of a semi-focused approach in which several features are targeted and see if such an approach can lead to accuracy improvement. If yes, then a semi-focused approach will be more practical for certain writing classes than the highly focused approach recommended and examined extensively by recent research.

Furthermore, most of the early research was conducted in writing classrooms where the teacher provided unfocused CF on students' class work. Although those early studies had many limitations and suggested negative or limited effect of CF, they still represent the authentic writing class. Much of the more recent research was not carried out in writing classes, but in other L2 classes, such as a speaking class, and involved asking the participants to produce

writing that does not resemble authentic writing. The current study, on the other hand, used an authentic classroom where the teacher provided her students with feedback on real writing tasks that were part of the class work. Since this is only one study, there is a definite need for research conducted in more authentic circumstances, because in the end the findings of this type of research should benefit writing classrooms.

The topic of the effect of CF on writing development and second language acquisition is complicated as it involves many variables that could affect its results. Teachers spend considerable time providing their students with written CF. Providing effective CF can ensure that this time is not wasted. Although recent research has already established that written CF has a positive effect on L2 writing and learning, many practices related to how to use CF to its maximum benefit are still not clear or have not yet been fully explored. The current study was an attempt to address one of the important issues related to CF practices- how to address different types of linguistic features through CF. Based on the results, the study supported recommendations for specific practices to help teachers provide more effective CF on their students' writing.

References

- Agudo, J. (2012). Investigating Spanish EFL students' beliefs and preferences regarding the effectiveness of corrective feedback. *International Journal of Humanities and Social Science*, 2 (19), 121-131.
- Alamargot D., & Chanquoy L. (2001). *Through the models of writing*. London: Kluwer Academic Publishes.
- Amrhein, H., & Nassaji, H. (2010). Written corrective feedback: What do students and teachers prefer and why? *Canadian Journal of Applied Linguistics*, 13, 95-127.
- Ashwell, T. (2000). Patterns of teacher response to student writing in a multi-draft composition classroom: Is content feedback followed by form feedback the best method. *Journal of Second Language Writing*, 9, 227-257.
- Azar, B. (1986). *Understanding and using English grammar*, Middle East Edition. London: Prentice-Hall International.
- Bitchener, J., Young, S. & Cameron, D. (2005). The effect of different types of corrective feedback on ESL student writing. *Journal of Second Language Writing*, 14, 191-205.
- Bitchener, J. (2008). Evidence in support of written corrective feedback. *Journal of Second Language Writing*, 17, 102-118.
- Bitchener, J., & Knoch, U. (2008). The value of written corrective feedback for migrant and international students. *Language Teaching Research Journal*, 12(3), 409-431.
- Bitchener, J., & Knoch, U. (2010a). The contribution of written corrective feedback to language development: A ten-month investigation. *Applied Linguistics*, 31 (2), 193-214.
- Bitchener, J., & Knoch, U. (2010b). Raising the linguistic accuracy level of advanced L2 writers with written corrective feedback. *Journal of Second Language Writing* 19, 207-2017.

- Bitchener, J. (2012). Written corrective feedback for L2 development: current knowledge and future research. *TESOL Quarterly*, 46(4), 855-860.
- Butterfield, E.C., Hacker, D.J., & Albertson, L.R. (1996). Environmental, cognitive and metacognitive influences on text revision: assessing the evidence. *Educational Psychology Review*, 8(3), 239-297.
- Celce-Murica, M., & Larsen-Freeman, D. (1999). *The grammar book: An ESL/EFL teacher's course*, 2nd Edition. New York: Heinle.
- Chandler, J. (2000, March). The efficacy of error correction for improvement in the accuracy of L2 student writing. *Paper presented at the AAAL conference*, Vancouver, BC.
- 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.
- Chen, S., Nassaji, H., & Liu, Q. (2016). EFL learners' perceptions and preferences of written corrective feedback: a case study of university students from Mainland China. *Asian-Pacific Journal of Second and Foreign Language Education*, 1, 1-17.
- Cohen, A. D., & Cavalcanti, M. C. (1990). Feedback on written compositions: Teacher and student verbal reports. In B. Kroll (Ed.), *Second language writing: Research insights for the classroom* (pp. 155-177). Cambridge, UK: Cambridge University Press.
- DeKeyser, R. M. (1994). Implicit and explicit learning of second language grammar: A pilot study. *TESOL Quarterly*, 28(1), 188-194.
- DeKeyser, R. M. (1998). Beyond focus on form: Cognitive perspective on learning and practicing second language grammar. In C. Doughty & J. Williams (Eds.), *Focus on form*

- in classroom second language acquisition* (pp. 42-63). Cambridge: Cambridge University Press.
- DeKeyser, R. M. (2003). Implicit and explicit learning. In C. J. Doughty & M. H. Long (Eds.), *Handbook of second language acquisition* (pp. 313-348). Oxford: Blackwell.
- Diab, R. (2005). Teachers' and students' beliefs about responding to ESL writing: A case study. *TESL Canada Journal*, 23(1), 28-43.
- Ellis, R. (1997). *SLA Research and Language Teaching*. Oxford: Oxford University Press.
- Ellis, R. (2005). Principles of instructed language learning. *System*, 33(2), 209-224.
- Ellis, R. (2008). Learner beliefs and language learning. *Asian EFL Journal*, 10(4), 7-25
- Ellis, R. (2009a). Implicit and explicit learning, knowledge and instruction. In R. Ellis, S. Loewen, C. Elder, R. Erlam, J. Philip, & H. Reinders. *Implicit and explicit knowledge in second language learning, testing and teaching* (pp. 3-17). Bristol, UK: Multilingual Matters.
- Ellis, R. (2009b). A typology of written corrective feedback types. *ELT Journal*, 36 (2), 97-107.
- Ellis, R., & Barkhuizen, G. (2005). *Analyzing learner language*. Oxford: Oxford University Press.
- Ellis, R., Loewen, S., & Erlam, R. (2006). Implicit and explicit corrective feedback and the acquisition of L2 grammar. *Studies in Second Language Acquisition*, 28 (2), 339-368.
- 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 (3), 353-371.

- Fathman, A. K. & Whalley, E. (1990). Teacher response to student writing: focus on form versus content. In B. Kroll (Ed.), *Second language writing: Research insights for the classroom* (pp. 178-190). Cambridge: Cambridge University Press.
- Ferris, D. (1999). The case for grammar correction in L2 writing classes: A response for Truscott (1996). *Journal of Second Language Writing*, 8, 1-10.
- Ferris, D. (2002). *Treatment of error in second language student writing*. Ann Arbor: University of Michigan Press.
- Ferris, D. (2006). Does error feedback help student writers? New evidence on the short- and long-term effects of written error correction. In K. Hyland & F. Hyland (Eds.), *Feedback in second language writing: Contexts and issues* (pp. 81-104). New York: Cambridge University Press.
- Ferris, D. (2010). Second language writing research and written corrective feedback in SLA: Intersections and practical applications. *Studies in Second Language Acquisition*, 32 (2), 181-201.
- Ferris, D., Liu, H., Sinha, A., & Senna, M. (2013). Written corrective feedback for individual L2 writers. *Journal of Second Language Writing*, 22 (3), 307-329.
- Ferris, D., & Roberts, B. (2001). Error feedback in L2 writing classes: How explicit does it need to be? *Journal of Second Language Writing*, 10, 161-184.
- Flower, L., & Hayes, J. (1981). A cognitive process theory of writing. *College Composition and Communication*, 32, 365-387.
- Gass, S. (2003). Input and Interaction. In Doughty, C. S., & Long, M. H. *Handbook of Second Language Acquisition*. Oxford: Blackwell. 224-253.

- Goldstein, L. (2004). Questions and answers about teacher written commentary and student revision: Teachers and students working together. *Journal of Second Language Writing* 13, 63-80.
- Hacker, D. & Sommers, N. (2011). *Rules for writers*, 7th Edition. Boston: Bedford & St. Martin's.
- Harley, B. & Swain, M. (1978). An analysis of the verb system by young learners of French. *Interlanguage Studies Bulletin*, (3), 35-79.
- Hayes, J. & Flower, L. (1980). Identifying the organization of writing processes. In L. Gregg, & E. Steinberg (Eds.). *Cognitive processes in writing: An interdisciplinary approach* (pp. 3-30). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Hayes, J., Flower, L., Schriver, K., Stratman, J., & Carey, L. (1987). Cognitive processes in revision. In S. Rosenberg (Ed.). *Advances in applied psycholinguistics*, Vol. 2: Reading, writing and language learning (pp. 176-240). Cambridge: Cambridge University Press.
- Hogue, A. (2008). *First steps in academic writing*, 2nd Edition. US: Pearson Education.
- Hyland, F., & Hyland, K. (2001). Sugaring the pill: Praise and criticism in written feedback. *Journal of Second Language Writing*, 10, 185-212.
- Kartchava, E. & Ammar, A. (2014). Learners' beliefs as mediators of what is noticed and learned in the language classroom. *TESOL Quarterly*, 48(1), 86-109.
- Kepner, C. (1991). An experiment in the relationship of types of written feedback to the development of second-language writing skills. *The Modern Language Journal*, 75(3), 305-313.
- Krashen, S.D. (1982). *Principles and practice in second language acquisition*. Oxford: Pergamon

- Lalande, J. (1982). Reducing composition errors: an experiment. *Modern Language Journal*, 66, 140-149.
- Lee, I. (2005). Error correction in the L2 writing classroom: What do students think? *TESL Canada Journal*, 22(2), 1-16.
- Lizotte, R. (2001, Winter). Quantifying progress in an ESL writing class. *MATSOL Currents*, 27(1), 7-17.
- Long, M. (1996). The role of the linguistic environment in second language acquisition. In *Handbook of second language acquisition*, eds. W. R. Ritchie & T. J. Bhatia, 413-468. San Diego, CA: Academic Press.
- Long, M., & Robinson, P. (1998). Focus on form: Theory, research and practice. In *Focus on form in classroom second language acquisition*, eds. C. Doughty & J. Williams, 15-41. Cambridge: Cambridge University Press.
- Mori, Y. (1999). Epistemological beliefs and language learning beliefs: What do language learners believe about their learning. *Language Learning*, 49, 377-415.
- Park, H. (2010). Teachers' and learners' preferences for error correction (Unpublished master's thesis). California State University, Sacramento, CA.
- Pavlik, C. & Segal, M. K. (2007). *Interactions: Paragraph development & introduction to the essay*, Middle East Edition. UK: McGraw-Hill Education.
- Peacock, M. (1999). Beliefs about language learning and their relationship to proficiency. *International Journal of Applied Linguistics*, 9(2), 247-265.
- 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.

- Reichelt, M. (2001). A critical review of foreign language writing research on pedagogical approaches. *The Modern Language Journal*, 85, 578-598.
- Robb, T., Ross, S., & Shortreed, I. (1986). Salience of feedback on error and its effect on EFL writing quality. *TESOL Quarterly*, 20(1), 83-93.
- Scardamalia, M., & Bereiter, C. (1983). The development of evaluative, diagnostic and remedial capabilities in children's composing. In M. Martlew (Ed.), *The psychology of written language: A developmental approach*. New York: Wiley.
- Schmidt, R. (1990). The role of consciousness in second language learning. *Applied Linguistics* 11(2), 129-158.
- Schmidt, R. (1994). Deconstructing consciousness in search of useful definitions for applied linguistics. *AILA Review* 11, 11-26.
- Schmidt, R. (1995). *Attention and awareness in foreign language learning*. Honolulu: University of Hawaii Press.
- Schmidt, R. (2001). Attention. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 3-32). New York: Cambridge University Press.
- Schmidt, R. (2010). Attention, awareness, and individual differences in language learning. In W. M. Chan, S. Chi, K. N. Cin, J. Istanto, M. Nagami, J. W. Sew, T. Suthiwan, & I. Walker (Eds.), *Proceedings of CLaSIC* (pp. 721-737). Singapore: National University of Singapore, Centre for Language Studies.
- Schwartz, B. D. (1993). On explicit and negative evidence effecting and affecting competence and "linguistic behavior." *Studies in Second Language Acquisition*, 15(2), 147 – 163.
- Semke, H. (1984). The effect of the red pen. *Foreign Language Annals*, (17), 195-202.

- Sharwood Smith, M. (1991). Speaking to many minds: On the relevance of different types of language information for the L2 learner. *Applied Linguistics*, (7), 239-256.
- Sheen, Y. (2007). The effect of focused written corrective feedback and language aptitude on ESL learners' acquisition of articles. *TESOL Quarterly*, 41 (2), 255-283.
- Sheen, Y. (2010). Differential effects of oral and written corrective feedback in the ESL classroom. *Studies in Second Language Acquisition* (32), 203-234.
- Sheen, Y. (2011). Corrective feedback, individual differences and second language learning. Springer Dordrecht Heidelberg London New York.
- Sheen, Y., Wright, D., & Moldawa, A. (2009). Differential effects of focused and unfocused written correction on the accurate use of grammatical forms by adult ESL learners. *System*, 37 (4), 556-569.
- Sheppard, K. (1992). Two feedback types: do they make a difference? *RELC Journal*, 23, 103-110.
- Shintani, N., & Ellis, R. (2013). The comparative effect of direct written corrective feedback and metalinguistic explanation on learners' explicit and implicit knowledge of the English indefinite article. *Journal of Second Language Writing*, 22, 286-306.
- Storch, N. (2010). Critical Feedback on Written Corrective Feedback Research. *IJES, International Journal of English Studies*, 10(2), 29-46.
- Swain, M. (1985). Communicative competence: Some roles of comprehensible input and comprehensible output in its development. In S. M. Gass & C. G. Madden (Eds.), *Input in second language acquisition* (pp. 235-253). Rowley, MA: Newbury House.

- Swain, M. (1991). French immersion and its offshoots: Getting two for one. In B. Freed (Ed.), *Foreign language acquisition: Research and the classroom* (pp. 91-103). Lexington, MA: Heath.
- Tanaka, K. & Ellis, R. (2003). Study abroad, language proficiency, and learner beliefs about language learning. *JALT Journal*, 25, 63–85.
- Truscott, J. (1996). The case against grammar correction in L2 writing classes. *Language Learning*, 46, 327-369.
- Truscott, J. (1998). Noticing in second language acquisition: A critical review. *SLA Research*, 14(2), 103-135.
- Truscott, J., & Hsu, A. Y.-p. (2008). Error correction, revision, and learning. *Journal of Second Language Writing*, 17(4), 292-305.
- Van Beuningen, C. (2010). Corrective feedback in L2 writing: theoretical perspectives, empirical insights, and future directions. *International Journal of English Studies*, 10(2), 1-27.
- Van Beuningen, C., De Jong, N., & Kuiken, F. (2008). The effect of direct and indirect corrective feedback on L2 learners' written accuracy. *International Journal of Applied Linguistics*, 156, 279-296.
- Van Beuningen, C., De Jong, N., & Kuiken, F. (2012). Evidence on the effectiveness of comprehensive error correction in second language writing. *Language Learning*, 62 (1), 1-41.
- Yoshida, R. (2008). Teachers' choice and learners' preference of corrective feedback types. *Language Awareness*, 17(1), 78–93.

Appendix A

Informed Consent Form

Study name

The Effect of Direct and Indirect CF on the Acquisition of Rule-based and Item-based Linguistic Features.

Researcher

Researcher name: Fatimah Alkhawajah

Doctoral Candidate

Graduate Program in Linguistics and Applied Linguistics

Email address: fia121@yorku.ca

Purpose of the research

This research study examines the effect of written corrective feedback on the students' grammatical accuracy in writing. The purpose of the study is to help instructors choose the most effective way of giving corrective feedback to students developing their writing skills.

The results of the study will be reported in a doctoral dissertation and shared with teachers and researchers through reports and presentations.

What you will be asked to do in the research

If you agree to participate in this research, a research assistant will ask to you to complete a questionnaire about your English learning background. During the course, you will participate in the same instruction as the other students and there will be no difference between you and the other students. This includes completing writing tasks and responding to corrective feedback from your instructor. At the end of the study, the research assistant will ask you to complete another questionnaire asking about your experience in the study.

The estimated time commitment for participation is twelve weeks.

Risks and discomforts

The risks are minimal. Since the writing tasks are adapted from writing textbooks, they should not cause more anxiety than what you usually go through in your writing classes. Also, your instructor will not know if you agreed to participate or not until after the course ends and the grades have been submitted. The research assistant will keep all the consent forms and questionnaires until after the course ends, and you can contact the assistant directly if you decide to withdraw from the study while you are in the course.

Benefits of the research and benefits to you

This research study may inform both writing pedagogy and second language writing research.

You, the participant, will benefit from the given corrective feedback on your writing and the writing practice offered in the writing tasks.

Voluntary participation

Your participation in the study is completely voluntary and you may choose to stop participating at any time. Your decision not to volunteer will not influence your relationship with the class instructor/researcher, study staff, or the nature of your relationship with York University either now, or in the future. If you have questions about the study, you can contact the research assistant and this will not be shared with your instructor during the course.

Withdrawal from the study

You can stop participating in the study at any time, for any reason, if you so decide. You can do so by contacting the research assistant directly and without the knowledge of the instructor. Your decision to stop participating, in part or entirely, or to refuse to answer particular questions, will not affect your relationship with the teacher/researcher, York University, or any other group associated with this project. In the event you withdraw from the study, all associated data collected will be immediately destroyed wherever possible.

Confidentiality

All data (e.g. questionnaires and writing tasks) will not be associated with identifying information. All written data will be stored and archived in a locked cabinet accessible only to the researcher. All electronic data will be stored on a password protected computer accessible only to the researcher. They will be stored for two years after the completion of the study. Afterwards, they will be destroyed. Confidentiality will be provided to the fullest extent possible by law.

Questions about the research?

This research has been reviewed and approved by the Human Participants Review Sub-Committee, York University's Ethics Review Board and conforms to the standards of the Canadian Tri-Council Research Ethics guidelines. If you have any questions about this process, or about your rights as a participant in the study, you may contact the Senior Manager and Policy Advisor for the Office of Research Ethics, 5th Floor, York Research Tower, York University, telephone 416-736-5914 or e-mail ore@yorku.ca

Legal rights and signatures:

I, _____,
consent to participate in a research study titled "The Effect of Direct and Indirect CF on the Acquisition of Rule-based and Item-based Linguistic Features"
conducted by Fatimah Alkhawajah. I have understood the nature of this project and wish to participate. I am not waiving any of my legal rights by signing this form. My signature below indicates my consent.

Signature _____
Participant

Date _____

Signature _____
Principal Investigator

Date _____

Appendix B

Student Questionnaire 1

Name: ----- Group: -----

The following questions will help the researcher understand your English learning background.

Fill in the blanks and choose the most suitable answer for your situation. If a question does not apply to your situation, write “N/A” or leave it blank.

1. First language: -----
2. Other language(s) known: -----
3. I have spent ----- years studying English.
4. Apart from school, I learned English in -----

5. I have visited a foreign country(-ies) where I had to speak English
 - a. Yes, for ----- (length)
 - b. No
6. I have visited an English speaking country(-ies)
 - a. Yes, for ----- (length) in order to -----(purpose)
 - b. No
7. I have enrolled in an English program abroad.
 - a. Yes, for ----- (length) in ----- (country)
 - b. No

Appendix C

Student Questionnaire 2

Name: ----- Group: A B C D

Choose the most suitable answer for your situation and fill in the blanks.

1. In your opinion, what is the best way of providing corrective feedback on your writing?
 - a. Underline and correct all errors
 - b. Underline all errors without correcting them
 - c. Underline and correct only the most serious (difficult) errors
 - d. Underline only the most serious (difficult) errors without correcting them.
2. What is your reaction to the teacher's feedback when she underlines your error(s) and also provides the correct form?
 - a. You like it because it provides you with the correct form immediately.
 - b. You don't like it because you want to figure out the correct form by yourself.
3. What is your reaction to the teacher's feedback when she underlines your error(s) but does not provide the correct form?
 - a. You like it and you try to figure out the correct form by yourself.
 - b. You don't like it but you still try to figure out the correct form.
 - c. You don't like it and you don't try to figure out the correct form.
4. What are the error types that you want your teacher to underline and correct?
 - a. All errors.
 - b. All grammatical errors.
 - c. Only the grammatical errors that are difficult for students at my English level.
 - d. Vocabulary errors.

5. What are the error types that you want your teacher to underline only without providing the correct form?
- a. All errors.
 - b. All grammatical errors.
 - c. Only the grammatical errors that are difficult for students at my English level.
 - d. Vocabulary errors.
6. Do you feel that your writing has improved as a result of the given corrective feedback?
- a. Yes
 - b. No

Reason for your answer:

Appendix D

Survey Completed by Writing Instructors Prior to Study

Name: -----

Email address: -----

I am interested in carrying out research that examines the effect of written corrective feedback on the acquisition of difficult versus easy linguistic features in English and would like to know more about the problematic linguistic features that students face in English writing.

By linguistic features I mean "aspects of language that your students need to learn and that you make a point of addressing in your classes, either by direct instruction or through feedback".

Thus, my question is:

As a writing instructor, what are the challenging linguistic features that your students face when writing? Please rank them starting with the most difficult ones.

1. -----
2. -----
3. -----
4. -----
5. -----
6. -----
7. -----
8. -----
9. -----
10. -----

How many linguistic features do you correct in every assignment? Why?

- a. I try to correct all linguistic errors
- b. In every assignment, I correct a limited number of linguistic features (How many? -----)

Your reason -----

How do usually provide corrective feedback to your students? Why? (You can choose more than one answer)

- a. Underline the errors without correcting them
- b. Underline the errors and provide the correct form
- c. Provide written metalinguistic explanation for some of the rules in the margins of students' writing.
- d. Provide written metalinguistic explanation for some of the rules in a handout distributed to all students.
- e. Provide oral metalinguistic explanation on the most common errors for the whole class.
- f. Provide oral conferences private to every student.
- g. Other. Explain -----

Your reason for the chosen answer(s)

Tell me a little about your background in teaching English (e.g. your degree of education, years of expertise in teaching English, years of expertise in teaching writing)

Appendix E

Summary of Writing Tasks and Tests

Pretest 1

Small colleges, like your college, have certain advantages and disadvantages. For example, small colleges offer small-sized classes in which students get more attention from their professors. Write about the advantages and disadvantages of small colleges. Support every main point you present with details and examples.

Helping vocabulary/ideas: size of classes and different facilities, availability/ unavailability of courses, majors, relationship with professors, attention, making friends, peer connection, un/friendly, experience, evaluation, types of exams.

Pretest 2/ Task1

Large universities have certain advantages and disadvantages. For example, large universities have a wide variety of courses and majors to choose from. Write about the advantages and disadvantages of large universities/colleges. Support every main point you present with details and examples.

Helping vocabulary/ideas: courses, majors and academic departments, size of classes, relationship with faculty, making friends, experience, social life, extracurricular activities, evaluation, types of exams.

Task 2

Living with your parents, in comparison to living in a dormitory, has certain advantages and disadvantages. For example, living with parents helps the student with expenses. Write about the advantages and disadvantages of living with your parents. Support every main point you present with details and examples.

Helping vocabulary/ideas: clean, cook, laundry, rest, depend, dependent, house chores, rules, share, saving money, freedom, independence.

Task 3

Living in a student dormitory has certain advantages and disadvantages. For example, living in a dorm gives the student more freedom and privacy. Write about the advantages and disadvantages of living in a dorm. Support every main point you present with details and examples.

Helping vocabulary/ideas: clean, cook, laundry, rent, independent, independency, experience, freedom, chores, dorm rules, rest, share, paying bills, away from family.

Immediate Posttest

Studying abroad has certain advantages and disadvantages. For example, studying abroad helps the student learn another language faster than studying it at school. Write about the advantages and disadvantages of studying in a foreign country. Support every main point you present with details and examples.

Helping vocabulary/ideas: learn, expand, language, history, culture, knowledge, appreciate, tolerant, tolerance, flexible, ways of living, habits, traditions, experience, explore, family, homesick, different life style, making friends, expenses, immigration/visa rules.

Delayed Posttest

Studying in your country, in comparison to studying abroad, has certain advantages and disadvantages. For example, making friends in a familiar environment is easy. Write about the advantages and disadvantages of studying in your country. Support every main point you present with details and examples.

Helping vocabulary/ideas: family& friends, expenses, culture and traditions, same life style,
experience, explore, language, appreciate.

Appendix F

Pilot Study: Evaluation Sheet for the Writing Tasks and Tests [Filled out by Instructors]

Name: -----

Degree of education: -----

Years of expertise in teaching English: -----

Years of expertise in teaching English writing: -----

I. Please rank the following writing tasks starting with the most difficult ones for the target classes, i.e. Composition 1 classes:

1. Write about the advantages and disadvantages of the facilities of the Arts College.
2. Write about the advantages and disadvantages of living with your parents in comparison to living in a dormitory.
3. Write about the characteristics of successful teachers.
4. Write about the characteristics of successful nurses.
5. Write about the characteristics of your perfect destination for tourism.
6. Write about the advantages and disadvantages of living in a foreign country.

II. What do you think of the tasks above in terms of their suitability to the target classes?

If you think that some of them are not suitable for the students' level in English,
please specify them and suggest other topics, whether from the textbook or external
sources, that are more suitable for the target classes?

Appendix G

Pilot Study: Evaluation Sheet for the Questionnaires [Filled out by Students]

Name: ----- Group: -----

Please choose the most suitable answer for your situation.

1. Were the questions clear?

- Yes
- No. Which questions do you find ambiguous (not clear) and why?

2. Was the vocabulary used in the questions easy to understand?

- Yes
- No. Which words do you find difficult?

Appendix H

Pilot Study: Evaluation Sheet for the Writing Tasks and Tests [Filled out by Students]

Name: ----- Group: -----

Please choose the most suitable answer for your situation.

1. Were you familiar with the topic?

- Very familiar
- Familiar
- Not familiar

2. Did you find the topic interesting?

- Yes.
- No. Why not?

3. Was the vocabulary used in the question easy to understand?

- Yes.
- No. Which words do you find difficult?

4. Was the given time enough to complete the task?

- Yes.
- No. Why not? How much time do you think you need?

Appendix I

Sample of Students' Writing (Scoring for Present Simple)

Group: A F. Alkhwajah Immediate Posttest

Studying abroad has certain advantages and disadvantages. For example, studying abroad helps the student learn another language faster than studying it at school. Write about the advantages and disadvantages of studying in a foreign country. Support every main point you present with details and examples.

Helping vocabulary/ideas: learn, expand, language, history, culture, knowledge, appreciate, tolerant, tolerance, flexible, ways of living, habits, traditions, experience, explore, family, homesick, different life style, making friends, expenses, immigration/visa rules.

There are a lot of students like to study in a foreign country. Studying abroad has certain advantages and disadvantages. For example, when the student goes to a foreign country, he/she can learn a lot of things like the country's cultures and other things. In my opinion, this thing is very important for every person who lives in this life. For example, when he/she discusses with any person about the country's culture or any thing, he/she can talk about a lot of information and correct information. Another thing is that a lot of people like to try to live with different life style. For example, some people like to stay with different families in the same home to see how they cook, eat, and deal with other people. These are some advantages of studying abroad.

In my opinion, studying abroad has very few disadvantages. For example, some people can't live away from their families. So, they can't focus on their goal of traveling to an other countries which is studying. Also, some of them go to an other countries to explore these countries not to the studying. Another thing is that some people can't deal with the people of different nationalities. So, they face a lot of problems when they need to talk with the others. In addition to immigration problems, visa rules, and etc. So, these are some disadvantages of studying abroad.

⇒

Finally, studying abroad²² has some advantages and disadvantages. It is dependent²³ on the student him/herself. He/she can only decide if studying abroad is²⁴ good or bad for him/herself. For me, studying abroad²⁵ is very effective way to study.

$$\frac{23.5}{25} \times 100 = 94\%$$

Appendix J

Sample of Students' Writing (Scoring for Prepositions)

Group: A

F. Alkhwajah Pretest 2

Large universities have certain advantages and disadvantages. For example, large universities have a wide variety of courses and majors to choose from. Write about the advantages and disadvantages of large universities/colleges. Support every main point you present with details and examples.

Helping vocabulary/ideas: courses, majors and academic departments, size of classes, relationship with faculty, making friends, experience, social life, extracurricular activities, evaluation, types of exams.

1 After graduating 2 from high school, every student enters a college or a university. Each one looks 3 for the best university to get the best certificate. Usually, best universities are large universities. They are better than small universities even though they have some disadvantages.

One 4 of the disadvantages 5 of large universities is their classes is large so there may be over 100 students 6 in one class. That hinders the student to communicate easily with 7 the faculty that make the relationship between them weak 8.

Although large universities have some disadvantages, they have more advantages. 9 For example, they provide more majors and academic departments so the student have more choices and they can find their desired major. Also, large universities receive students from 10 foreign countries so you can meet people from 11 with different nationalities, so you have more options to make friends.

¹²In the end, I prefer large ~~university~~ ¹³universities to small universities because ¹⁴of their advantages.

$$\frac{11}{14} \times 100 = 78.57\%$$