THE EFFECT OF WRITTEN CORRECTIVE FEEDBACK AND REVISION ON INTERMEDIATE CHINESE LEARNERS' ACQUISITION OF ENGLISH

DAVID JON FREAR

A THESIS SUBMITTED FOR THE DEGREE OF DOCTOR OF PHILOSOPHY
THE UNIVERSITY OF AUCKLAND, 2012
DEPARTMENT OF APPLIED LANGUAGE STUDIES
AND LINGUISTICS

ABSTRACT

The study reported in this thesis builds on the research of Ellis, Sheen, Murakami and Takashima (2008) by investigating the effectiveness of corrective feedback (CF) in terms of its focus and directness. It also examines the effectiveness of CF with and without opportunities for revision. As such, the research allows for an inspection of theoretical claims made by Truscott (1996) and Ferris (1999).

The research was undertaken in a Taiwanese university context. Adopting a quasi-experimental and step-up design, intact classes were assigned to the following four CF treatment groups: Focused Direct CF, Unfocused Direct CF, Focused Indirect CF, Unfocused Indirect CF. A Control Group received no CF. The effects of the CF on two grammatical structures, English regular verbs forms and irregular verb forms, were investigated. The participants were required to complete five narrative writing tasks and a revision task over a seven week period with corrections provided on two tasks. The design of the study allowed for an analysis of the effectiveness of CF on new pieces of writing, the effectiveness of CF on a revised version of a text and the effectiveness of a combination of CF and revision on new pieces of writing.

The results demonstrated that there were improvements in accuracy in new pieces of writing for the regular past tense. However, there were none for the irregular past tense. A single episode of CF led to short-term improvements for all four CF groups; however, only

the Focused Direct CF Group showed continued improvements in the long-term (i.e. in a piece of writing completed two weeks later). The accuracy of the learners in the other three CF groups atrophied over the same period. There were no improvements in accuracy for the Control Group. There were also group differences evident in the long-term. The combined focused CF groups outperformed both the unfocused CF groups and the Control Group.

The CF groups improved in accuracy in the use of both structures in the revised version of their text. There were also group differences evident. The combined direct groups produced more accurate revisions of both structures than the combined indirect groups and the Control Groups.

The investigation of the effectiveness of CF plus revision on new pieces of writing found improvements over time for the regular past tense but none for the irregular past tense. There were short-term improvements in accuracy for the Focused Direct CF Group and the Unfocused Direct CF Group. Both these groups in fact continued to improve in the long-term. There were no improvements over time in the Control Group. In the long-term, the combined direct CF groups outperformed the indirect CF ones.

The findings provide evidence to refute Truscott's (1996) claim that CF has no effect on accuracy in new pieces of writing. However, this was only clearly evident for Focused Direct CF and only for the regular past tense. The study failed to support Truscott's assertion that writing practice alone can lead to improved accuracy. Ferris' (1999) argument that some error

types are treatable while others are untreatable was supported by the results of the study; only the regular past tense benefited from CF.

To Gary Frear, who never got to see me complete my Ph.D.,

En-Mien Liao, Mark Frear, Rachel Frear, Jan Bakker and Marten Bakker,

who accompanied me throughout the journey,

and Jack and Joy Frear, who joined us part way through.

ACKNOWLEDGEMENTS

First and foremost, I would like to thank Distinguished Professor Rod Ellis. His intellect and support were constant companions throughout the passage of this doctorate. As a supervisor and a mentor, his guidance has always been impeccable.

Dr. Rob Batstone's co-supervision was also greatly appreciated. His feedback on drafts has helped me as a writer. I would additionally like to acknowledge the long-term support and belief Rob has shown in me since he first taught and supervised me as a Master's student.

Associate Professor John Read and the Department of Applied Language Studies and Linguistics are deserved of thanks for providing me with employment during my studies. Both the financial support and experience were invaluable.

I would like to thank the staff and students of Minchuan University for allowing me to conduct the pilot study and to the staff and students of Taipei Business College for allowing me to conduct the main study with them.

I would like also to acknowledge Education New Zealand for awarding me a Post Graduate Study Abroad Award. This provided me with some of the financial assistance required to conduct the research in Taiwan.

Last of all, special mention goes also to my brother Mark Frear. I am grateful for his friendship and advice over the years. His help with the care-giving of my son was invaluable. His artwork appreciated in the production of the writing tasks.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ACRONYMS

CF	Written Corrective Feedback
L1	First Language
L2	Second Language
EFL	English as a Foreign Language
ESOL	English as a Second or Other Language
SLA	Second Language Acquisition

CHAPTER ONE

BACKGROUND TO THE STUDY

1.1 Overview

The study presented here was undertaken in a Taiwanese university context. To frame the research, a general overview of the education system in Taiwan will be presented. I will discuss how I fit into this context and how my experiences in Taiwan and New Zealand ultimately led to this study. I will then present the theoretical basis for the study and a review of key issues in written corrective feedback research before finally providing an outline of the thesis.

1.2 Educational context

Taiwan has a well-developed education system spanning preschool education through to university. Instruction of English is provided or available at all levels. There are many privately owned English language immersion kindergartens, and there are also state-run kindergartens, which have Mandarin Chinese as the medium of instruction. Students spend six years at primary school and then three years at junior high school. They need to undertake an external examination at the end of junior high school, and this exam determines the status of the senior high school the students can enter. Public senior high schools rather than private ones are deemed to be more desirable as they are perceived as having a higher quality of education. Students can, at this stage, select different tracks such as liberal arts or sciences. Junior high school students can choose to take an external exam for vocational school which

like senior high school offers a three year program. Students in such schools study subjects with a more applied orientation, for example, engineering. There are also public and private vocational schools and the performance on the aforementioned entrance exam also dictates which institution the students can enter. Again the state schools are considered better than the private ones. Following senior high school or vocational school students need to undertake another entrance exam - this time for university. The score a student receives decides whether they gain entry to the prestigious public universities or the private universities. Students can also test into technical institutes providing vocational training. It is not surprising, moreover, that with the pressure students are under to perform well in external exams that a thriving business in additional study or exam preparation exists in privately owned cram schools.

1.3 Personal context

My experience of teaching English as a foreign language (EFL) started shortly after I arrived in Taiwan in the late 1990s. A friend, who had lived in Taiwan for many years, invited me to come and stay for a while to see if I liked the culture. I did. I had always liked the idea of teaching English abroad and while I had two undergraduate degrees, I had no training or experience in English language teaching. There was, however, demand for native speakers of English. I got a job in an English immersion kindergarten during the day and a cram school at night teaching junior high school and senior high school students. After a few months, I was offered a job teaching adults at a business English firm, so I quit the other jobs and found myself teaching general English to adults onsite. It was at this time I had my first training as a teacher. It was a two day in-house training session. The teacher trainer was a firm advocate of

Krashen's (1985) Input Hypothesis and the pedagogy derived from it. He taught the group The Natural Approach (Krashen & Terrell, 1983), showed us how to plan a lesson, instructed us in board work and demonstrated pair work and group work. I was "ready" to teach. I soon became aware, however, that The Natural Approach was near impossible for me to implement as the text books we were required to use actually required task-supported language teaching (Ellis, 2003). With its notional-functional syllabus built around different grammar points, the first book I taught was "New Interchange 1" (Richards, Hull & Proctor, 1997). I also became aware that students did not respond well to those tasks I designed with The Natural Approach in mind. They asked for grammar instruction and wanted feedback on their errors. I decided that I would discard this approach and put my trust in the books I was assigned to teach.

Six months and a series of chance encounters later I found myself teaching in the language centre of the preeminent university in Taiwan, National Taiwan University. Over the next two and half years two significant events occurred. The first was I began teaching grammar classes. In accordance with the texts I was using, my grammar instruction tended to follow a present - practice - produce model. Still the learners continued to make errors with grammar, and it seemed to me that the teacher-centred method I was employing was not succeeding. Every time I taught a structure some of the students had already acquired it and had tuned out. In fact, more often than not most of the learners had. Students would be talking to friends or gazing into the middle distance. I began circulating amongst the students during the production phase providing recasts. Recasts are when a speaker makes an error and another more competent speaker provides the correct form. I would also note down common errors

and write them on the board after the production session and then provide grammar explanations. Still the errors continued. The second significant event was that I began teaching writing. With writing, I was certain that I had found the answer to my concerns through CF. Such feedback could address individual learners' errors without me standing at the front of the class boring the learners with my definitions, examples and diagrams. The type of CF that tended to be offered in the textbooks then (as it is now) was metalinguistic coded corrective feedback (e.g. Hedge, 1988; Oshima & Hogue, 2006). This comprises the provision of a code such as v.t. for verb tense when an error with a verb tense is made. I diligently corrected my students' texts and had learners produce a revised version in line with the process approach to writing. That is, the CF was provided on the penultimate draft. By the end of a course, I could see that learners had produced more accurate revisions. I felt at last I was connecting with my students, and they were learning. However, I was acutely aware of the huge gaps in my technical knowledge about language teaching. I needed further study.

I returned to the University of Auckland where I had completed my undergraduate studies. There I enrolled in an M.A. studying language teaching and learning (applied linguistics). One of my lecturers there was Dr. Rob Batstone. I had read his book on teaching grammar, "Grammar" (Batstone, 1994), and was very keen on him supervising my M.A. thesis. He agreed and introduced me to sociocultural theory, showed me different approaches to teaching grammar and helped me orient more effectively to students in the Taiwanese context - an experience for which I am still grateful.

I took myself back to Taiwan to apply what I had learned under the tutelage of Dr. Batstone and the other educators on the M.A. course. I got a position as a lecturer (instructor) at a university and began plying my trade. It was during this time that another significant event occurred. Professor Rod Ellis was presenting at a conference located not far from where I was teaching, so as I had taken some of his classes during my M.A., I decided to attend the conference. I had an opportunity to talk with Professor Ellis and shortly after he offered to supervise my doctorate. I finished my teaching contract and returned to study at the University of Auckland.

After extensive reading, I found that while there was evidence that CF was effective in a revised version of a draft, there was only limited evidence that such improvements were transferred to subsequent new pieces of writing. My earlier held beliefs that I was helping my students may not in fact have been well-founded as the evidence I had seen in support of improved accuracy in writing was largely in the form of revised versions of texts. I also found little support for the effectiveness of metalinguistic coded corrective feedback outside of revised texts. There appeared to be some evidence though that direct CF (providing a correction for an error) was successful and indirect CF (underlining that an error had occurred) might be effective. There was also research that showed that focused CF (providing feedback on a single error type or small number of preselected errors) was also advantageous for improving the accuracy of student writing. It would appear that my certainty about the effectiveness of CF was an assumption. Further reading and discussions with Professor Ellis eventually led to the design of the study presented here. In general, the study investigates the

effectiveness of written corrective feedback and revision on intermediate Chinese learner's acquisition of English. More specifically it investigates the effectiveness of corrective feedback in new pieces of writing. The CF treatments I investigated were both focused and unfocused and direct and indirect. The CF was provided with and without opportunities to produce a revised version of a text.

1.4 Theoretical context

Like oral corrective feedback, CF allows opportunities to investigate the theories and cognitive processes deemed to underlie second language acquisition (SLA). According to the revised version of the Interaction Hypothesis (Long, 1996), the cognitive processes associated with acquisition can be promoted through corrective feedback. While Long discussed these with oral communication in mind, they are also applicable to writing. Requiring learners to attend to direct or indirect CF by studying it for a period of time and having learners incorporate direct CF into a revised version of a text allows for an investigation of the likely contributions of noticing (Schmidt, 1990) and perhaps noticing plus understanding (Schmidt, 2001). When learners are asked to incorporate indirect CF into a revised version of a text or when they are asked to write without CF affords insights into the role of pushed output (Swain, 1985, 1995).

1.5 Key issues in written corrective feedback research

A central issue in written CF research concerns whether written CF is only effective on a revised version of a text or on new pieces of writing. In the case of the effect that CF may

have on new writing, a second key issue is whether different types of written CF are differentially effective.

Studies investigating the effectiveness of CF on a revised version of a text have demonstrated that by and large the provision of CF does lead to more accurate revisions than providing no CF (e.g. Fathman & Whalley, 1990; Lee, 1997; Ferris & Roberts, 2001;Truscott & Hsu, 2008). There is evidence to show that different types of CF are more or less effective depending on the explicitness of the CF (i.e. Chandler, 2003). However, demonstrating that CF is effective in a revised version of a text is of relatively little interest. Much more important is whether it results in improved accuracy in new pieces of writing as Truscott (1996) pointed out. An interesting question, however, is whether CF combined with the opportunity to revise is more effective in promoting accuracy in new writing than CF alone. This has received almost no attention, however.

The results of early studies comparing the relative effectiveness of different CF types are mixed – in part because of problems with the design of the studies. They mainly failed to provide a control group so no claims as to the effectiveness of CF on acquisition can be made. Some also failed to require learners to attend to the CF and so it is not known if the CF was actually processed or not (e.g. Semke, 1984). Still other studies did not include delayed post-tests so it is not known whether the CF had any effect on long-term learning (e.g. Lalande, 1982; Polio, Fleck & Leder, 1998, Chandler, 2003). Despite these limitations, however, some conclusions can be drawn from these studies. When learners are required to revise, they

would appear to make greater improvements over time than when corrected texts were just returned to them and they were not required to attend to the CF in some way (Chandler, 2003).

Later research examining the effectiveness of CF on acquisition has largely dealt with the design flaws evident in studies comparing the relative effectiveness of different types of CF. These studies indicate that CF is effective in improving accuracy in new pieces of writing (Ellis et al., 2008; Bitchener, 2008; Bitchener & Knoch, 2008; Sheen et al., 2009, Bitchener & Knoch, 2010a; Bitchener & Knoch, 2010b). Only one study to date has required learners to attend to CF by asking them to revise the errors that were corrected (Van Beuningen, De Jong & Kuiken, 2012). This study will be discussed in greater detail in Chapter 2.

1.6 Outline of thesis

The thesis is made up of eight chapters. Chapter 2 provides a review of the literature pertaining to CF. This comprises an evaluation of both theoretical and empirical issues in the study of written CF. Some of the gaps in the literature identified in the literature review were investigated in the pilot study presented in Chapter 3. The pilot study served as an important vehicle to develop the design and procedures of the main study. Chapter 4 contains a complete presentation of the methodology used in the study. The next three chapters present the results and a discussion of the results for the research questions investigating the effect of CF on new pieces of writing (Chapter 5), the effect of CF on a revised version of a text (Chapter 6) and the effect of CF plus revision on new pieces of writing (Chapter 7). In the

final chapter, Chapter 8, a conclusion is presented.

CHAPTER TWO

THEORETICAL AND EMPIRICAL ISSUES IN THE STUDY OF WRITTEN CORRECTIVE FEEDBACK

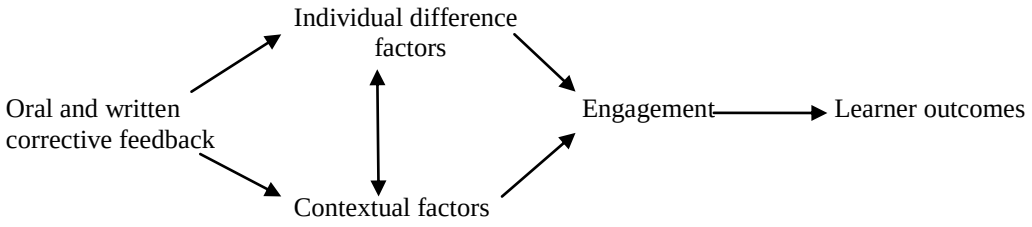
2.1 Overview

I will first present a framework for investigating oral and written corrective feedback. Then theoretical perspectives on CF will be presented before providing an overview of the history of research into CF. Next I will present a typology of CF then a discussion of the key issues in the current debate on CF in the literature. This will be followed by a review of the research into CF. Last of all, I will present the state of current knowledge on the effectiveness of CF and in so doing identify gaps in the literature which the study reported here was designed to address.

2.2 A framework for investigating oral and written corrective feedback

Ellis (2010) presents a framework for investigating both oral corrective feedback and written corrective feedback. One can see in Figure 1 that Ellis identifies six areas relevant to the study of corrective feedback and also specifies the relationships between these different parts. A brief discussion of these six areas and how they relate to each other will provide an opportunity to contextualize this review of the literature and the study that follows.

Figure 1: A componential framework for investigating oral and written corrective feedback



Source: Ellis (2010, p. 337)

The study of corrective feedback involves the development of taxonomies to identify and classify the different feedback strategies. By individual factors, Ellis refers to individual differences such as language aptitude, memory, and motivation. Such variables interact with contextual factors which can be at both the macro level (e.g. EFL and ESOL) and the micro levels (the particular setting in which the corrective feedback is received). The interaction of the individual and contextual factors serve "...to mediate between the CF the learners receive and their engagement with the CF and thereby influence learning outcomes" (p. 339). By engagement Ellis refers to how learners respond to the corrective feedback and this is influenced by the previously stated three factors. Ellis acknowledges furthermore that engagement can be seen cognitively (i.e. the way in which learners attend to the corrective feedback), behaviourally (e.g. whether learners revise following corrective feedback) or affectively (i.e. how the learners respond to the corrective feedback attitudinally). With regard to learning outcomes, Ellis notes that there has been a difference between oral corrective feedback, which has been concerned with acquisition, and written corrective feedback, which has until recently been concerned with ways to make learners better writers. The research reported in this thesis will consider types of written corrective feedback, engagement (i.e.

revision following feedback) and learning outcomes.

2.3 Theoretical perspectives on written corrective feedback

2.3.1 A definition of written corrective feedback

CF comprises the provision of written feedback on linguistic errors in writing. It excludes concerns with content and organization. It only concerns linguistic errors relating to grammar (e.g. tense and aspect), lexis (e.g. vocabulary and idiomatic expressions), mechanics (e.g. punctuation and capitalization), syntax (e.g. sentence structure) and discourse features (e.g. sentence cohesion).

2.3.2 The roles of positive and negative evidence in L2 acquisition

The roles of positive and negative evidence have been presented by Long (1996) as a means to explain his Interaction Hypothesis.

> As positive evidence, in the process of communicating they [writers and speakers] offer models of what is grammatical and acceptable (not necessarily the same) in the L2, but also instances of ungrammatical language use [e.g. simplification] at a time when learners do not know which is which. Under certain conditions they adapt their speech or writing in ways that make those models comprehensible to the learner and thereby usable for acquisition. As negative evidence, they [writers and speakers] provide direct or indirect information about what is ungrammatical. This may be explicit (e.g., grammatical explanation or overt correction) or implicit (e.g., failure to understand incidental error correction in a response, such as a confirmation check...). (p. 413)

Positive evidence then provides information about what is grammatical through comprehensible input, that is, input that a learner can understand. Ellis (2003, 2009b) notes that there are a number of different ways that input can be made comprehensible. The context can be used to make the meaning more apparent. Items can be simplified by modifying linguistic structures such as grammar to make them less complex, or they can be simplified by modifying discourse features, for instance, the provision of appositives. Input can also be made comprehensible by negotiating whether a learner does not understand or misunderstands. Comprehensible input then can be provided prior to communication (pre-modified input) or during interaction as with corrective feedback.

Negative evidence comprises the provision of information about what is ungrammatical. It is interactionally-modified input in the form of corrective feedback. Consistent with Long's definition above, Sheen and Ellis (2010) propose that oral corrective feedback can be divided on the basis of whether it is explicit (the learner is aware of the information being conveyed) or implicit (the learner is not aware) and by whether it is direct (input-providing) or indirect (output-prompting). For oral corrective feedback, there can then be implicit input-providing feedback (conversational recasts), explicit input-providing feedback (e.g. explicit correction), implicit output-prompting feedback (e.g. clarification requests) or explicit output-prompting feedback (e.g. metalinguistic clues). In relation to writing, written CF can be explicit input-providing feedback (e.g. explicit correction in the form of direct CF if a learner is required to attend to the CF) or explicit output-prompting feedback (e.g. underlining errors in the form of indirect CF if a learner is required to incorporate the corrections into a revised version of a

text).

Long (1996) provided an extensive review of the literature on positive evidence finding that comprehensible input alone was insufficient for acquisition. Evidence from Canadian French immersion programs, for example, demonstrated that learners were able to perform as well on listening and reading comprehension tests as L1 speakers (e.g. Hart & Lapkin, 1989). Learners in these contexts, however, continued to make extensive grammatical errors (e.g. Swain,1984; Lapkin, Hart & Swain, 1991). There is additionally the issue of the learnability of certain structures. White (1989) noted that certain structures or hypotheses from a L1 can be overgeneralized onto L2 structures and positive evidence cannot deal with such errors. She claimed negative evidence is required.

In light of these findings, Long suggested that the role of positive evidence is to provide interactionally-modified comprehensible input that is usable for acquisition while that of negative evidence is to provide interactionally-modified corrective feedback that can facilitate acquisition. He identified two cognitive processes that can assist acquisition through corrective feedback. These are noticing and modified-output through pushed output. Modified output occurs in uptake with repair. Long suggests these roles for positive and negative evidence occur during the negotiation of meaning, that is, when a breakdown of communication occurs between a less competent and a more competent speaker which then results in interactional adjustments that in the process of focusing on meaning draw a less competent speaker’s attention to the form.

The Interaction Hypothesis was developed with oral communication in mind, yet the roles of positive and negative evidence are equally applicable to writing; however, the delayed nature of CF means there are no opportunities to negotiate meaning in Long's sense of this term. Positive evidence through writing comes in the form of written texts that can be pre-modified to make them more comprehensible. The role of positive evidence then is also to make input comprehensible and hence usable for acquisition. Negative evidence in the form of CF is delayed. Its role is to facilitate acquisition by drawing a learner's attention to linguistic errors or, if required, pushing a learner to produce modified output by requiring the learner to revise following correction.

2.3.3 How negative evidence through written corrective feedback can assist L2 acquisition

Ellis (1994) identifies two senses of acquisition - the acquisition of a new structure and the acquisition of a partially acquired one. Negative evidence through CF can assist L2 acquisition by promoting the cognitive processes claimed to contribute to acquisition as defined in both of these ways. Each of these processes will be discussed separately.

2.3.3.1 Noticing

Schmidt (1990) defined noticing as comprising the largely conscious registration of the surface features of a structure. CF facilitates noticing by either indicating or locating that an error has occurred. It can provide the correction (direct CF) to be noticed or signal that an error has occurred (indirect CF) and learners then need to draw on their own existing linguistic resources in order to notice the correct form.

2.3.3.2 Noticing-the-gap

Providing negative evidence through CF allows for noticing-the-gap. Schmidt and Frota (1986) defined noticing-the-gap as identifying the difference between the input to which learners are exposed and the output that they can produce. It is seen as an important catalyst in the acquisition process. When learners are provided CF, they are given additional input on output they have already produced and this should trigger noticing-the-gap. Learners then have an opportunity to notice the correction if they are required to attend to the CF by studying it. Depending on the type of CF, requiring learners to attend to the CF by incorporating it into a revised version of a text can lead to opportunities to notice the correction (direct CF) or it can lead to chances for pushed output (indirect CF).

2.3.3.3 Noticing with metalinguistic understanding

Schmidt (2001) acknowledged that noticing items can also lead to metalinguistic awareness (i.e. noticing plus metalinguistic understanding). He drew on the study by Loew (1997) to illustrate this. Leow conducted a study of 28 Spanish as a L2 learners using think-aloud protocols as they completed a problem solving task (a crossword puzzle). He was able to show that learners did indeed notice structures and demonstrate rule formation and hypothesis testing as they developed metalinguistic awareness about the noticed structures. These two types of awareness can be seen as involving low levels of awareness (noticing) or high levels of awareness (noticing plus metalinguistic understanding). Leow also reported a positive correlation between the learners' levels of awareness and performance on post-task recognition and written production tests. These findings then suggest that noticing with

metalinguistic understanding is an important process that facilitates the incorporation of intake into long-term memory.

2.3.3.4 Pushed output

According to the Comprehensible Output Hypothesis (Swain,1985, 1995), learners can, during production, become aware of a gap in their own knowledge between what they want to express and what they can express if they push their output to the outer extreme of their linguistic ability. At such a moment, learners move from semantic processing to syntactic processing. That is, they go from focusing on meaning to focusing on form. Attention to the gap in their knowledge can be brought about by external feedback leading to modified output. Swain and Lapkin (1995) were able, furthermore, to demonstrate that learners engaged in rule formation and hypothesis testing when producing pushed output. As discussed above in Loew (1997), rule formation is an important factor deemed to lead to the transfer of uptake to long-term memory.

To better understand modified output, it is again useful to first consider oral corrective feedback. Lyster and Ranta (1997) discuss uptake in response to corrective feedback. This can comprise a ‘repaired’ response when a learner produces a correct structure or a ‘needs repair’ response when a learner produces an incorrect one or fails to respond to the correction. The former type leads to modified output. The claim that modified output may contribute to acquisition draws on the notion of pushed output discussed above. Lyster (2004), furthermore, takes the position that modified output represents essential practice in line with Skill Learning

Theory (e.g. DeKeyser, 1998). According to this, acquisition starts with explicit knowledge (knowledge about the language) and then is proceduralized through practice until it is finally automatized. That is, it becomes implicit knowledge (subconscious knowledge available for use in language production). Associated with Skill Learning Theory is transfer appropriate processing.

> ...according to the principle of transfer appropriate processing, the learning environment that best promotes rapid, accurate retrieval of what has been learned is that in which the psychological demands placed on the learner resemble those that will be encountered later in natural settings. (Segalowitz & Lightbrown, 1999, p. 51)

Lyster and Saito's (2010) meta-analysis of oral corrective feedback types is illustrative of the potential role of transfer appropriate processing. They found that output-prompting corrective feedback led to greater increases in accuracy than positive evidence (i.e. recasts). This they suggest may be explicable by transfer appropriate processing.

These perspectives on oral corrective feedback and modified output are equally applicable to written CF. The provision of CF such as indirect CF allows for opportunities for modified output in writing if learners are given an opportunity to 'uptake' the CF in a revised version of a piece of writing. This may provide an opportunity for learners to push their output to the limit of their current linguistic ability. In so doing, learners may move from semantic processing to syntactic processing. This pushed output can lead to hypothesis testing and rule

formation. It can also be argued from a skill-learning perspective that modified output through revising and correcting an initial draft can provide practice required for the proceduralization of explicit knowledge. Guenette (2007) argued that learners need to notice the CF and have enough opportunities to apply the corrections. Pushed output following indirect CF caters to the acquisition of partially acquired structures.

2.3.4 Conditions required for written corrective feedback to be successful

The conditions required for CF to be successful centre on the issue of allocation of attentional resources. Schmidt (2001) argues that attended processing is required for a structure to be noticed. Drawing on the work of Posner and Peterson (1990), Tomlin and Villa (1994) identified three mechanisms of attention: attention as alertness, attention as orientation and attention as detection. Attended processing then can be seen as comprising these three mechanisms. Detection, furthermore, represents the registering of stimuli in input, and for this to occur the conditions of attention as alertness and attention as orientation have to be successfully met. In other words, alertness primes orientation which, in turn, enables detection.

2.3.4.1 Attention as alertness

Attention as alertness does not directly relate to the study reported here; however, its definition is of significance to contextualize the other mechanisms of attention. Schmidt (2001, p. 17) reinterpreted Tomlin and Villa's (1994) notion of alertness as being "…related to motivation, interest in the L2, and readiness to learn".

2.3.4.2 Attention as orientation

Schmidt (2001) reconceptualised Tomlin and Villa's (1994) notion of orientation in terms of whether or not a learner's attention is directed towards form or meaning. This is consistent with the notion that attention has a limited capacity and therefore cannot deal with both form and meaning at the same time. In the case of CF, all feedback is explicit so learners are aware that they are receiving negative feedback. The very nature of written CF itself means that learners are likely to orient themselves to form over meaning. Sheen (2011), furthermore, suggests that CF can be seen as a type of input enhancement. Input enhancement comprises making input more salient through such techniques as underlining or capitalizing different structures (Rutherford & Sharwood Smith, 1985). While learners may choose to simply ignore the corrections, if they are required to attend to the CF in some way (e.g. by producing a revised version of a text), they will need to attend to form over meaning.

2.3.4.3 Attention as detection

There is some controversy over whether the role of detection is a conscious or sub-conscious event. Schmidt (2001) refers to subconscious detection as registration and conscious detection as noticing. Originally Schmidt (1990) claimed that there was no learning without noticing. Tomlin and Villa (1994), on the other hand, argued that detection can occur as registration. More recently Schmidt (2001) has acknowledged that while learning occurs through conscious noticing, it may be possible that some could occur through subconscious registration. There is some evidence that supports the effectiveness implicit learning (e.g. Doughty, 1991; Gass, Svetics & Lemelin, 2003) and studies investigating explicit learning

have also demonstrated its effectiveness (e.g. Rosa & O'Neill, 1999). There is furthermore consistent evidence that explicit learning outperforms implicit learning (e.g. N. Ellis, 1993; Robinson, 1996). However, the aforementioned study by Loew (1997) identified that there were two levels of awareness. The lower level of noticing and the higher level of noticing with metalinguistic knowledge and these differing levels of awareness correlated with the recognition and production measures used in the study. Leow (2000), furthermore, was able to demonstrate that learners showing awareness at the level of noticing or noticing plus metalinguistic understanding outperformed learners with no such awareness on post-test recognition and written production tests. These studies then provide evidence that learning is more effective when it occurs with awareness and lend support to Schmidt's position. CF, moreover, can provide input specifically designed to facilitate noticing (i.e. direct CF) or noticing with metalinguistic understanding (i.e. direct CF plus metalinguistic explanation).

2.3.4.4 Written corrective feedback, alertness, orientation and detection

There is a case to be put forward that alertness as motivation is, in fact, the most important condition for successful engagement with CF. While little research has been undertaken on CF and motivation, Uzum and Yazan (2011) found that learners with high intrinsic motivation made several moves in response to semantic corrections with oral corrective feedback whereas learners with low intrinsic motivation made fewer moves. CF constitutes a type of input enhancement, so as long as the instructions are clear and learners are required to attend to the CF in some way, they will probably focus on form over meaning. Thus, as long as learners are motivated, CF allows for an excellent opportunity to investigate the

aforementioned cognitive processes claimed to underlie acquisition.

2.3.5 The issue of delay

Unlike oral corrective feedback, which is online (immediate), CF is offline (delayed). Arguably learners have more time to engage offline. As such, when learners are required to study direct and indirect CF for a period of time or provide a revised version of a text that incorporates direct CF, learners may be able to engage in the processes that facilitate the transfer of noticed features from short-term memory to long-term memory. Similarly, asking learners to incorporate indirect CF into a revised version of a text encourages pushed output in a way that is less likely to occur with just writing practice. However, the lack of a speaker to immediately prompt a learner to produce the correct structure means that learners have to rely on their own resources. Learners who incorrectly repair their errors in a revised version of a text may develop incorrect hypotheses which could lead to the storing of an incorrect form of a structure in long-term memory.

2.4 An overview of research on written corrective feedback

Hendrikson (1978, p. 389) presented five questions about written CF. These same five questions have continued to inform research not only into oral corrective feedback but also written CF. The questions represent key areas in the existing CF literature, and, as such, an overview of research on CF will be presented in relation to the five questions. Sheen and Ellis (2010) and Sheen (2011) both reviewed the current state of the literature regarding these questions, and these reviews inform the following discussion.

2.4.1 Should learner errors be corrected?

There has been considerable debate in the literature surrounding this question. Truscott (1996) claimed that grammatical errors should not be corrected. He argued that while the existing research base provided some evidence as to the effectiveness of CF on a revised version of a text, there was none that demonstrated such an effect for new pieces of writing (i.e. acquisition). Truscott also claimed that CF actually has an inhibiting effect on acquisition. Ferris (1999) disagreed with Truscott's position arguing instead that CF has value if it is provided in an appropriate manner, namely, if it was "selective, prioritized and clear" (p. 4). She suggested, however, that more research was required, and on this issue Truscott (1999) agreed. Ferris (2004) subsequently reviewed the literature on CF also finding little evidence in support of its effectiveness. She identified, however, the key question relating to CF. Namely, "Does error feedback help L2 student writers?" (p. 50). She also presented a number of other specific areas in need of research. Research into CF since her publication has by and large investigated the question. As the review below will demonstrate, there is now evidence that CF is effective under some conditions.

2.4.2 When should learner errors be corrected?

As already noted, unlike online oral corrective feedback, written CF is offline. It is delayed. Learners typically receive CF some time after they originally produced an error. There is less certainty, however, about when to provide CF within the process approach to writing which typifies many EFL and most ESOL writing classrooms. With the process approach, writing is seen as a recursive and generative process in which learners write multiple drafts (Flower &

Hayes, 1981). Early researchers reported that the writing processes of L1 and L2 learners were the same (e.g. Zamel, 1982, 1983) leading some to suggest the CF was not necessary (Krashen, 1984). Subsequent research presented evidence supporting CF (e.g. Raimes, 1985, 1987; Arndt, 1987), and such research lead to the established position that feedback on early drafts of a composition should focus on issues of content and structure while the penultimate draft should receive CF. Recent research has led to a reassessment of the role of CF in the process of writing. Ashwell (2000) found that his measures of content and accuracy were not affected by whether the placement of content feedback or CF was on a first draft or second draft or if they were provided together. Such findings suggest that rather than dealing with issues of form and content in separate drafts, they could be addressed according to the critical needs of the students (Ferris, 2003, p. 23). Namely, if issues of form or content need to be attended to in the early drafts, they should be. Thus, there are different positions about when to provide CF within the process approach. Until research demonstrates a clear advantage for providing CF at a particular stage in the writing process, it would appear sensible to respond to errors whenever they occur.

2.4.3 Which learner errors should be corrected?

Ferris (1999) noted that CF should be selective rather than comprehensive. Selective CF comprises choosing some errors on the basis of some pre-determined criteria while comprehensive CF involves the provision of CF on all errors. Ferris (2002) went on to identify a number of criteria upon which this selection can be based. She suggested that any decision about which errors should be corrected should be informed by knowledge about the

most common errors, the effect of learners' English language learning background, their level of proficiency and the influence of the learners' L1. She recommended that teachers should select global errors rather than local errors for correction. Global errors are those that interfere with a reader's comprehension of a text while local errors do not. Another factor to consider is the frequency of errors, and those that a particular assignment elicits. In her earlier article, Ferris (1999) also proposed that some errors are treatable in that they follow some type of rule or system while other errors are untreatable in that they are lexical or idiomatic (e.g. They do not follow a regular rule-governed system).

Selective written CF is 'focused'. Focused CF comprises the provision of CF on one or a few preselected structures whereas unfocused CF provides CF on all or a range of structures. Indeed, the studies that have investigated the effectiveness of CF on acquisition have largely adopted a focused approach to providing CF.

2.4.4 How should learner errors be corrected?

One can discern two perspectives in the literature about how learner errors should be corrected. The first of these can be seen as a pedagogic perspective. Lalande (1982, p. 140) suggested that indirect CF (e.g. metalinguistic codes such as v.t. for verb tense) is beneficial because it permits 'guided learning' and 'problem solving'. This prompts learners to self-correct their errors by editing their texts. Ferris (2002) has also suggested that the use of indirect CF is best used with more advanced learners whereas direct CF (the provision of the correction) may be more beneficial for learners of lower levels of proficiency. However, the

provision of direct CF may not be the best option as this provides learners with the corrections and thus fails to demonstrate what they can write based on their own linguistic competence.

The second perspective can be considered a SLA perspective informed largely by the aforementioned cognitive processes claimed to underlie acquisition. Pedagogy needs to take account of these processes and whether some are more effective than others. Referring to oral CF, Long (1996) proposed that input-providing oral corrective feedback in the form of recasts is the most effective type of corrective feedback because it provides both positive and negative evidence. In other words, it provides opportunities for noticing. Lyster (2004), on the other hand, suggested output-prompting corrective feedback should be the more effective as the instances of modified output provide opportunities for practice that help proceduralize a structure. Lyster and Saito's (2010) meta-analysis of oral corrective feedback, in fact, found that both input-providing and output-prompting CF were effective over time; however, the effect sizes were larger for output-prompting CF.

While the research base is still incomplete for CF, it is evident that the two perspectives on how learner errors should be corrected may potentially be in conflict. Indeed, looking for a universal answer of how to correct errors may not be the best solution. Sheen and Ellis (2010) suggest that CF cannot be separated from the educational and social goals of learners; that is, it is context dependent. Thus, no one type of CF will suit all situations. They suggest rather than looking for a universal cure-all, teachers may want to explore different strategies in their own unique contexts.

2.4.5 Who should correct learner errors?

In relation to the issue of who should correct learner errors, there are a number of possible choices. These are the teacher provides the correction, the student self-corrects following CF, students peer correct each other's compositions or the students self-correct without CF. There can also be different combinations of these options. Hendrikson (1980) suggested that the provision of corrections by teachers should not predominate but rather learners be given the opportunity to self-correct their errors, so they can become independent self-editors. It is also suggested that peer-correction is of value (e.g. Zamel, 1985). The problem with both of these positions is that the students may not be able to identify and correct errors without assistance. A possible solution is to combine student self-correction or peer-correction with teacher correction. In other words, first have the students attempt to correct the errors and then have the teacher check the answers and provide corrections for those responses that were incorrect (e.g. Chandler, 2003). Krashen (1984) suggested that students should try to self-correct without access to any external CF. This approach has the same problem discussed above. Students may or may not be able to correct errors. They may furthermore fail to identify their errors. It is of course possible to have learners attempt to correct their own errors, then provide indirect CF so they can self-edit and finally provide direct corrections on the remaining errors but this would be very time-consuming and impractical in many teaching contexts.

2.5 Ellis' typology of written corrective feedback

Ellis' (2009a) typology of CF is informed by the cognitive theories of CF outlined above. It

draws on existing pedagogic systems of classification of CF (e.g. Ferris & Hedgcock, 1998); however, Ellis' typology differs from such pedagogically derived typologies in a number of ways. As can be seen in Table 1, it distinguishes teachers' strategies for providing CF and students' options for responding to them; it also isolates additional strategies and responses. It distinguishes direct, indirect and metalinguistic CF, and for the CF responses, it presents options involving whether or not to revise. The table also provides broad descriptions of some of the options available to teachers in implementing the strategies, and descriptions of some responses available to learners. The different types of CF can be used separately or in combination with each other.

Table 1: Types of teacher written corrective feedback (Ellis, 2009a)

Type of CF	Description	Studies
A Strategies for providing corrective feedback		
1 Direct CF	The teacher provides the student with the correct form.	e.g. Lalande (1982) and Robb et al. (1986).
2 Indirect CF	The teacher indicates that an error exists but does not provide the correction.	
a indicating + locating the error	This takes the form of underlining and use of cursors to show omissions in the student's text.	Various studies have employed indirect correction of this kind (e.g. Ashwell 2000; Ferris and Roberts 2001; Chandler 2003).
b indication only	This takes the form of an indication in the margin that an error has taken place in a line of text.	Fewer studies have employed this method (e.g. Robb et al. 1986)
3 Metalinguistic CF	The teacher provides some kind of metalinguistic clue as to the nature of the error.	
a Use of error code	Teacher writes codes in the margin (e.g. ww = wrong word; art = article).	Various studies have examined the effects of using error codes (e.g. Lalande 1982; Ferris and Roberts 2001; Chandler 2003).
b Brief grammatical descriptions	Teacher numbers errors in text and writes a grammatical description for each numbered error at the bottom of the text.	Sheen (2007) compared the effects of direct CF and direct CF + metalinguistic CF.
4 The focus of the feedback	This concerns whether the teacher attempts to correct all (or most) of the students' errors or selects one or two specific types of errors to correct. This distinction can be applied to each of the above options.	Most studies have investigated unfocused CF (e.g. Chandler 2003; Ferris 2006). Sheen (2007), drawing on traditions in SLA studies of CF, investigated focused CF.
a Unfocused CF	Unfocused CF is extensive.	
b Focused CF	Focused CF is intensive.	

5 Electronic feedback	The teacher indicates an error and provides a hyperlink to a concordance file that provides examples of correct usage.	Milton (2006).
6 Reformulation	This consists of a native speaker's reworking of the students' entire text to make the language seem as native-like as possible while keeping the content or the original intact.	Sachs and Polio (2007) compared the effects of direct correction and reformulation on students' revisions of their text.
B Students' response to feedback	For feedback to work for either redrafting or language learning learners need to attend to the corrections. Various alternatives exist for achieving this.	
1 Revision required		A number of studies have examined the effect of requiring students to edit their errors (e.g. Ferris and Roberts 2001; Chandler 2003).
2 No revisions required		
a Students asked to study corrections		Sheen (2007) asked students to study corrections.
b Students just given back corrected text		A number of studies have examined what students do when just given back their text with revisions (e.g. Cohen and Cavalcanti 1990; Sachs and Polio 2007). No study has systematically investigated different approaches to revision.

Source: Ellis (2009a, pp. 98-99)

2.5.1 Corrective feedback strategies

2.5.1.1 Direct corrective feedback

Direct CF comprises the provision of the correct form by a teacher. It involves the crossing out of errors and the writing of corrections above them or nearby in the margin. The corrections can be directed at errors in morphemes, words or phrases, and missing elements can be inserted into writing using cursors. Examples of direct CF are provided in Table 2.

Table 2: Examples of Direct Corrective Feedback

Direct CF		walked home ^ and sandwich I ~~walk~~ ~~hame~~ yesterday ate a ~~sandwiche~~s.

2.5.1.2 Indirect corrective feedback

Indirect CF involves signaling that an error has occurred but does not provide a correction. There are two types. Indirect indicated CF provides some indication in the margin that an error exists; however, it does not show the location. This can be done by writing a √ for lines with no errors and an X for those with errors. Learners have to first locate the errors before being able to attend to them. Indirect located CF shows the location but does not provide the correction. It can comprise the underlining or circling of errors as well as the use of cursors for missing words. As the errors are located, a learner needs only attend to the errors. Learners have to draw on their own linguistic resources and thus are likely to be most

effective if they have already partially acquired the features being corrected. Table 3 provides examples of indirect CF.

Table 3: Examples of Indirect Corrective Feedback

Indirect CF	a. Indirect indicated CF	X	I walk hame yesterday ate a sandwiches.
	b. Indirect located CF		^ I walk hame yesterday ate a sandwiches.

2.5.1.3 Metalinguistic corrective feedback

Metalinguistic CF entails the provision of metalinguistic clues about an error, but it does not provide a direct correction. There are two types. The first comprises metalinguistic feedback that utilizes error codes. The error codes can include such codes as ‘v.t.’ for verb tense, ‘sp.’ for spelling, ‘conj.’ for conjunctions and ‘pl.’ for plurals. Metalinguistic coded CF varies according to whether an error is located in the text or not. Metalinguistic coded CF involves just writing the codes in the margin. Located metalinguistic coded CF specifically locates the errors by writing the codes above the errors or by writing the codes in the margin and underlining or circling them in the text. The second type of metalinguistic feedback locates the errors and provides brief metalinguistic explanations about them. This can involve writing a number above the errors and listing the numbers and explanations at the end of the text, or,

alternatively, it can comprise underlining errors and writing the explanations next to the errors in parentheses. As the name suggests, metalinguistic CF can be seen as developing metalinguistic knowledge about a linguistic structure. Learners have to draw on their existing linguistic resources so it is likely to be more effective for partially acquired structures. Examples of metalinguistic corrective feedback are presented in Table 4.

Table 4: Examples of Metalinguistic Corrective Feedback

Metalinguistic CF	a. Metalinguistic coded CF	v.t.; sp.; conj.; pl.	I walk hame yesterday ate a sandwiches.
	b. Located metalinguistic coded CF		conj. v.t. sp. ^ pl. I walk hame yesterday ate a sandwiches.
	c. Metalinguistic explanation		1. 1 2 3 4 I walk hame yesterday ate a sandwiches. 1. Use the past tense for completed actions and states in the past. 2. The correct spelling is home not hame. 3. A conjunction is needed between two main verbs in a single independent clause. 4. A plural noun form is not used with a singular determiner.

2.5.1.4 Focused versus unfocused feedback

Feedback can be either focused or unfocused. Focused CF involves concentrating on one or a few pre-selected forms. Unfocused CF involves providing feedback on all or many errors. The focused-unfocused distinction is a continuum, and it can be applied to the various CF strategies discussed here. Table 5 presents examples of focused direct CF and unfocused direct CF. The focused condition provides feedback on only a single form, the regular simple past tense. It has been suggested that focused as opposed to unfocused CF potentially reduces the 'attentional strain' on learners (Sheen, 2007). As such when focused CF is combined with direct CF, it should increase the likelihood of noticing errors, and when it is combined with indirect CF, it could enhance the chances for modified output.

Table 5: Examples of the Focus of Feedback

The focus of the feedback	a. Focused CF (e.g. focused direct CF of regular simple past tense)		walked I ~~walk~~ hame yesterday ate a sandwiches.
	b. Unfocused CF (e.g. focused direct CF)		and walked home ^ sandwich I ~~walk~~ ~~hame~~ yesterday ate a ~~sandwiches~~.

2.5.1.5 Electronic corrective feedback

Electronic feedback refers to the utilization of corpora of written English for the purpose of

feedback. This is achieved by providing learners with the resources to self-edit through comparing their usage with that available in a corpora. In other words, students can refer to more experienced writers' constructions. Learners then subsequently revise a draft. Such an approach encourages independent learning. Ellis discussed Milton's (2006) software package "Mark my Word" as an example of electronic feedback. This software package has a store of 100 common lexical, grammatical and style errors made by Chinese users of English. These errors have a comment about their correct usage and links to examples of correct usage. It also generates an error log of learners' frequent errors. Thus a teacher can insert a metalinguistic comment and a link to correct forms. This software provides corrections and metalinguistic information and as such caters for the acquisition of both new and partially acquired structures.

2.5.1.6 Reformulation

Reformulation involves a native speaker reformulating a whole text while attempting to maintain the integrity of the meaning expressed by the writer. It can be seen as a kind of direct CF. The learner then needs to identify the changes that have been made from the original by comparing it with the reformulated version. Reformulation caters for the acquisition of both new and partially acquired structures.

2.5.2 Learner responses to written corrective feedback

2.5.2.1 Revision required

Revision involves requiring students to produce a new draft incorporating corrections, and

learners typically have access to their corrections during this process. Irrespective of the type of CF, when revision is required, the CF becomes output-prompting CF; however, such CF promotes different cognitive processes. When revision is required with direct CF, direct CF plus metalinguistic explanation or reformulation, it represents a mechanical process of accepting and then copying the corrections. It can therefore be seen as providing an opportunity to notice a structure and perhaps notice with metalinguistic understanding. If revision is required with indirect CF or metalinguistic CF, it allows an opportunity for a learner to provide modified output. Revising a draft without any CF provides an opportunity for self-correction and also for modified output.

2.5.2.2 Revision not required

This category is separable into two further kinds of response. These are studying the CF and students being given back corrected texts. Sheen (2007), for example, required students to study the CF for a period of time. This option, then, is an attempt to ensure that the students attend to CF. For all types of CF, requiring the learners to attend to the CF by studying the CF promotes noticing and perhaps noticing with metalinguistic understanding. Requiring learners to study their writing when no CF is provided allows for learners to self-correct.

The second response is that students are returned their corrected texts but are not required to attend to the CF. Whether a student attends to the correction or not is left up to the individual. Thus, learners may or may not notice the corrections.

2.6 The debate about the efficacy of written corrective feedback

The debate in the literature started with Truscott's (1996) article that claimed CF directed at grammatical structures was pointless as it does not lead to improvement in the accuracy of writing. He gave the following reasons in support of his stance.

> (a) Research evidence shows that grammar correction is ineffective; (b) this lack of effectiveness is exactly what should be expected, given the nature of the correction process and the nature of language learning; (c) grammar correction has significant harmful effects; and (d) the various arguments offered for continuing it all lack merit. (p. 328)

Each of these claims will be considered separately and in so doing, responses from key contributors to the debate discussed. In some instances, relevant research relating to these claims will be drawn upon to assess their validity. Last of all, a brief discussion of the value of the debate will be presented.

2.6.1 The claim that 'research evidence shows that grammar correction is ineffective'

Truscott (1996) reviewed the L1 and L2 literature on CF finding little evidence in support of its effectiveness. Of the L2 research, he initially looked at those studies that found no effect for CF (e.g. Semke, 1984; Robb et al., 1986). By inspecting and comparing the methodologies employed in these studies, he dismissed the influence of the English as a foreign language (EFL) versus English as a second or other language (ESOL) context, the target languages, the learners L1, the type of CF and the focus of the feedback. He additionally discounted the possibility of a delayed effect for grammar on grammar learning, the influence of the

measures of accuracy employed in the studies, the instructional approaches used in the classes, the requirement to have learners revise or not and the proficiency of the learners. All in all, he found the studies robust enough in their methodological designs to support their findings that CF had no influence on improving learners' written accuracy.

Truscott also addressed the studies that found some positive effect for CF. He rejected some studies as not being research-based then dismissed the findings from research into oral CF as not directly related to writing. This left two remaining studies. One found a positive effect for CF on a revised version of a text (Fathman & Whalley, 1990). Truscott questioned whether such benefits would have any subsequent effect on improving writing. Another study found one type of CF better than another (Lalande, 1982). In this instance, Truscott noted that the failure to provide a control group receiving no CF meant it was impossible to tell whether it was the CF or writing practice that was responsible for the improvements in accuracy.

Ferris (1999) disagreed with some of the claims made by Truscott. She cited studies that suggest CF is effective for some learners if it is focused, prioritized and clear (e.g. Ferris, 1995; Ellis 1998). She also questioned the validity of Truscott's conclusions in his review of the research. She argued the studies were not comparable due to the different types of subjects. In particular, she pointed out that only one study was in an ESOL context - a context in which the process approach motivates learners to correct and revise errors. She doubted the comparability of the studies' research methodologies and approaches to teaching. Some studies, for example, were cross-sectional and some longitudinal. She suggested finally that

Truscott overstated the findings about the ineffectiveness of CF and understated those studies that found in favour of CF. Ferris called for further research to better understand the complex nature of CF.

Truscott (1999) responded to Ferris' concern about the comparability of studies due to their different samples, instructional methodologies and design. He argued that such variability in fact makes the findings all the more valid as the results were constant irrespective of this variability. He rejected Ferris' claim that he overstated evidence in support of his argument and understated those in favour of CF. He agreed with Ferris, however, about the need for further research.

Consistent with her assertions regarding Truscott's interpretation of the existing research base, Ferris (2004) undertook a review of the studies examined by Truscott and those published since his review. She made three observations.

> "(1) the research base on the 'big question' - does error feedback help L2 student writers? - is inadequate; (2) the previous studies on error correction are fundamentally incomparable because of inconsistencies in design; and (3) existing research predicts (but certainly does not conclusively prove) positive effects for written error correction." (p. 50)

In light of these observations and drawing on some of Truscott's earlier statements, Ferris made some recommendations for the future direction of CF research. Ferris proposed that in

order to answer the big question about the efficacy of CF, there was a need for longitudinal controlled studies, that is, studies that investigated whether the accuracy of student writing improves over time, and she noted that such studies require a control group receiving no CF. She also suggested, in contrast to Truscott's view, that studies should be comparable in design and also reported clearly so as to be replicable. Last of all, she suggested that there is a need to undertake 'finely tuned' studies that investigate the following issues: the role of revision on learners' written accuracy; the effect of supplementary grammar instruction on students' written accuracy; the function and effect of error logs; the treatable or untreatable nature of different types of error; the effect of the explicitness of CF on the accuracy of learners' writing.

2.6.2 The claim that 'grammar correction is ineffective because of the nature of the correction process and the nature of language learning'

While Truscott originally divided this section into two parts, a closer inspection of his article identifies three areas that can be considered under this heading: individual differences (learner variation), theoretical problems and practical problems.

2.6.2.1 Individual difference issues

Truscott claimed that individual differences were so complex and interrelated that it is unlikely that these factors could be isolated so CF could be used to benefit learners. Recently, however, Sheen (2011) investigated the role played by individual differences. She reports on a study that investigated the mediating effects of language analytic ability (a component of

language aptitude), anxiety and learner attitudes. She found significant correlations between language analytic ability and total test scores in a metalinguistic explanation plus direct CF group and in a direct CF group in both the short- and long-term. She reported that language analytic ability had a mediating effect on the two types of CF investigated. She also demonstrated non-significant and low-level correlations between short-term and long-term gain scores and her measure of anxiety for the two CF groups. She concluded that anxiety does not have a mediating effect on written CF. In relation to learner attitudes, Sheen established significant large correlations in the short-term and the long-term between positive attitudes to CF and gain scores for the two CF groups. Sheen found that learners with positive attitudes towards CF received greater benefits from CF. All in all, this research casts some doubt on Truscott's claims as to individual differences being too complex and inter-related as to be isolated so CF could be used to benefit learners. However, there is clearly a need for further studies that examine the mediating role of individual difference factors on different samples.

2.6.2.2 Theoretical issues

Truscott's first concern was that syntactic, morphological and lexical knowledge is more than likely acquired in different ways, so one CF type will probably not be effective with all three. This was a position with which Ferris (1999) agreed. Indeed, she introduced the notions of treatable and untreatable errors as a way of dealing with these differences. Ferris' view is consistent with SLA perspectives on learning. Ellis (2009b) distinguishes between structures that involve system-learning and item-learning. Treatable errors then can be seen as involving

system-learning whereas untreatable errors involve item-learning.

Another of Truscott's theoretical concerns involved the notion of psuedolearning. Psuedolearning represents the learning of knowledge about a language (explicit knowledge) and is distinguishable from knowledge as acquisition represented by the successful production of a language which requires implicit knowledge. Reflecting Krashen's (1984) view on grammar correction and his Monitor Hypothesis (1985), Truscott suggests that CF directed at grammar may contribute to such psuedolearning and ceded that such knowledge may have benefits for editing; however, he then argued that the temporary and isolated nature of psuedolearning means that such benefits would be limited. In essence, Truscott is taking a non-interface position about the relationship between explicit knowledge and implicit knowledge. However, this is an area of dispute in the literature. The non-interface position claims there is no relationship between explicit and implicit knowledge (e.g. Krashen, 1981). This stance posits that positive evidence is nearly wholly responsible for acquisition; however, as Long (1996) showed in his comprehensive review of the role of positive evidence, positive evidence is insufficient to account for acquisition. The interface position posits a direct relationship between the two types of knowledge (e.g. Sharwood Smith, 1981). The weak non-interface position views explicit knowledge as facilitating implicit knowledge (e.g. Ellis, 1994). Certainly the aforementioned studies by Loew (1997, 2000) would appear to suggest that explicit knowledge does in fact facilitate implicit knowledge. Truscott's position then is contentious.

The final theoretical area that Truscott discussed related to developmental sequences. While he is careful to hedge many of his claims, Truscott would appear to be taking the stance that there is a natural order of acquisition. He states for example that "...correction that respects natural sequences of acquisition is not realistic now and is not likely to be soon" (1996, p. 345). In his review of studies investigating developmental sequences, Ellis (2009b) reported that there is some evidence that tense, aspect and syntax would appear to follow recognizable sequences of development in a particular L2. Truscott's position in this instance would appear to be supported by the research. However, if one adopts a weak non-interface position concerning the relationship between explicit knowledge and implicit knowledge, positive evidence in unison with CF as negative evidence can be seen as facilitating L2 development. In other words, CF may work because it raises learners' consciousness about what is correct and incorrect and this explicit understanding feeds into implicit knowledge by facilitating noticing and noticing-the-gap over time. From this perspective, then, the natural sequence of acquisition is not an impediment to CF.

2.6.2.3 Practical issues

Truscott noted that identifying errors is difficult, errors are often missed by teachers and teachers are ill-equipped to explain errors. He argues that poor explanations could lead students to make further errors in the future. Truscott also suggests that correction is a de-motivating experience for learners. Ferris (1999) agreed with Truscott's concerns regarding the effectiveness of CF in light of the practical shortcomings of both teachers and students. She argued, however, that such problems are surmountable. For teachers, specific training and

the selective prioritizing of error types can overcome the problems identified by Truscott. Students' practical problems with CF can be alleviated through consideration of the learners' L1, their English language proficiency and prior experience with grammar pedagogy and editing practice. Truscott's claim as to the de-motivating effect of CF has not been demonstrated. Some research has been undertaken on CF and motivation (e.g. Uzum & Yazan, 2011), and other research is suggestive. Ferris, for example, identified survey research that reports on learners positive opinions towards receiving CF (e.g. Cohen, 1987; Leki, 1991), and as discussed above Sheen (2011) found positive correlations between students attitudes towards CF and their performance. However, there is clearly a need for further research that investigates motivation in relation to CF.

2.6.3 The claim that 'grammar correction has significant harmful effects'

Truscott finally presented his view on why grammar correction has harmful effects. He argued that the provision of CF induces stress in learners and this in turn affects the learners' ability to learn. He additionally suggested that in the face of correction learners will avoid errors by shortening and simplifying their written texts. Another concern he voiced relates to time. Students have to spend a considerable amount of time on correcting errors and this time could be spent on more productive learning activities. Teachers also have to spend sizable portions of their time correcting errors and this energy he argues could better be used with feedback on issues of content and organization.

2.6.3.1 Stress

With regard to Truscott's claim that the provision of CF induces stress which, in turn, affects students' ability to learn, he would again appear to be drawing on the work of Krashen, in particular, Krashen's (1985) Affective Filter Hypothesis. Like Krashen, Truscott's position is that CF creates anxiety and thus impedes learning. However, as mentioned above, Sheen's (2011) research does not support Truscott's claim. Sheen found that learners' levels of anxiety were not affected by written CF.

2.6.3.2 Avoidance

Truscott (2004) developed his second concern relating to avoidance of errors by critiquing Chandler's (2003) studies. Truscott rejected Chandler's interpretation of her results and argued that in addition to the ESOL context and writing practice, avoidance provides an alternative explanation for her results. According to Truscott, learners avoid certain constructions and thereby reduce the complexity and length of their writing. In so doing, the likelihood of producing certain errors is reduced. Truscott also questioned Chandler's finding that fluency increased over time. As a measure of fluency, Chandler used student reports on how long it took them to write a text. Truscott doubted the accuracy of self-reports. He suggested furthermore that avoidance can account for the increase in speed. He claimed that by avoiding certain grammatical structures the learners actually simplified their use of the language and this is why they wrote quickly.

Chandler (2004) offered a response to this critique. She noted that both groups had the same

amount of writing practice and neither groups' writing was shorter and both, in fact, increased on her measure of fluency. As for the issue of avoidance leading to shorter and less complex writing, she reiterated that her students' writing did not reduce in size. She maintained that the findings of her holistic measure of writing quality provided some evidence that the learners writing did not become simpler. She also noted that due to the unfocused nature of her feedback, it would be difficult for learners to avoid so many different types of error. The debate has continued in both Chandler (2009) and Truscott (2007, 2009). Truscott largely reiterated his position. However, Chandler responded to Truscott's concerns regarding error avoidance by conducting a post-hoc analysis of her data. She demonstrated that for her study error avoidance can largely be discounted as a variable affecting fluency and complexity.

2.6.3.3 Written corrective feedback is an inefficient use of time

The last of Truscott's concerns regarding the harmful nature of CF involves his claim that correction uses up valuable time that could be used on issues of content and organisation. While undeniably the provision of CF is a time consuming act, this is particularly the case with unfocused CF. Focused CF is less time consuming. Nevertheless, irrespective of CF, well-trained and dedicated teachers also provide appropriate feedback on content and organisation.

2.6.4 The claim that 'various arguments offered for continuing it all lack merit'

Last of all, Truscott outlines his view that there is no reason to correct grammatical structures in writing. Here he introduces the burden of proof argument according to which the onus is on

teachers and researchers who advocate the use of CF to demonstrate that it is effective. Strong intuitions about its effectiveness are not enough. He rejects claims that CF is beneficial because learners cannot identify their errors without it, and he rejects the assertion that the failure to correct will lead to fossilization. In support of this he refers to the evidence and arguments presented above. He also refutes the evidence that students want to be corrected, arguing again that teachers should educate learners on the best possible strategies to improve their learning. Truscott suggested that grammatical accuracy develops through extensive reading and writing practice, not through CF.

Ferris (1999) considered reasons why error correction should be continued. She cites survey research showing L2 learners' affirmative opinions about grammar correction (e.g. Ferris, 1995; Hedgecock & Lefkowitz, 1994). She also notes that some L1 content faculty are less tolerant of L2 learner errors than of L1 speaker errors. Most importantly she notes that L2 learners need to become more proficient at self-editing their writing.

2.6.5 The value of the debate

The debate has helped refine and identify key issues surrounding the provision of CF. It is clear that Truscott's position has largely been informed by Krashen's (1985) Input Hypothesis which claims positive evidence alone is responsible for acquisition - a position that has largely been discredited as a result of extensive research (e.g. Long, 1996). This is one reason why Truscott's claims have largely been challenged. As a result of the debate between Truscott and Ferris, however, well-designed studies are now being conducted to address the

efficacy of CF. Some of Ferris' 'fine tuned' research questions though have yet to be investigated. Issues relating to individual differences are now being examined, however. As a result of the practical problems involved with CF, concerns surrounding the procedures for implementing CF through extensive training of teachers have been addressed. Truscott and Chandler's debate has led to studies investigating whether avoidance results from CF and whether CF affects the fluency and complexity of writing. Clearly the debate has led to a clearer identification of the central issues surrounding CF.

2.7 Issues of research methodology

The literature on CF can be separated into research that has examined the effectiveness of CF on a revised version of a text and the effectiveness of CF on new pieces of writing. Of the latter, there are two types. First, there are those studies that have compared the relative effectiveness of different CF types. These studies often failed to provide a control group. Second, there are studies that have investigated the effectiveness of CF on acquisition. Such studies did provide a control group. These two types of studies will be reviewed before presenting a discussion of the methodological limitations of the research.

2.7.1 The effect of written corrective feedback on a revised version of a text

Fathman and Whalley (1990) conducted research with 72 intermediate-level ESOL learners at a U.S. college. Adopting a quasi-experimental design (pre-test - treatment - post-test), the treatment comprised four conditions - unfocused indirect located CF, content comments, content comments with unfocused located CF and a control group. The pre-test involved a

picture sequence task, and the revised version of the text constituted the post-test. The participants had access to their feedback when they revised their text. Grammatical scores were generated based on the number of errors per text. The researchers found that both conditions receiving indirect located CF led to a significant improvement in accuracy between the two time periods whereas the other two conditions did not.

Lee (1997) undertook research with 141 low intermediate-level EFL students at a Hong Kong university. She investigated her participants' ability to correct a newspaper article seeded with 20 grammatical errors of different types (e.g. articles, tense, agreement, spelling). After dividing her participants into three conditions - unfocused indirect located CF, unfocused indirect indicated CF and no CF - she then required the members of these groups to correct the texts. Mean scores were calculated based on the number of successfully corrected errors. The indirect located CF group significantly outperformed the other two conditions.

Ferris and Roberts (2001) conducted a study with 72 intermediate-level ESOL students at an American university with the purpose of investigating the participants' ability to revise following three types of CF - unfocused located metalinguistic coded CF, unfocused indirect located CF and no CF (the control group). The participants wrote one of two persuasive essays. These were corrected and the participants had to subsequently correct their own essays. The learners had access to the CF when they corrected their essays. A score was generated by calculating the percentage of errors successfully corrected. They found that there were significant differences in the accuracy scores between the control group and both the

indirect CF and metalinguistic CF conditions. That is, the revised versions for both the CF groups were significantly more accurate than the no CF group.

Chandler (2003) investigated the effectiveness of four CF conditions on a revised version of a text. The groups were an unfocused direct CF group, an unfocused located metalinguistic coded CF group, an unfocused metalinguistic coded CF group and an unfocused indirect located CF one. Thirty-six intermediate-level ESOL students from an American university were included in the quasi-experimental study. The study used a partially balanced incomplete block design. That is, the participants completed four different autobiographical journal texts and for each text they received a different type of CF. Thus, they all experienced the four types of CF but received the CF in different orders. They were required to revise the texts following the receipt of the CF. This design followed a pre-test - treatment - post-test structure. The pre-test was a journal entry, the treatment the CF and the post-test the revised text following the CF. The measure of accuracy used was mean error rate. Chandler reported that there were significant differences between the groups; however, she failed to undertake any post-hoc between-group analyses for group differences. She reported that there were differences in the levels of accuracy reflected by the explicitness of the CF. The unfocused direct CF group was the most accurate followed by the metalinguistic coded CF group then the unfocused metalinguistic CF group and finally the unfocused indirect located indirect CF group.

Truscott and Hsu (2008) conducted research on 47 EFL students at a Taiwanese university.

Adopting a quasi-experimental design, the treatment comprised an unfocused indirect located CF condition and a no CF one. Following the completion of a picture sequence narrative writing task (pre-test), the CF was supplied and the participants asked to revise the original draft (post-test). Students had access to the CF when revising. Both sets of writing were scored in terms of error rate calculated as the total number of errors divided by total number of words written. A one-way between groups ANOVA was then undertaken on this score. It was found that the unfocused indirect located CF group outperformed the control group.

It would appear from this review of studies that the provision of CF enhances the accuracy of learner revisions. There is also evidence that when learners wrote a revised version of a text that incorporates CF, learners who received direct CF wrote more accurate revisions than learners receiving metalinguistic CF and indirect CF who in turn produced more accurate revisions than learners who received no CF (control groups). A concern with these studies, however, is that demonstrating that CF is effective in a revised version of a task does not necessarily mean that such improvements in accuracy will be passed on to subsequent new pieces of writing (Truscott, 1996). That is, there is no evidence that improvements in accuracy will lead to the acquisition of linguistic structures.

2.7.2 The effect of written corrective feedback on new pieces of writing

2.7.2.1 Studies that have compared the relative effectiveness of different corrective feedback types

Lalande (1982) conducted research on 60 intermediate-level German as a foreign language

students at a U.S. college. The design was quasi-experimental (pre-test - treatment - post-test). Five compositions were completed over a ten week period. They were described as plot summaries, and the first and fifth of these represented the pre-test and post-test, respectively. There were two treatment conditions: unfocused direct CF and located metalinguistic coded CF. CF was provided on the second, third and fourth compositions, and the participants were required to revise these texts. Members of the metalinguistic CF condition were also allowed to use reference grammar texts and if unable to find the correct answer they could ask for teacher or peer assistance. For lexical errors, written explanations were added to the writing. Finally, this group was also required to complete error logs. The tests were scored for the percentage of errors and Fisher paired t-tests were undertaken on the pre-tests and post-tests. Lalande reports that there were no significant increases in accuracy over time for either of the groups; however, for orthographic errors (spelling, capitalization and punctuation) there were for the metalinguistic CF group. He also reports there were no group differences in the pre-test but there were in the post-test. The metalinguistic condition significantly outperformed the direct CF group.

Semke (1984) undertook research with 141 German as a foreign language students at an American university, using a quasi-experimental design (pre-test - treatment - post-test). The pre-test and post-test both comprised free-writing samples. These were timed and the participants had ten minutes to complete them. The treatment comprised four conditions - unfocused direct CF, content comments, unfocused direct CF with content comments and unfocused located metalinguistic CF. The feedback was given on weekly journal entries over

a ten week period. Only the participants in the metalinguistic condition were required to revise. Accuracy was scored on the writing tests by using the ratio of errors to the number words produced. Using a one-way ANCOVA on the post-test scores and using the pre-test scores as the covariate, the researcher found that there were no differences in accuracy between the conditions.

The study by Polio et al. (1998) involved 65 ESOL learners at an American university. It was quasi-experimental in design (pre-test - treatment - post-test). The two tests involved writing two descriptive essays. The pre-test was conducted in week two and a revised version of the text was completed. The post-test was 13 weeks later. The treatment involved two groups. An unfocused direct CF group completed two journal entries a week, undertook editing practice and had grammar instruction. Unfocused direct CF was provided on the journal entries and the editing practice, and the participants were required to revise only one of the two entries per week. An extensive free-writing practice group completed four journal entries a week with no feedback whatsoever on their writing. Using measures based on error-free T-units, the researchers found that both conditions improved significantly over time but there were no group differences between them. There was also no time-group interaction.

Chandler (2003) reports on two studies she undertook. In the first study, she conducted research on 31 post-intermediate level ESOL students at an American college and used a quasi-experimental design comprising a pre-test, treatment and a post-test. The pre-test represented the first journal entry and the post-test the fifth, which was completed 10 weeks

after the first. There were two treatment conditions. The first group received unfocused indirect located CF, were required to revise following feedback and were subsequently given direct CF on any errors they had missed in their revision. The learners were not required to attend to this direct CF. The second group had unfocused indirect located CF and were not required to revise following CF. That is, their writing was just returned and they were not required to attend to it. Error rates were calculated as the mean number of errors per 100 words. The study found that students who were required to revise increased in accuracy over time while those that were not failed to do so. There were, furthermore, no group differences for the pre-test, but for the post-test, group differences were demonstrated between the two conditions - the required revision group being significantly more accurate than the group that was not required to revise.

Chandler's (2003) second study presents a re-analysis of data collected for the earlier study. The population and sample of the study were the same. However, in this instance she measured improvements between five journal entries, thus examining the comparative effectiveness of different types of CF over time. The study had a quasi-experimental design (pre-test - treatment - post-test). The tests were the journal entries and the treatment consisted of unfocused direct CF, unfocused located metalinguistic coded CF, unfocused metalinguistic coded CF and unfocused indirect located CF. Students in all four conditions were required to revise. Direct CF was given on incorrectly revised errors. However, the learners were not required to attend to the direct CF in any way. She reported that the direct CF group and indirect CF group improved significantly over time whereas the located metalinguistic coded

CF group and metalinguistic coded CF group did not. There were also group differences. Both the direct and indirect groups were significantly more accurate in the post-tests than both of the metalinguistic ones.

Conducting research on 53 post-intermediate level ESOL learners at a New Zealand university, Bitchener, Young and Cameron (2005) investigated the comparative effectiveness of CF. Their research had a quasi-experimental design involving a pre-test - treatment - post-test (2 weeks later) - treatment - post-test (2 weeks later) - treatment - post-test (4 weeks later) design. For each of the four tests, the participants were required to write an informal letter. The three treatment conditions comprised focused metalinguistic explanation plus a five minute oral conference, focused metalinguistic explanation alone and feedback on content and organisation. None of the groups were required to revise their compositions. The target structures were prepositions, the simple past (regular, irregular and copula) and the definite article. Accuracy was measured using obligatory occasion analysis. The researchers found the focused metalinguistic explanation plus conferencing group significantly outperformed the content and organisation group in the use of the definite article and the simple past. However, it was not reported at which time this occurred. There was no time-group interaction. For the simple past, the metalinguistic explanation plus conference group outperformed the written metalinguistic explanation group, but again the time that this occurred was not reported.

What is at first evident from these studies is that the results are mixed. This can to some extent be explained by referring back to Ellis' typology. A number of studies failed to require

the learners to attend to the CF. In other words, it is not known whether the learners processed the CF or not. Another issue pertains to the failure to provide any delayed post-tests. It has been suggested that in order for learners to restructure their interlanguage following the receipt of CF an incubation period is required; this allows for the permanent storage of items in long-term memory (Gass, 1997). If this is the case, the failure to provide delayed post-tests may fail to reveal underlying patterns of performance. None of the studies can furthermore make any claim as to the effectiveness of CF on acquisition as no control group, completing only the pre-test and post-test, was included. As will be discussed in greater detail in a later section, nearly all of these studies in fact suffered from critical methodological flaws in their design or in the analysis of the data.

However, there are some conclusions that can be drawn from the findings of the research that compared the effectiveness of different types of CF. It would appear that when learners are required to revise, greater improvements over time occur than when the corrected texts are just returned without learners having to attend to the corrections (Chandler, 2003). This is an important finding. For the one study (Chandler, 2003) that consistently had the learners attend to the CF by requiring a revised version of a draft, the results are therefore informative (Chandler, 2003). The unfocused direct CF plus revision and unfocused indirect located CF plus revision groups outperformed the unfocused located metalinguistic coded CF plus revision and unfocused metalinguistic CF plus revision groups.

2.7.2.2 Studies that have examined the effect of written corrective feedback on acquisition

Sheen (2007) undertook research on 111 intermediate-level ESOL students at an American community college. She adopted a pre-test, treatment, immediate post-test and delayed post-test quasi-experimental design. For the tests, picture sequence tasks were used, and for the treatment, CF was provided on two dictogloss tasks. The treatment examined two types of CF: focused direct CF plus metalinguistic explanation and focused direct CF. There was a control group that had no treatment including no writing practice. The two CF conditions were required to solely attend to the CF by studying it, and the control group undertook oral fluency conversation tasks with no CF. The target form that she investigated was the English article system, specifically, the functions of first mention and anaphoric reference. The texts were scored using target-like-use analysis (Pica, 1983). Unlike obligatory occasion analysis, target-like-use analysis takes into consideration the overuse of a particular form. Sheen reported that there were group differences in the immediate post-test between the direct CF plus metalinguistic explanation group and the control group. There were also group differences between the direct CF with metalinguistic explanation group and both the direct CF group and the control group in the delayed post-test.

Ellis et al. (2008) also required their participants to study the CF. Forty-nine intermediate-level EFL students at a Japanese university participated in the study. The study was quasi-experimental with a pre-test - treatment - immediate post-test - delayed post-test design. The tests utilized picture-based writing tasks, and the treatment comprised CF on three stories that the participants were required to write based on a series of pictures reproducing the content of

the original texts. The CF conditions were unfocused direct CF, focused direct CF and a control group receiving no CF. The control group had oral communication classes during the treatment sessions. The linguistic form that was investigated was also the English article system for expressing anaphoric reference and first mention. Following the scoring of the texts through obligatory occasion analysis, the scores were subject to a series of ANOVAs. The researchers found that there were significant differences for the main effects of time and group. For time, the unfocused condition demonstrated significant improvement from the pre-test to the immediate post-test and this was maintained in the delayed post-test whereas the focused CF condition resulted in significant improvement in the accuracy scores between the pre-test and the delayed post-test. In the delayed post-test, both the focused and unfocused direct CF groups outperformed the control group.

Conducting research on 75 low-intermediate level ESOL students at two private language schools in New Zealand, Bitchener (2008) investigated the effectiveness of CF on acquisition. His research was quasi-experimental (pre-test - treatment - immediate post-test - delayed post-test). The tests comprised picture-based tasks, and there were four treatment conditions: focused direct CF, focused direct CF with written metalinguistic explanation, focused direct CF with written and oral metalinguistic explanation and a control group (no CF). The study investigated the use of first mention and anaphoric reference for English articles in writing. Feedback was provided on the writing used for the pre-test only. That is, the initial writing task doubled as the pre-test and the means for eliciting a sample of writing. The participants were required to attend to the CF by studying it. Those participants that received focused

direct CF as part of their treatment also received this feedback on their second writing task but were not required to attend to it. The texts were scored using obligatory occasion analysis. Bitchener found that there were significant effects for time, group and there was a time-group interaction. He reported only the main effect for time but not for the individual groups. There was significant improvement between the pre-test and the immediate post-test and this was maintained in the delayed post-test. The group scores for both the focused direct CF with written and oral metalinguistic explanation and the focused direct CF group alone significantly outperformed the control group whereas the direct CF plus written metalinguistic explanation group did not. There was also a significant time-group interaction. This occurred between the immediate post-test and the delayed post-test and represented the increase in scores of the control group and a slight decrease in the scores of the focused direct CF group and the focused direct CF plus written metalinguistic explanation group.

Bitchener and Knoch (2008) conducted research on 141 low-intermediate level ESOL learners at two private language schools and a university in New Zealand. The design was quasi-experimental with a pre-test - treatment - immediate post-test - delayed post-test design. Three picture-based writing tasks were used for the tests, and there were four treatment conditions. One group received focused direct CF, another focused direct CF plus written metalinguistic explanation, a third direct CF plus written metalinguistic explanation and oral conferences and the final group was a control group receiving no CF. Additionally, the three CF groups received focused CF following the immediate post-test. This comprised ticks above correct forms and crosses above incorrect ones. The participants were not required to

attend to this CF. Obligatory occasion analysis was used to score the texts. The researchers found significant differences for time, group and there was a time-group interaction. The participants in all three CF conditions produced significantly more accurate writing than the control group in the immediate post-test and this effect continued in the delayed post-test. All the experimental groups outperformed the control group. The time-group interaction resulted from an increase in the accuracy of the control group and decreases in the accuracy of the three treatment conditions between the two post-tests.

Xu (2009) critiqued the methodology of two studies investigating the effectiveness of CF on the accuracy of new pieces of writing, these being Ellis et al. (2008) and Bitchener (2008). Xu begins by questioning the sample sizes used in the two studies. The small samples sizes it is claimed could have allowed chance incidents to affect the results. The findings for the pre-test ANOVAs, which found no group differences between the groups in either study, were also questioned. Xu noted the different groups in fact had different pre-test scores and the control groups in both studies had the lowest scores to begin with. As the pre-tests were not comparable, the post-treatment between-group variations in accuracy might not have been attributable to the CF. Xu also raised the issue of whether the increases in accuracy for the two functions of English articles may have resulted in deterioration of the accuracy of other linguistic structures. The design of the studies was also brought under scrutiny. It is claimed that the design of the studies would have made the focus of the studies, two functions of English articles, apparent. This Xu claims would have happened throughout the study but particularly after the provision of CF. It is argued furthermore that the provision of the exit

questionnaire in Ellis et al.'s study before the delayed post-test may have made the focus evident. The exit questionnaire was used to assess what the participants thought the focus of the research was. Xu suggests, furthermore, that its design was biased. Next Xu questions the value of using obligatory occasion analysis to score the learners' writing. Such an analysis fails to account for the overuse of a particular structure. The discussion of the results was also critiqued. For Ellis et al.'s study, Xu questioned how there could be group differences in a delayed post-test but not in an immediate one. Xu suggested that some other uncontrolled variable is accountable for the group differences evident in the delayed post-test, namely, the existing pre-test differences between the groups and the learners becoming aware of the focus of the research in the delayed post-test. With regard to Bitchener's study, Xu first discounted one of his groups as having multiple treatments - one of which he asserted is not a CF treatment. He also argued that the learners were able to figure out the research focus and this, not the treatment, was responsible for the gains evident in the study.

Bitchener (2009) responded to Xu's critique. He first of all addressed Xu's concern as to CF activating pre-existing knowledge about the target structures. Bitchener agreed that indeed this was the case as the pre-test scores demonstrated the learners had only some control over the accurate use of the target structures. In other words, the CF in this study was intended to improve the accuracy of partially acquired forms. Bitchener acknowledged that some types of CF may also assist the acquisition of new knowledge but this was not the purpose of his study. Xu's concern about the sample size was discussed next. Bitchener stated that he mentioned the limitation of the sample size in his study but also noted that the purpose of

inferential statistics was to make claims as to the generalizability of the findings to larger populations if the population were similar in composition. Bitchener additionally rejected Xu's assertions that the findings of his study were overgeneralized. He noted that the findings were relevant to the context in which the study was undertaken and to the structures investigated. Bitchener also reported that the study investigated group scores and suggested that future studies could include a socio-cultural aspect to the design so as to address the performance of individual learners. Bitchener also rejected Xu's claim that the groups' pre-test scores were different. He referred to the fact that one-way between groups ANOVAs demonstrated that there were not a significant difference between the groups in the pre-test. The next of Xu's criticisms to be addressed surrounded the claim that learners became aware of the focus of the study and hence claims as to the effectiveness of CF in the post-tests were invalid. Bitchener rejected this assertion as he noted precautions were undertaken to protect the focus of the study until the first CF session. At this stage the learners should have become aware of the focus of the study. That is, if the CF was effective, the learners should have been able to identify their problems with the target structure and then use this knowledge in the two post-tests. Xu's concern with the focused nature of the CF was next discussed. Bitchener reported that throughout the debate on the efficacy of CF it has been pointed out that different types of structures may respond in different ways to different types of CF. As such, different structures need to be examined with different types of CF and with different focuses. Xu's concern about investigating multiple CF types was then discussed. Bitchener acknowledged that there is a need to investigate individual types of CF; however, he justified the inclusion of his multiple treatment group as it was ecologically valid. In other words, it represented

common practice in the ESOL classroom. Finally, Bitchener discussed the issue of whether focused CF in fact resulted in the reduced accuracy of other non-target structures. Bitchener suggested that the results of his study showed a single episode of CF can lead to improvements in accuracy for selected error categories and these improvements remained in the long-term. It is not known how such CF affects other error categories though. It may actually bring about positive change in other error categories. As a single episode was effective with certain error categories, it may well be with others. Still other error types may require multiple episodes of CF, different types of CF or direct instruction.

Sheen, Wright and Moldawa (2009) undertook research on 80 ESOL students at an American college. The study adopted a quasi-experimental design with a pre-test, treatment, immediate post-test and delayed post-test. The tests involved written narratives and the treatment groups included a focused direct CF group, an unfocused direct CF group, a writing practice group and a control group. The target structures for the focused group were once again two functions of the English article system, namely, first mention and anaphoric reference. For the unfocused group, these two functions for articles, the irregular past tense, the regular past tense, the copula past (be) and both temporal and locative prepositions were investigated. The writing practice group undertook the same two narrative writing tasks on which the CF groups received corrections. Texts were scored using the ratio of correct to incorrect uses. The researchers found that when assessing the accuracy of articles, all four groups significantly improved over time. There were group differences between the focused direct CF group and both the unfocused direct CF group and the control group in the immediate post-test. In the

delayed post-test, the focused direct CF group was more accurate than the control group although this difference only approached significance. The writing practice group was significantly superior in accuracy to the control group. There was also a time-group interaction. For the analyses of the five grammatical structures, all four groups' accuracy improved over time yet there were only group differences in the immediate post-test between the focused direct CF group and the control group as well as the writing practice group and the control group. There was also time-group interaction. It is difficult, however, to draw conclusions from this study for two reasons. First of all, Sheen et al. used a different measure of accuracy than all the other studies investigating the effectiveness of CF on the acquisition of new pieces of writing. Thus, her findings cannot reliably be compared with these other studies. The second reason surrounds the structures investigated for the five grammatical structures. One of them, the irregular past tense, is item-learned rather than system-learned; hence, learners may not have used the same corrected structure in new pieces of writing. They could have used an alternate structure. This may have affected the results.

Sheen (2010) compared the effectiveness of oral and written corrective feedback. 143 ESOL learners were included in the study which was situated in an American community college context. The design was quasi-experimental with a pre-test, a treatment, an immediate post-test, and a delayed post-test (3-4 weeks after the treatment). The treatment groups were drawn from intact classes and included an oral recast group (oral reformulation of an error), an oral metalinguistic CF group (the correct form of an error is provided with a rule for its use), a focused direct CF group and a metalinguistic explanation group. Control groups were

required only to complete the tests and received no CF. For the oral CF groups, learners were given the opportunity to repeat corrections (uptake). During the written CF sessions students were asked to study the CF for five minutes. Sheen used three separate tests which were repeated over three times (pre-test, immediate post-test and delayed post-test): a narrative writing picture composition task (scored with Target-Like-Use Analysis), a speed dictation test (scored with Target-Like-Use Analysis) and an error correction test (discrete item). The three separate test types generated percentages that were then combined and averaged for a total test score. Sheen first aimed to establish whether there were any differences between oral recasts and direct CF. Using one-way ANOVAs, Sheen found that there were significant improvements over time for the control group and subsequently completed two-way ANOVAs and found a time-group interaction and group differences. The time-group interaction can be explained by the deterioration of scores from the immediate post-test to the delayed post-test for the direct CF group and control while the oral recast group continued to improve over this time. With regard to the group differences, it was established the direct CF group outperformed both the oral recast group and the control group at the immediate post-test and delayed post-test while there were no significant differences between the oral recast and the control group. Sheen also looked into whether there were any differences between oral metalinguistic explanation and written metalinguistic explanation. Two-way ANOVAs revealed a time-group interaction and a main effect for group differences. The time-group differences can be explained by the control group's accuracy deteriorating from the immediate post-test to the delayed post-test while the other two CF group's accuracy (both oral and written) continued to improve in accuracy. For both post-tests, group differences

were evident. Both metalinguistic explanation groups (oral and written) outperformed and the control group and there were no significant differences between the two corrective feedback groups.

Undertaking research on 54 ESOL students in a New Zealand university context, Bitchener and Knoch (2010a) investigated the longitudinal effects of three types of CF and a control group (no CF). One CF group received focused direct CF plus written and oral metalinguistic explanation, another focused direct CF plus written explanation and the last CF group, focused direct CF alone. It is perhaps important to note direct CF was operationalized in a slightly different manner in this study. Bitchener and Knoch not only provided the corrections for errors of the target structure but also provided a tick (√) for correct uses of the target structure. Two functions of the English article system were investigated and these were first mention and anaphoric reference. The texts were scored using obligatory occasion analysis. The study used a quasi-experimental design. This involved the participants first completing a pre-test then the CF treatments were provided on the texts produced for these tests. Following the treatments, the learners completed an immediate post-test (1 week after the pre-test) and three delayed post-tests: one eight weeks after the pre-test; one six months after the pre-test; the final post-test 10 months after the pre-test. The tests were picture based tasks and these were scored using obligatory occasion analysis. The authors found that there was a significant main effect for the both time and group as well as there being a time-group interaction. One way between groups ANOVAs with post-hoc Tukey pairwise comparisons were applied to all the post-tests, and these found that for all four post-tests the three CF groups outperformed the

control group. One can largely explain the time-group interaction as a result of the accuracy of the CF groups improving over time while that of the control group tended to atrophy.

Bitchener and Knoch (2010b) conducted research on 63 ESOL learners in an American University. The design of the study was quasi-experimental with a pre-test - treatment - immediate post-test - delayed post-test (10 weeks following the pre-test). There were three treatment groups: a focused metalinguistic explanation group, a focused metalinguistic explanation group plus oral instruction group, and a focused located indirect CF group. A control group only completed the tests and received no CF. The tests comprised picture description tasks which were scored using obligatory occasion analysis. The structure investigated was the English article system and in particular the functions of first mention and anaphoric reference. The researchers found that there was no significant time-group interaction. There were, however, significant effects for both time and group. One-way between group ANOVAs demonstrated there were group differences in both the immediate post-test and the delayed post-test. Post-hoc Tukey pairwise comparisons were conducted and for the immediate post-test it was found that all three CF groups outperformed the control group while in the delayed post-test only the two metalinguistic CF groups did.

Van Beuningen et al. (2012) undertook a study with 268 secondary school and prevocational school students in a Dutch as a second language context. Adopting a quasi-experimental design, the participants undertook a pre-test, a post-test (one week after the treatment) and a delayed post-test (four weeks after the treatment). Intact classes were assigned to one of four

groups - an unfocused direct CF group, an unfocused located metalinguistic coded CF (nine error categories), a self-correction group (control 1) and a writing practice group (control 2). The writing tasks involved outlining the process of metamorphosis for butterflies (pre-test), ladybugs (post-test) and wasps (delayed post-test). Learners were required to attend to the CF by incorporating corrections into a revised version of a text. The tasks were scored for accuracy by using an error ratio where the number of linguistic errors were divided by the total number of words and multiplied by 10. They were also scored for structural complexity and lexical diversity. The researchers used ANCOVAs. They found that neither structural complexity nor lexical diversity were affected by the CF. The researchers established that the accuracy of the two treatment conditions outperformed that of two control groups in a revised version of a text and this was also evident in both the short-term and the long-term. They additionally conducted an analysis of the accuracy of grammatical structures and non-grammatical structures.

> The first measure [grammatical errors] was a ratio calculated on the basis of the sum of the number of article errors, inflectional errors, word order errors, omissions of necessary elements, additions of non-necessary elements, pronominal errors, and other grammatical errors [errors that did not fit into existing error categories]. Lexical errors, orthographical errors, appropriateness/pragmatic errors, and other non-grammatical errors [errors that did not fit into existing error categories], on the other hand, were included in the non-grammatical accuracy ratio. (pp. 17-18)

This additional study presented mixed results. With regard to non-grammatical structures in the short-term, both the unfocused direct CF group and the unfocused located metalinguistic coded CF group outperformed the self-correction group and the writing practice group. In relation to grammatical accuracy in the short-term, the unfocused direct CF group alone produced a more accurate piece of writing compared to the writing practice group. In the long-term for the non-grammatical structures, the unfocused located metalinguistic coded CF outperformed the self-editing group and the writing practice group; that is, the group effects demonstrated in the short-term were maintained in the long-term. This was not the case for the unfocused direct CF group which failed to demonstrate any significant group differences with non-grammatical structures in the delayed post-test. The unfocused direct CF group did however, maintain the effects demonstrated in the short-term to the long-term for grammatical structures. Namely, there were group differences between an unfocused direct CF group and a writing practice group at both post-tests.

Shintani and Ellis (mimeograph) completed a study with 49 low intermediate-level ESOL learners in an American University. The quasi-experimental design included a pre-test, treatment, immediate post-test and a delayed post-test (two weeks after the treatment). Five classes were randomly assigned to three groups: a focused direct CF group, a metalinguistic explanation group and a control group. Feedback was given on the initial writing task. The CF groups were required to attend to the feedback by first studying the corrections or the metalinguistic explanation and then revising their text. The control group was given five minutes to study their initial piece of writing during the treatment session in preparation for

completing the revised version. The tests comprised new pieces of writing in the form of picture composition tasks, the target structure was a single function of the indefinite article (specific referent unknown to the hearer), and the texts were scored using obligatory occasion analysis. Between the pre-test and the revised version of the pre-test, two-way ANOVAs demonstrated significant main effects for time, group and time-group interaction. It was demonstrated that only the metalinguistic explanation group improved over time, and was significantly more accurate than the control group. It also tended to outperform the direct CF group as the difference in this instance approached significance. The time-group interaction is explicable by the fact that while both groups improved in accuracy, the control group remained the same. For the effect of the feedback plus revision on new pieces of writing, it was found that there was a significant effect for time and time-group interaction. Pairwise comparisons found the metalinguistic group improved in accuracy in the short-term alone and its accuracy, like that of the focused direct CF group, in fact, atrophied from the short- to the long-term. The time-group interaction can largely be explained by the improvement in the accuracy of both experimental groups in the short-term while the control group's accuracy deteriorated and as a result of the experimental groups deteriorating in accuracy from the short-term to the long-term while the control group's accuracy improved. Shintani and Ellis also conducted eye tracking and stimulated recall to ascertain how learners responded to the two treatments and what they focused on during a revision of an initial test. For this, they recruited six additional students at the same level (three for focused direct CF and three for metalinguistic explanation). They demonstrated that while a focused direct CF group did fixate on the corrections, they were unable to form an explicit rule for the target structure

whereas the metalinguitsic group were able to form such a rule.

Later studies investigating the effectiveness of CF on acquisition have largely addressed the problems of the earlier studies comparing the relative effectiveness of CF. All of the studies required the learners to attend to the CF by studying the CF or revising the CF, they all had control groups undertaking only the tests and they all provided delayed post-tests. The following represent the main findings of the research investigating the effectiveness of CF on acquisition.

1. There are several findings related to changes over time.
    a. Several studies reported a main effect for time (Ellis et al., 2008; Bitchener, 2008; Bitchener & Knoch, 2008; Sheen et al., 2009, Bitchener & Knoch, 2010a; Bitchener & Knoch, 2010b; Shintani & Ellis, mimeograph).
    b. Two studies reported changes over time for individual groups. Ellis et al. (2008) reported both a focused direct CF group and an unfocused direct CF group improved over time while a control group did not. Sheen et al. (2009) showed that a focused direct CF group, unfocused direct CF group, writing practice group and a control group all improved over time. There are then mixed results regarding improvements over time for control groups.
    c. Two studies reported when the improvements over time occurred (i.e. between the pre-test, immediate post-test and delayed post-test). Ellis et al. (2008) demonstrated that both a focused direct CF group and an unfocused direct CF group improved in the short- and long-term. Only the focused direct CF group

showed a tendency of continued improvement from the short-term to the long-term whereas the unfocused direct CF group's accuracy deteriorated. Shintani and Ellis (mimeograph) showed a metalinguistic explanation group improved in the short-term; however, this was not maintained in the long-term.

2. There were a number of findings related to group differences.
    a. Several studies reported main group effects (Ellis et al., 2008; Bitchener, 2008; Bitchener & Knoch, 2008; Sheen et al., 2009, Sheen, 2010; Bitchener & Knoch, 2010a; Bitchener & Knoch, 2010b).
    b. Group differences were reported in the short-term (i.e. immediate post-test).
        i. Focused direct CF plus metalinguistic explanation outperformed a control group (Sheen, 2007; Bitchener & Knoch, 2010a).
        ii. Focused metalinguistic explanation was more accurate than a control group (Sheen, 2010; Bitchener & Knoch, 2010b).
        iii. Unfocused located metalinguistic coded CF outperformed a control group (Van Beuningen et al., 2012).
        iv. Unfocused located metalinguistic coded CF outperformed a self-editing group (Van Beuningen et al., 2012).
        v. Unfocused located metalinguistic coded CF for non-grammatical structures performed better than a self-editing group (Van Beuningen et al., 2012).
        vi. Unfocused located metalinguistic coded CF group examining non-grammatical structures produced a more accurate piece of writing than

a control group (Van Beuningen et al., 2012).

vii. Focused direct CF performed better than a control group (Sheen, 2007; Bitchener & Knoch, 2008; Sheen et al., 2009; Sheen, 2010**;** Bitchener & Knoch, 2010a), and in one study focused direct CF led to superior accuracy compared to unfocused direct CF (Sheen et al., 2009).

viii. Unfocused direct CF outperformed a control group (Van Beuningen et al., 2012).

ix. Unfocused direct CF outperformed a self-editing group (Van Beuningen et al., 2012).

x. Unfocused direct CF for grammatical structures produced a more accurate piece of writing than a control group (Van Beuningen et al., 2012).

xi. Focused indirect located CF outperformed a control group (Bitchener & Knoch, 2010b).

xii. A writing practice group was shown to be more accurate than a control group (Sheen et al., 2009).

c. Group differences were also reported in the long-term (i.e. delayed post-test or delayed post-tests).

i. Focused direct CF plus metalinguistic explanation outperformed both a focused direct CF group and a control group (Sheen, 2007).

ii. Focused metalinguistic explanation lead to superior accuracy compared to a control group (Sheen, 2010; Bitchener & Knoch 2010b).

iii. Unfocused located metalinguistic coded CF outperformed a control group (Van Beuningen et al., 2012).

iv. Unfocused located metalinguistic coded CF outperformed a self-editing group (Van Beuningen et al., 2012).

v. Unfocused located metalinguitsic coded CF investigating non-grammatical structures produced a more accurate piece of writing then a self-editing group (Van Beuningen et al., 2012).

vi. Unfocused located metalinguistic CF group, when examining non-grammatical structures, outperformed a control group (Van Beuningen et al., 2012).

vii. Focused direct CF performed better than a control group (Ellis, et al., 2008, Bitchener & Knoch, 2008; Sheen et al., 2009; Sheen, 2010; Bitchener & Knoch, 2010a)

viii. Unfocused direct CF demonstrated superior accuracy compared to a control group (Ellis et al., 2008;Van Beuningen et al., 2012).

ix. Unfocused direct CF outperformed a self-editing group (Van Beuningen et al., 2012).

x. Unfocused direct CF group investigating grammatical structures produced more accurate writing than a control group (Van Beuningen et al., 2012).

xi. Writing practice was shown to be more accurate than a control group (Sheen et al., 2009).

3. Some studies reported time-group interactions.
    a. Bitchener (2008) demonstrated that the accuracy of a focused direct CF plus metalinguistic explanation group and a focused direct CF group deteriorated between an immediate post-test and a delayed post-test while the scores of control groups increased across the same two tests.
    b. Bitchener and Knoch (2008) also found that a focused direct CF group plus metalinguistic explanation and a focused direct CF group deteriorated in accuracy from an immediate post-test to a delayed post-test while the accuracy of a control group improved over the same period.
    c. Sheen et al. (2009) conducted an analysis of two sets of structures. For the two functions of English articles (first mention and anaphoric reference), they found that between an immediate post-test and a delayed post-test a focused direct CF group, an unfocused direct CF group and a writing practice group's accuracy improved whereas the accuracy of a control group atrophied. Their analysis of five different structures (the two aforementioned functions for English articles, the irregular past tense, the regular past tense, the copula past and both temporal and locative prepositions) produced different results. Between an immediate post-test and a delayed post-test, a focused direct CF group and a writing practice group's accuracy deteriorated while a control group and an unfocused direct CF group improved in accuracy.
    d. Sheen (2010) found two time-group interactions. The first can be explained by her direct CF group deteriorating between an immediate post-test and a

delayed post-test while a control group improved in accuracy across the same time period. The second is explicable by the metalinguistic explanation group improving in accuracy from an immediate post-test while a control group's accuracy atrophied.

e. Bitchener and Knoch (2010a) found that between an immediate post-test and the final of three delayed post-tests a focused direct CF plus metalinguistic explanation group and a focused direct CF group improved in accuracy whereas the accuracy of a control group did not.

f. Shintani and Ellis (mimegraph) showed that a group that received metalinguistic explanation plus revision and focused direct CF plus revision led to improved accuracy in the short-term while the control group's accuracy deteriorated. However, from the short-term to the long-term, the control group's accuracy improved whereas the two CF groups' accuracy atrophied.

2.7.3 Methodological limitations

There are, as with most research, methodological limitations with these studies - some of which limit the generalizability of the findings. One of these is the confounding of independent variables. This involves the inclusion of two or more variables that operate as a single independent variable yet the conclusions of the studies may refer to only one of these variables as being responsible for the results. Some studies have confounded their independent variables as a result of poor design. In Lalande (1984), for the unfocused located metalinguistic CF condition it is unclear whether it was the CF, access to grammar texts, peer

feedback, oral teacher feedback, written explanations of lexical items or the use of error logs that accounted for the group differences. Similarly, Polio (1998) et al.'s unfocused direct CF group also had editing practice. Both of the studies reported in Chandler (2003) confounded variables. In the first study, the variables were an unfocused indirect located CF plus revision group and an unfocused indirect located CF group that was not required to attend to the CF in any way. In the second study, they comprised an unfocused indirect located CF plus revision group, an unfocused located metalinguistic coded CF plus revision group and an unfocused metalinguistic coded CF plus revision one. Following the receipt and revision of the CF, the learners were given direct CF on any existing errors; however, they were not required to attend to the CF. They may or may not have attended to the CF. It is not known then to what extent it was the revised unfocused CF or returned unfocused direct CF that influenced the new piece of writing. Similarly, Bitchener (2008) provided focused direct CF on all the targeted errors for the two post-tests. While the participants were not required to attend to this CF, they may or may not have attended to it. Thus, for the direct CF plus metalinguistic explanation group, it is not known whether it was the initial direct CF plus metalinguistic explanation or the subsequent direct CF on the immediate post-test that was responsible for the results in the delayed post-test. Bitchener and Knoch (2008) provided feedback on the post-tests in the form of a tick (√) for correct uses of target structures and a cross (×) for incorrect uses. The learners were not required to attend to this additional feedback but they may well have. It is not known then whether this extra feedback influenced performance of the focused direct CF plus metalinguistic explanation group or the focused direct CF group. Bitchener and Knoch (2010a) also confounded variables. Here the authors combined direct

CF with ticks (√) above correct forms. It is not known the extent to which it was the direct CF or the additional feedback that was responsible for the results for the focused direct CF plus metalinguistic explanation group or the focused direct CF group.

Methodological issues also exist regarding the methods used to score texts. In general, the comparability of studies is brought into question by the use of difference measures for calculating accuracy scores. Studies investigating the effectiveness of CF on a revised version of a text used the number of errors per text (Fathman & Whalley, 1990), the number of successfully corrected errors (Lee, 1997), the percentage of successfully corrected errors (Ferris & Roberts, 2001), the mean error rate per 100 words (Chandler, 2003) and error rate as the total number of errors divided by the total number of words (Truscott & Hsu, 2008). The research comparing the relative effectiveness of different types of CF used percent of errors (Lalande, 1982) the ratio of errors to number of words (Semke, 1984), the mean number of errors per 100 words (Chandler, 2003), obligatory occasion analysis (Bitchener, et al., 2005) and error free t-units (Polio et al., 1998). The measure used by Polio et al., is often used for measuring complexity rather than accuracy. One could argue that error free t-units is quite a coarse measure of accuracy that conceals overall accuracy within a t-unit. Those studies investigating the effectiveness of CF on acquisition have tended to use obligatory occasion analysis (Ellis et al., 2008; Bitchener, 2008; Bitchener & Knoch, 2008; Bitchener & Knoch, 2010a; Bitchener & Knoch, 2010b). One used target-like use analysis (Sheen, 2007) and yet another has used the ratio of correct to incorrect uses (Sheen, 2007). A single study by Van Beuningen et al. (2012), however, used an error ratio to measure accuracy. This was the same

as those used in Chandler (2003) and Truscott and Hsu (2008). It comprised counting the number of linguistic errors divided by the total number of errors and this figure was then multiplied by 10.

There are concerns surrounding sample size. A number of studies had group sizes smaller than 30, did not report whether the groups met the assumption of normality and used parametric tests (Fathman & Whalley, 1990; Ferris & Roberts, 2001; Chandler, 2003; Bitchener et al. 2005; Truscott & Hsu, 2008; Ellis et al. 2008; Sheen et al, 2009; Bitchener & Knoch, 2010a; Bitchener & Knoch, 2010b; Shintani & Ellis, mimeograph). Pallant (2007) points out that it is generally accepted that group sizes greater than 30 are large enough so parametric tests can be undertaken. In other words, the assumption of normality is not an issue. For groups smaller than 30, tests for normality should be undertaken. If the groups are not all of a normal distribution, non-parametric tests should be employed.

The next issue involves the statistical tests employed by researchers. Two studies used inappropriate types of tests for pre-tests or post-tests. Lalande (1982) used a paired samples t-test to test for group differences in a pre-test when he should have used an independent samples t-test. Truscott and Hsu (2008) reported using a Wilcoxon Signed Rank Test to test for group differences in a pre-test when they should have used a Mann-Whitney U Test. This may have affected the results.

There are problems involving issues of scoring reliability. Two studies failed to report inter-

rater or intra-rater reliability scores (Fathman & Whalley, 1990; Lee, 1997). Such measures of reliability are used to ensure that the rater or raters score consistently. Inconsistent scoring or incorrect scoring would affect results.

Another issue surrounds the failure to determine if there were group differences in a pre-test (Chandler, 2003; Bitchener et al., 2005; Sheen, 2007). With these studies, it is unknown whether initial group differences affected between-group differences in post-tests

Some studies failed to counterbalance the tasks (Lalande, 1982; Polio et al., 1998; Bitchener, et al.; 2005; Bitchener, 2008; Bitchener & Knoch, 2008; Bitchener & Knoch, 2010a; Bitchener & Knoch, 2010b; Van Beuningen et al., 2012). Counterbalancing is a technique by which different writing tasks are equally divided between students at each testing time yet each participant completes a different task across the different testing times. Counterbalancing is used to take account of varying degrees of task difficulty.

Sheen's (2010) study used three separate tests repeated three times across a pre-test and two post-tests. It is possible that such a procedure might lead to a testing effect as the participants did not complete different tests. The scores for the speeded dictation test, the narrative writing test and the error correction test were combined to provide an averaged total test score. In this instance, it is not clear what effect the CF types had on the learners' accuracy in new pieces of writing. Similarly, Van Beuningen et al. (2012) had the learners write the same task three times, which also may have led to a testing effect. As the researchers examined unfocused

feedback as a whole as well as on the basis of non-grammatical and grammatical structures, it is possible that some testing effect occurred. This may account for the researchers choosing ANCOVAs which report only group differences rather than accounting for any changes over time which one-way within groups ANOVAs with post-hoc pairwise comparisons or two-way ANOVAs with post-hoc within groups pairwise comparisons may have demonstrated.

One study had concerns regarding the pre-existing knowledge of the target structures for one of the groups in their study (Bitchener & Knoch, 2010b). One of the groups had a pre-test score of over 90. This value is commonly recognized as the value deemed to represent the acquisition of a particular structure (Brown, 1973). Having a group start with such a value also means that the participants have less room to improve over time and hence between groups.

Concern also exists about the choice of tests. One study used timed writing tasks of 10 minutes for pre-tests and post-tests (Semke, 1984). This may have influenced the results. It is possible, for example, that time pressure may have affected the learners' ability to draw on their existing metalinguistic knowledge during the writing process.

The final area of concern surrounds contextual issues. The majority of studies were undertaken in L2 contexts (Polio et al., 1998; Chandler, 2003; Bitchener, et al., 2005; Sheen, 2007; Bitchener, 2008; Bitchener & Knoch, 2008; Sheen et al., 2009; Sheen, 2010**;** Bitchener & Knoch, 2010a; Bitchener & Knoch, 2010b; Van Beuningen et al., 2012; Shintani & Ellis,

mimeograph). L2 contexts provide access to additional input in other classes and via the larger L2 context. It is not known then to what extent it is the independent variables or the L2 context that are responsible for improvements evident in these studies. In fact few studies investigating the effect of CF on acquisition have been conducted in a foreign language environment (Ellis et al., 2008; Truscott & Hsu, 2008).

2.8 State of current knowledge about written corrective feedback

It is evident that requiring learners to incorporate CF into a revised version of a piece of writing results in more accurate drafts compared to not providing any CF at all. The provision of direct CF would appear to lead to more accurate revisions than learners receiving indirect CF and metalinguistic CF which in turn produce more accurate drafts than learners who received no CF (control groups). However, demonstrating that CF is effective in a revised version of a text does not mean that such CF will lead to the acquisition of structures as demonstrated in new pieces of writing.

Little can reliably be concluded from most of the studies that investigated the relative effectiveness of CF. This is due to the various methodological limitations of those studies. However, there are some findings. When learners were required to revise their CF, they produced a more accurate new piece of writing compared to learners not required to revise (i.e. the CF was just returned to the learners). Indeed, when learners were consistently required to incorporate CF into a revised draft, it appears that the provision of unfocused direct CF and unfocused indirect located CF led to superior accuracy in a new piece of writing

compared to learners receiving unfocused located metalinguistic coded CF and unfocused metalinguistic coded CF. However, the unfocused located indirect CF group represented a confounded independent variable as the learners received additional direct CF on those errors they incorrectly revised. Despite the learners not being required to attend to this additional CF, it is not known the extent to which the additional direct CF affected the results for this group.

Studies investigating the effectiveness of CF on acquisition have, on the other hand, tended to have more robust methodological designs and procedures. All of these studies required the learners to attend to the CF, had control groups and provided delayed post-tests. There now exists research that has begun to investigate Ferris' big question regarding the efficacy of CF. It would appear that by and large CF is effective in improving learners written accuracy in the short or long-term whereas this is largely not the case for writing practice alone. In the short-term (immediate post-test), the following types of CF outperformed the control groups: focused direct CF plus metalinguistic explanation, focused metalinguistic explanation, metalinguistic explanation following writing plus revision, unfocused located metalinguistic coded CF plus revision**,** focused direct CF, unfocused direct CF, focused indirect located CF and writing practice (5 episodes of writing practice). In the long-term (delayed post-test or delayed post-tests), several types of CF also outperformed control groups: focused direct CF plus metalinguistic explanation, focused metalinguistic explanation, unfocused located metalinguistic coded CF, focused direct CF, unfocused direct CF and a writing practice group (5 episodes of writing practice). There were also mixed results in the short-term and long-

term. One study reported a focused direct CF group did not outperformed an unfocused direct CF group in an immediate post-test (Ellis et al., 2008) while another reported that it did (Sheen et al., 2009). Another study reported significant differences between a focused direct CF group plus metalinguistic explanation and a direct CF group in the long-term (Sheen, 2007) while others did not report such a finding (Bitchener, 2008; Bitchener & Knoch, 2008; Bitchener & Knoch, 2010a).

With regard to time-group interactions, few conclusions can reliably be drawn at this stage as the results are largely mixed or only a single study has demonstrated such an effect for certain types of CF. Shintani and Ellis' study demonstrated an interaction in the short-term as a metalinguistic explanation plus revision group and a focused direct CF plus revision group improved in accuracy while a control group deteriorated. All of the studies reporting interactions demonstrated patterns between an immediate post-test and a delayed post-test. Two studies reported a deterioration of the accuracy of both direct CF plus metalinguistic explanation groups and direct CF groups (Bitchener, 2008, Bitchener & Knoch, 2008) while the accuracy of control groups improved; however, two other studies reported the opposite pattern (Sheen, 2010; Bitchener & Knock, 2010) and still another for both focused direct CF and unfocused direct CF (Sheen et al., 2009). Sheen et al. (2009) investigated two functions of the English article system (i.e. first mention and anaphoric reference), and five grammatical structures (i.e. the two functions of the English article system, the irregular past tense, the regular past tense, the copula past (be) and both temporal and locative prepositions). In both instances there were time-group interactions. For the English article system, a focused direct

CF group, an unfocused direct CF group and a writing practice group's accuracy improved whereas the accuracy of a control group atrophied; on the other hand, for the five grammatical structures, a focused direct CF group and a writing practice group's accuracy deteriorated while a control group and an unfocused direct CF group improved in accuracy. One could argue that there are differences between these analyses because some structures are system-learned (treatable) such as the English article system while others are item-learned (untreatable), for instance, the irregular past tense. One other study found that a direct CF group's accuracy atrophied over the same time period while control group's accuracy improved (Sheen, 2010).

2.9 Gaps in the literature

What is immediately evident is that CF has been shown to be effective in a revised version of a text. However, few studies investigating the effectiveness of CF on acquisition have also required learners to revise the CF. The majority of studies required participants to attend to the CF by studying it for a period of time. There is clearly a need to investigate whether the improvements evident in revised versions of a text transfer to the accuracy of new pieces of writing in studies that have a control group.

There are furthermore still some types of CF that need to be investigated in relation to the big question, that is, for changes over time. No study has looked into unfocused direct CF plus metalinguistic explanation, focused indirect indicated CF, unfocused indirect indicated CF, unfocused indirect located CF, focused metalinguistic coded CF, unfocused metalinguistic

coded CF, unfocused located metalinguistic coded CF, focused located metalinguistic coded CF and unfocused metalinguistic explanation.

Bitchener and Knoch (2010b) looked into the effectiveness of focused indirect located CF and focused metalinguistic explanation. As this is the only study that has investigated these types of CF, arguably further research is required. No study has compared the effectiveness of direct CF and indirect CF in the accuracy of new pieces of writing.

Those studies that have provided mixed results are also worthy of further examination. Ellis et al. (2008) and Sheen et al. (2009) provided conflicting results in their comparisons of focused and unfocused direct CF. Sheen et al.'s (2009) study provided conflicting results for the same types of CF yet different structures. Indeed, this study alone has looked into structures other than the English article system. In light of the findings of Sheen et al.'s study, further research would appear to be needed on the effect of CF on treatable and untreatable grammatical features.

Of these gaps in the literature, this study is going to investigate the effectiveness of CF based on the degree of focus and directness. It will attempt to build on Ellis et al.'s (2008) study by examining the effectiveness of a focused direct CF group, unfocused direct CF group, focused indirect located CF group (henceforth referred to as focused indirect CF), unfocused indirect located CF group (henceforth referred to as unfocused indirect CF) and a writing practice group (control) on both revision and in new pieces of writing. In other words, the study will

investigate CF both with and without opportunities to revise. As most studies have investigated the English article system, this study will use the past tense and will look at changes over time and between groups.

CHAPTER THREE

PILOT STUDY

3.1 Overview

A pilot study was conducted at a University in Taipei, Taiwan. Its purpose was to test the validity of the study's design, instruments and procedures. It investigated the effectiveness of different kinds of CF plus students' response to the feedback as measured in terms of the learners' accuracy in a revised text and in new pieces of writing. Following the collection, processing and analysis of the data, the findings and limitations of the study are presented. These served to inform decisions about how to improve the design of the main study.

3.2 Research questions

The pilot study investigated three research questions.

RQ1. What are the relative effects of focused and unfocused direct CF plus students' response (i.e. just attending to errors) and focused and unfocused indirect CF plus students' response (i.e. students self-correct their errors indicated on their text) on learners' accurate use of past tense forms in a new piece of writing?

RQ2. What are the relative effects of focused and unfocused direct CF plus students' response (i.e. just attending to errors) and focused and unfocused indirect CF plus students' response (i.e. students self-correct their errors indicated on their text) on learners' accurate

use of past tense forms in a revised version of their original text?

RQ3.What are the relative effects of focused and unfocused direct CF plus students' response (i.e. just attending to errors) plus revision and focused and unfocused indirect CF plus students' response (i.e. students self-correct their errors indicated on their text) plus revision on learners' accurate use of past tense forms in a new piece of writing?

3.3 Target structures

The linguistic structure that was investigated was the simple past tense. The selection of this form was based on feedback from EFL teachers currently working at the University. They acknowledged that mastery of this form was varied with some students having acquired it but others still in the process of doing so. It was also chosen because the target structure occurs frequently in the selected written genre used in this study (i.e. narrative).

The simple past tense is not functionally complex although it does have numerous forms. The past tense is used to refer to a completed action or state in the past. This function can be expressed through the use of the past tense copula (*was* or *were*), regular verbs (e.g. *walked* and *talked*) and irregular verbs (e.g. *went* and *did*). It can be used in the active voice or the passive one, but it was decided to focus exclusively on active forms. This decision was made on the grounds that errors in the passive voice are potentially not treatable through corrective feedback (Ferris & Roberts, 2001). Therefore, the forms of the simple past examined in this study were the past tense copula verbs, irregular verbs and regular verbs.

These were all in the active voice.

3.4 Population and sample

The population for the pilot was university-level Chinese learners of English in Taiwan. In this context, classes are organized on the basis of the major of the students, and the students stay in these classes for all the subjects they are required to take and for the duration of their undergraduate study. The sample was drawn from one sophomore class majoring in computer science and industrial engineering, a sophomore class of information telecommunication engineering students and a junior-year class of business management majors. There was a range of levels of proficiency in the classes. Nevertheless, according to the teacher of these classes, the vast majority of the learners could be classified as intermediate in level.

English is a compulsory subject for all undergraduate students at the university. Students must have four years of English instruction, and the syllabus is built upon eight-levels of in-house texts with varying degrees of emphasis on vocabulary and grammar development and of varying degrees of focus on the skills of listening, speaking, writing and reading.

The ages of the participants were between 19 and 21 years of age, and they had, on average, been studying English for around 10 years. Of the sample, 65.00% were male and 35.00% female.

The number of students that undertook the initial writing task was 178; however, after the completion of the study, only 151 were included in the research. There are some reasons for this reduction in numbers. First of all, due to the range of levels, some students failed to provide at least two obligatory occasions for the past tense in all the written samples. If this was the case, they were removed from the study. There were also students who missed one of the data collection episodes or arrived to class late. Despite their participation in most of the study, their data had also to be removed.

For each of the three intact classes, the participants were randomly and equally assigned to one of five groups following the completion of the first writing task. These groups were focused direct CF plus response (29 students), unfocused direct CF plus response (30 students), focused indirect CF plus response (30 students), unfocused indirect CF plus response (29 students) and a Control Group (33 students).

3.5 Design

Four experimental groups were included in the design of the study, and one control group.

1. Focused direct CF + response (attending to errors)

walked	ate
I ~~walk~~ hame yesterday and ~~eat~~ a sandwiches.	

Focused direct CF involved crossing out only the incorrect past forms and providing the correct forms. Missing ones were indicated with a cursor (^) and then inserted. The students were required to study the feedback for five minutes upon the CF being returned to them in the CF response sessions. In these sessions, the teachers provided no additional comments.

2. Unfocused direct CF + response (attending to errors)

walked home	ate sandwich
I ~~walk~~ ~~hame~~ yesterday and ~~eat~~ a ~~sandwiches.~~	

This entailed providing the correct forms for all the linguistic errors by crossing out the errors and writing the corrections above the errors. Missing forms were indicated and inserted. The students were required to study the feedback for five minutes upon the CF being returned in the CF response sessions. At no time did the teacher give any additional comments during these sessions.

3. Focused indirect CF + response (self-correcting errors)

I walk_ hame yesterday and <u>eat</u> a sandwiches.

This consisted of underlining errors associated with the target form only or in the

case of missing words the use of a cursor and a line (^). The participants were asked to self-correct in the CF response sessions. During these sessions, the teachers gave no additional comments.

4. Unfocused indirect CF + response (self-correcting errors)

I walk__ hame yesterday and eat a sandwiches.

Unfocused indirect CF consisted of underlining all the errors in a text, or in the case of missing words, the use of a cursor and a line. The participants were asked to self-correct their errors in the CF response sessions. No comments were made by the teacher during the response sessions.

5. The Control Group

The Control Group received no CF and was required to undertake conversation tasks when the other groups were engaged in the CF response sessions or when they were undertaking the revised version of Task 2. The tasks were designed to enhance fluency. These tasks simply comprised a series of questions that the participants asked each other in pairs. The questions provided an opinion gap. None of the questions were in the past tense, and the teachers provided no oral feedback.

1. Are there too many people in the world?
2. Compare your two favorite cities?

Table 6 shows the design of the study. For step 1, all groups completed writing task 1 (Task 1). This served as a means of establishing the learners' initial level of accuracy in the use of past tense verb forms. In Step 2 the learners' written texts with their respective types of CF written on them were returned. The Control Group received no CF, so the learners did not have their texts returned. In Step 3, both direct CF groups were required to attend to the CF by studying it for five minutes. In step 4 the indirect CF groups were asked to self-correct the errors indicated. In the next step, step 5, all five groups completed writing task 2 (Task 2). In step 6, the texts produced in Task 2 were corrected in the same way as for Task 1. Again the Control Group received no CF, so their texts were not returned to them. Step 7 required the learners in the direct CF groups to attend to the CF by studying it for five minutes. The learners in the indirect CF groups were asked to self-correct the CF in step 8. All four CF groups undertook a revised version of writing task 2 in step 9. The learners did not have access to their CF when they completed their revised version but were given copies of their original piece of writing without CF. All five groups completed writing task 3 (Task 3) in step 10. In the pilot study, the writing tasks were not counterbalanced (i.e. a different task was used on each occasion).

Table 6: The Design of the Pilot Study

Steps	Focused Direct	Unfocused Direct	Focused Indirect	Unfocused Indirect	Control
Step 1 - writing task 1	Yes	Yes	Yes	Yes	Yes
Step 2 - corrective feedback	Yes	Yes	Yes	Yes	No
Step 3 - CF response of attending to CF	Yes	Yes	No	No	No
Step 4 - CF response of self-correction of CF	No	No	Yes	Yes	No
Step 5 - writing task 2	Yes	Yes	Yes	Yes	Yes
Step 6 - corrective feedback	Yes	Yes	Yes	Yes	No
Step 7 - CF response of attending to CF	Yes	Yes	No	No	No
Step 8 - CF response of self-correction of CF	No	No	Yes	Yes	No
Step 9 - revision of task 2	Yes	Yes	Yes	Yes	No
Step 10 - writing task 3	Yes	Yes	Yes	Yes	Yes

Table 7 shows how the design was used to answer the research questions:

- Research question 1 examined the effect of CF plus response on a new piece of writing. Task 1 operated as a pre-test and Task 2 as an immediate post-test. The experimental groups received CF which they had to respond to after Task 1. The Control Group received no CF and undertook oral fluency tasks between the two tasks.
- Research question 2 investigated the effect of CF plus response on a revised version of a text. Task 2, in this instance, was used as a pre-test and the revised version of Task 2 as an immediate post-test. The learners did not have access to their CF when

completing the revised version. After Task 2, the experimental groups received CF that they were required to respond to. The Control Group again received no CF and was required to complete oral fluency tasks in the time between Task 2 and a revised version of Task 2.

- Research question 3 looked into the effect of CF with response plus revision on a new piece of writing. Task 2 functioned as a pre-test and Task 3 as an immediate post-test. The experimental groups received CF and were required to respond to the CF. The participants were then asked to complete a revised version of Task 2. They did not have access to the CF when completing the revised version. The Control Group once again received no CF and undertook oral fluency tasks in the time between Task 2 and Task 3.

Table 7: The Design, Research Questions and Variables

	Task 1	Corrective feedback plus response 1	Task 2	Corrective feedback plus response 2	Revision Task	Task 3
Research question 1	√	√	√			
Research question 2			√	√	√	
Research question 3			√	√	√	√

Key: √ The dependent variables are the writing tasks (tests) and revision task.

√ The independent variables are the CF with response and CF with response plus revision.

3.6 Instruments

There were two types of instruments used to collect the data: a questionnaire and three writing tasks. The questionnaire was designed to collect the participants' background data including age, gender, field of study and years of studying English.

The writing tasks doubled as both the tests and the means of eliciting a writing sample to which CF was provided. Each task included instructions in English and the following materials within a task package: a text; pictures corresponding to the content of the text; a sheet with a rectangular box in which to write key words; a lined piece of paper for reproducing the text. This had the first sentence of the text written on it. The text involved a narrative genre in the form of a fictional newspaper article reporting an event unique to the Taiwanese context. In order to ensure the texts were not too difficult, the verbs used in the text were selected from the General Service List. This forms the basis of the vocabulary taught in Taiwanese high schools and universities.

There were four narrative tasks originally designed by the researcher. Each of them had a text and six pictures that corresponded to their content. The stories were told from the perspective of a policeman who reported on the events. 'The Lucky Dog' narrative told a story of how a dog caused a car accident in Taipei. 'A Landslide in Nantou' reported the events surrounding a woman being trapped in her house following a landslide. 'The Lost Bag' recounted the events of a woman who lost her bag at the Shi Lin Night Market, and 'Jade Mountain' comprised a report about two students who became lost while hiking on

Jade Mountain.

A number of procedures were employed to ensure that the tasks were designed well, and that they elicited the target structure. The texts were first given to two experienced ESOL tutors at Auckland University, and their feedback was used to revise the texts. Prior to the implementation of the pilot study, the four different writing tasks were trialled on a class of 31 students at the same university in Taiwan so as to ascertain whether they elicited sufficient obligatory occasions. One task, which elicited the fewest obligatory occasions, was removed from the pilot study. This left a total of three texts for the study. These were 'A Landslide in Nantou', 'The Lost Bag' and 'Jade Mountain'.

3.7 Procedures

3.7.1 Writing tasks

The procedure for completing the writing tasks was as follows:

1. The participants were given the task package including the aforementioned materials.
2. They were informed that they had to reproduce a text based on the narrative text provided to the best of their ability.
3. They were instructed to read the text once and underline any unknown vocabulary.
4. The students were put into groups of four and asked to discuss any unknown vocabulary.
5. Any vocabulary that the students did not know was explained by the teacher.
6. The students read the narrative again.

7. The narrative text was collected by the teacher.
8. The narrative was then read to the students twice during which the students noted down important words on the page designed for these key words.
9. The students then compared their key words with a partner and added any new ones to their list.
10. The participants were told that they could use the pictures and the key words to rewrite the text, and that they should double space their writing.
11. The completed writing was collected.

3.7.2 Response sessions

The response sessions had different procedures for both Task 1 and Task 2 and depending on whether or not the groups received direct CF or indirect CF.

1. Task 1
    a. Direct CF groups

       Task 1 was returned to members of the direct CF groups with corrections. They were required to study the CF for five minutes and then the tasks were collected by the teacher. No revision was required to be undertaken.

    b. Indirect CF groups

       The tasks were returned to the participants with indirect CF located on their original texts. They were instructed to self-correct their texts based on the CF provided. These corrected texts were then collected by the teacher. The participants were not required to write a revised version of the original text.

c. Control Group

The Control Group received no CF on their tasks. They therefore were not required to attend to or correct the CF. They were also not asked to complete a revised version of Task 1.They were instead placed into pairs and required to undertake oral fluency tasks. They were given no oral CF at this time.

2. Task 2

a. Direct CF groups

Task 2 was returned to the participants; this had been corrected with direct CF. The learners were required to study the CF for five minutes then the texts were collected by the teacher. Next, their original copies of their Task 2 were returned (i.e. without any CF). The learners were then required to write a revised version of their original text without access to any CF. The revised texts were collected by the teacher.

b. Indirect CF groups

First of all, Task 2 with indirect CF was returned to the learners. The participants were instructed to make corrections based on the CF. Upon completion, the corrected versions were collected. Then a copy of the original version of Task 2 without CF was returned to the participants. They were instructed to write a revised version of the original text. This revised version was collected by the teacher.

c. Control Group

The Control Group again received no CF. The learners were not required to

study their texts or self-correct them nor were they required to provide a revised version of Task 2. They were placed into pairs and completed a series of oral fluency tasks to which no oral CF was provided.

3.8 Data collection

The data were collected over a four week period as presented in Table 8. It involved the trialling of the tasks, the completion of a background questionnaire and the signing of the University of Auckland Human Participants Ethics Committee documents. Additionally, the three tasks were completed and two CF response sessions carried out.

Table 8: Data Collection Schedule for Pilot Study

Week 1	Week 2	Week 3	Week 4
Trial tasks	Ethics documents Background questionnaire Task 1 (new piece of writing)	Response session (task 1) Task 2 (new piece of writing)	Response session (task 2) Task 3 (new piece of writing)

3.9 Analysis

The analyses of the writing tasks involved the use of obligatory occasion analysis and the subsequent statistical analysis of these data through SPSS version 17 so as to test for the reliability of the scoring and to address the research questions.

The type of obligatory occasion analysis chosen for this study was Pica's (1983) Target-Like Use Analysis. It takes into consideration the overuse of a particular form. Ellis and Barkhuizen (2005) provide the following formula for Target-Like Use Analysis:

$$\frac{n \text{ correct suppliances in contexts}}{n \text{ obligatory occasions} + n \text{ suppliance in non-obligatory contexts}} \times 100$$

In the process of scoring simple past tense forms, a scoring issue was raised. This related to errors in obligatory contexts. If a sentence had a verb missing and the context established that a simple past tense form was needed, it was coded as an obligatory occasion but not a correct suppliance. On the other hand, if a verb was missing but there were a variety of possible past tense forms that could be used in that particular context (e.g. the past progressive tense or the past perfect tense), it was not coded as an obligatory occasion.

The results of the obligatory occasion analyses were examined for reliability. This comprised a second scoring of the data two months after the initial scoring by the researcher. Forty texts were randomly selected across the five experimental groups from the pre-test. An intra-rater reliability score of $r = .995$ was calculated using Pearson Product Moment Correlation.

The Target-Like-Use Analysis scores were subject to a variety of statistical analyses to

address the research questions.

1. For research question 1, the scores for Task 1 and Task 2 were analysed. A two-way repeated measures ANOVA (five groups x two times) with post-hoc one-way ANOVAs was conducted.
2. Research question 2 compared the scores of the experimental groups alone (no control) on Task 2 and their revised texts. A two-way repeated measures ANOVA (four groups x two times) with post-hoc one-way ANOVAs was calculated.
3. To investigate research question 3, a two-way repeated measures ANOVA (five groups x two times) with post-hoc one-way ANOVAs was undertaken on the scores for Task 2 and Task 3.

3.10 Results

3.10.1 The effect of corrective feedback plus response on a new piece of writing

The effect of CF plus response on a new piece of writing was investigated by comparing Task 1 and Task 2. Table 9 provides the descriptive statistics for these tasks, and the mean scores are presented graphically in Figure 2.

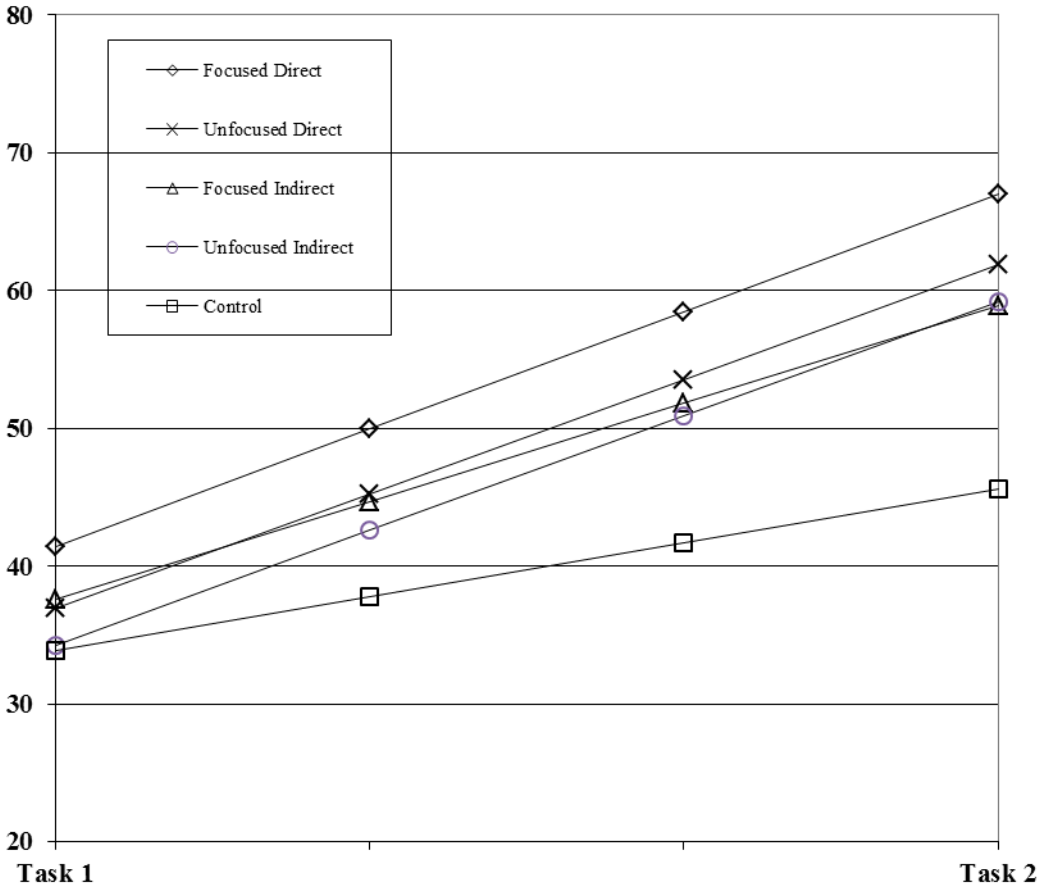
All the mean scores increased for the five groups from Task 1 to Task 2. A Mauchly's Test of Sphericity, however, revealed that the assumption of sphericity was violated so the Greenhouse-Geisser adjustment was used to correct the F values. There was a significant effect for time ($F = 92.33$; d.f. = 1; $p < .001$). There were not, however, for group ($F = 1.81$; d.f. = 4; $p = .13$) or for the time-group interaction ($F = 1.66$; d.f. = 4; $p = .16$).

Table 9: Descriptive Statistics for the Effect of Corrective Feedback plus Response on a New Piece of Writing

Corrective Feedback plus Response	*n*	Task 1		Task 2	
		M	SD	*M*	SD
Focused Direct	29	41.46	30.39	66.97	19.65
Unfocused Direct	30	36.97	25.84	61.85	21.60
Focused Indirect	30	37.62	24.38	58.88	28.42
Unfocused Indirect	29	34.25	28.49	59.23	23.77
Control	33	33.91	30.37	44.74	27.05

All five groups demonstrated significant improvement over time: the Focused Direct CF Group (F = 25.61; d.f. = 1; $p < .001$); the Unfocused Direct CF Group (F = 25.19;.d.f. = 1; $p < .001$); the Focused Indirect CF Group (F = 18.39; d.f. = 1; $p < .001$); the Unfocused Indirect CF Group (F = 24.55; d.f. = 1; $p < .001$); the Control Group (F = 5.86; d.f. = 1; p = .02).

Figure 2: Mean Scores for the Effect of Corrective Feedback plus Response on a New Piece of Writing



3.10.2 The effect of corrective feedback plus response on a revised text

By comparing Task 2 with a revised version of this task, the effect of CF plus response on a revised text could be examined. There was no control group for this comparison. Table 10 presents the descriptive statistics for these two pieces of writing.

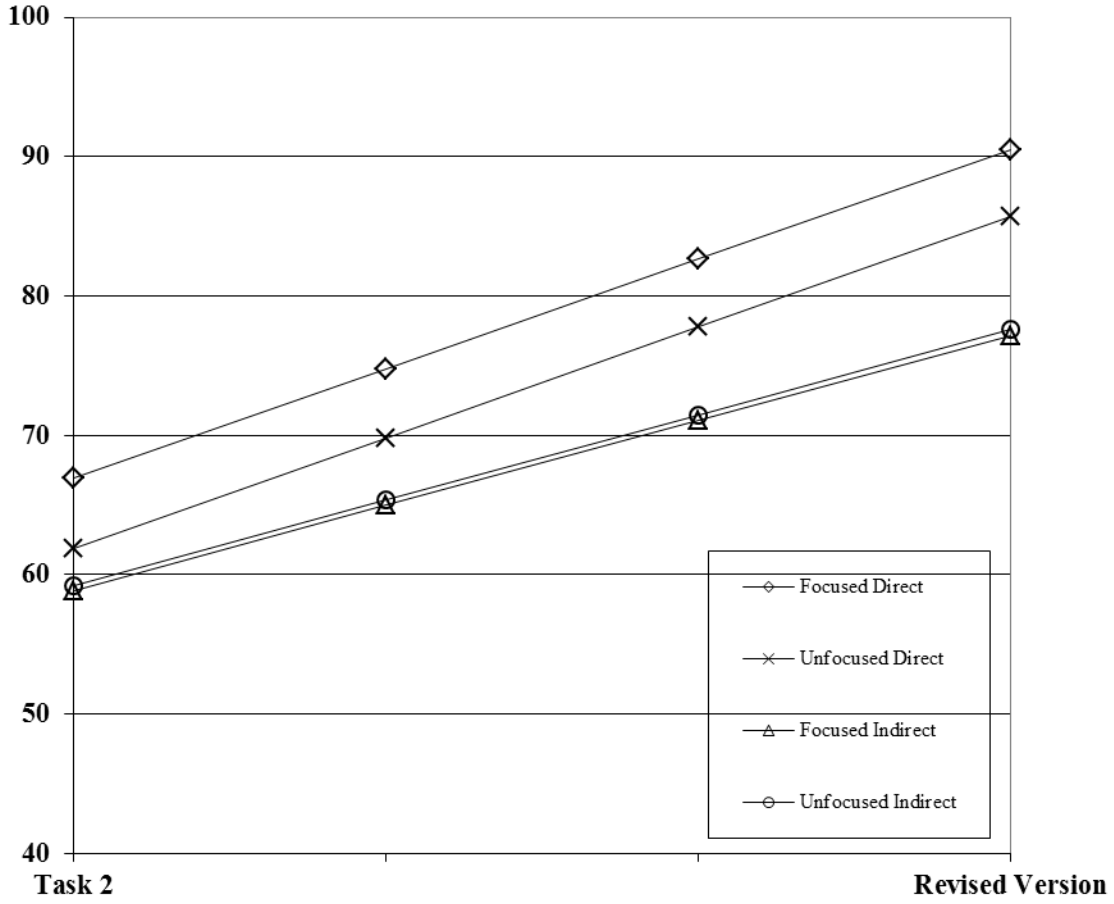
Table 10: Descriptive Statistics for the Effect of Corrective Feedback plus Response on a Revised Version of an Original Text

Corrective Feedback plus Response	*n*	Task 2		Revised Version	
		M	SD	*M*	SD
Focused Direct	29	66.97	16.65	90.46	10.88
Unfocused Direct	30	61.85	21.60	85.73	15.20
Focused Indirect	30	58.88	28.42	77.15	22.86
Unfocused Indirect	29	59.23	23.77	77.58	22.37

As is evident in Figure 3, all the mean scores increased between the task and the revised text. The Greenhouse-Geisser adjustment was used on the *F* scores to account for the violation of the assumption of sphericity following the calculation of Mauchly's Test of Sphericity. While there was found to be significant differences for time (F = 136.38; d.f. = 1; < .001), there were none for either the group effect (F = 2.11; d.f. = 3; p =.10) or the time-group interaction (F = .75; d.f. = 3; p =.53).

In relation to the time effect, all four conditions improved significantly from Task 2 to the revised version of Task 2: focused direct CF (F = 41.93; d.f. = 1; p <.001); unfocused direct CF (F = 44.84; d.f. = 1; p < .001); focused indirect CF (F = 26.25; d.f. = 1; p < .001); unfocused indirect CF (F = 25.06; d.f. = 1; p < .001).

Figure 3: Mean Scores for the Effect of Corrective Feedback plus Response on a Revised Text



3.10.3 The effect of corrective feedback plus response and revision on a new piece of writing

To assess the effectiveness of CF plus response and revision on a new piece of writing, Task 2 and Task 3 were compared. Table 11 presents the descriptive statistics for these tasks, and as is illustrated in Figure 4, all of the mean scores increased between these two tasks.

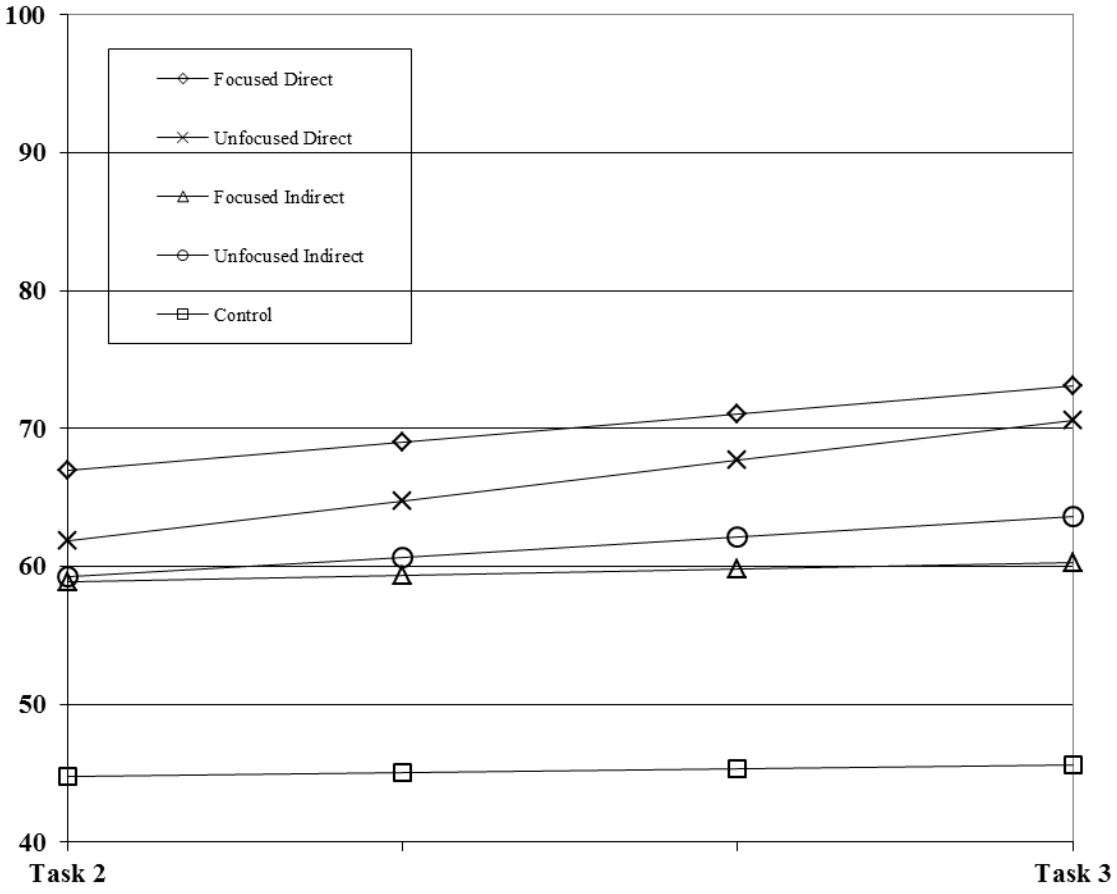
A Mauchly's Test of Sphericity demonstrated that the assumption of sphericity was violated, so the Greenhouse-Geisser adjustment was used to correct the *F* scores. The changes over time approached significance ($F = 5.79$; d.f. = 1; $p = .08$) and reached significance for the group differences ($F = 5.74$; d.f. = 4; $p < .001$); however, there was no significant difference for the time-group interaction ($F = .722$; d.f. = 4; $p = .58$).

Table 11: Descriptive Statistics for the Effect of Corrective Feedback plus Response and Revision on a New Piece of Writing

Corrective Feedback plus Response and Revision	*n*	Task 2		Task 3	
		M	SD	*M*	SD
Focused Direct	29	66.97	19.65	73.10	17.68
Unfocused Direct	30	61.85	21.60	70.64	21.00
Focused Indirect	30	58.88	28.42	60.28	26.00
Unfocused Indirect	29	59.23	23.37	63.61	27.82
Control	33	44.74	27.05	45.59	30.53

With regard to the time differences, there were no significant improvements over time for the Control Group ($F = .051$; d.f. = 1; $p = .82$), Focused Direct CF Group ($F = 2.33$; d.f. = 1; $p = .13$), Focused Indirect CF Group ($F = .126$; d.f. = 1; $p = .72$) and Unfocused Indirect CF Group ($F = 1.19$; d.f. = 1; $p = .28$). The Unfocused Direct CF Group demonstrated a significant improvement between Task 2 and Task 3 ($F = 4.98$; d.f. = 1; $p = .03$).

Figure 4: Mean Scores for the Effect of Corrective Feedback plus Response and Revision on a New Piece of Writing



For the group differences, a one-way ANOVA of Task 2 demonstrated significant differences between the five groups (F = 3.63; d.f. = 4; p < .01). These differences were between the Control Group and both the Focused Direct CF Group and the Unfocused Direct CF Group. To take account of the group differences when examining the performance on Task 3, a post-hoc one-way ANCOVA was conducted using Task 2 as the covariate,

Task 3 as the dependent variable and the five groups as the independent variable. This produced significant differences between the intervention groups ($F = 2.49$; d.f. = 4; $p = .05$). A post-hoc analysis of ANCOVA simple contrasts revealed these to be between the Control Group and the Focused Direct CF Group ($p = .01$) and between the Control Group and the Unfocused Direct CF Group ($p < .01$).

3.11 Discussion

Research question one asked if written CF plus response affected the accuracy of a new piece of writing. Between Task 1 and Task 2, experimental groups received CF and responded to it while the Control Group completed a series of conversation tasks. The results revealed that the writing for all five conditions significantly improved in accuracy between the two times, and there were no significant differences between the groups.

That all conditions improved between a writing task and a second similar one and that there were no group differences between the conditions are surprising findings for a number of reasons. First of all, one would have expected that as the Control Group received no CF, the members of this group would not significantly improve. That the opposite occurred would appear to lend support to Truscott's (1996) claim that writing practice alone can enhance the accuracy of writing. However, there are alternate explanations. The design of the study placed all five groups in the one class at the same time thereby increasing the chances of the participants in the Control Group becoming aware of the target structure. There is also the strong possibility that the design resulted in some type of testing effect or that the

implementation of the tasks themselves raised awareness of the target structures.

Another intriguing finding is the failure to find any group differences between the focused and unfocused types of CF. Ellis et al. (2008) also found no significant differences between their focused and unfocused direct CF conditions. They suggested that rather than CF being seen as focused or unfocused one can consider CF as varying in degree of focus depending on the quantity of corrections the participants' received. Based on such corrections, they were able to classify the groups as focused and less focused. The same may very well have been the case in the pilot study; however, without further analysis of the data, this claim cannot be verified at this stage.

An additional issue relates to the failure to find any group differences in Task Two for the four CF plus response groups. A possible reason for this may reside in the target structure itself. Ferris (2002) separates errors into treatable and untreatable types. Treatable errors include simple morphological grammar rules while untreatable ones are lexical or idiosyncratic in nature. The simple past actually comprises three separate forms: regular past tense verbs, irregular past tense verbs and the copula verb. Both the regular form and copula verb follow simple grammatical rules and can be viewed as treatable whereas irregular past tense verbs are essentially lexical. That is, they follow no grammatical rule and hence may be untreatable. It is quite possible that CF may only be effective with forms that follow a regular grammatical rule. Indeed, there is some evidence to support this in some recent research that has found group differences in a new piece of writing. Bitchener and Knoch

(2008) found group differences between a focused direct CF condition and a control group. Ellis et al. (2008) found such differences between an unfocused direct CF condition and a control group. Both these studies examined the English article system for the functions of first mention and anaphoric reference. This is arguably a treatable feature. Thus, the only evidence in support of the effectiveness of CF on a new piece of writing has come from studies that used target structures that follow simple grammatical rules. Further analysis of the different forms of the simple past tense may very well contribute to a greater understanding of this issue.

While there were no group differences, there is some evidence that the CF conditions provided different levels of accuracy. A way to demonstrate this is to look at the mean improvement scores between Task 1 and Task 2 for each group. As all conditions improved over time, this was calculated by subtracting the mean score for Task 2 from that of Task 1 for each group. These mean scores are presented in Table 9. The Focused Direct CF Group and Unfocused Direct CF Group achieved scores of 25.52 and 22.18, respectively. The Focused Indirect CF Group attained a score of 21.26 and the Unfocused Indirect CF one 24.98. The Control Group, on the other hand, only attained a score of 10.83. It is thus evident that the CF conditions resulted in greater improvement in accuracy than the control group. The Focused Direct CF Group not only achieved the highest mean achievement score, but as a group, the variability of the scores diminished. This is also evident in the descriptive statistics presented in Table 9. One can see that the standard deviation score dropped by a third for this group from 30.09 to 19.65, a pattern not repeated for any of the

other groups. Such a finding suggests that focused direct CF had a relatively even effect on the learners in that group.

The second research question investigated the effectiveness of CF plus response on a revised version of an original text. The four types of CF were given on Task 2 and the participants were required to attend to the direct CF and self-correct the indirect CF before writing a revised version of the original text. They did not have access to the CF during the revision. All four conditions significantly improved between Task 2 and the revised version of Task 2. There were no group differences between the four conditions.

As the participants did not have access to the CF when they wrote the revised version, this research question actually investigated how much of the CF they remembered following either attending to the CF or self-correcting the CF. All four conditions would appear to be as effective in the uptake of CF in a revised version irrespective of CF type or response. Again, however, an inspection of the mean improvement scores provides a more detailed insight into the groups' performance in their revised texts. The Focused Direct CF Group has a score of 23.49, Unfocused Direct CF Group 23.86, Focused Indirect CF Group 18.27 and Unfocused Indirect CF Group 18.35. For CF type plus response, the scores are remarkably similar. Both direct CF conditions achieved scores that were about 5 points higher than the indirect CF plus response groups. The standard deviation scores in Table 10 are also informative. The Focused Direct CF Group's standard deviation score diminished by 35%, Unfocused Direct CF by 30%, Focused Indirect by 19% and Unfocused Indirect by

6%. It would appear that attending to both focused and unfocused direct CF resulted in less variability in the scores than focused indirect CF and unfocused indirect CF.

Research question 3 examined the effect of CF plus response and revision on a new piece of writing. From Task 1 to Task 3, the participants received the four different types of CF, corresponding responses and revision. They had no access to the CF while revising. The members of the Control Group received no CF and completed oral fluency tasks. It was found that there were improvements over time only for the Unfocused Direct CF Group. For task 3, group differences existed between the Control Group and both the focused and unfocused direct CF conditions. The two direct CF groups outperformed the Control Group.

The group differences in Task 3 present the most interesting findings. It would appear that both focused and unfocused direct CF lead to improved accuracy in a new piece of writing when such CF is first attended to and the learners have the opportunity to revise their texts. While Sachs and Polio (2008) investigated attending to unfocused direct CF and revising without access to CF, they only measured the effectiveness on revisions rather than a new piece of writing. Hence, the findings of this study are unique. Direct CF provides the corrections while indirect CF requires learners to correct their own errors using the clues given. Such self-correction could be right or wrong. Thus, in the revision, the participants in the direct CF conditions could draw on correct input whereas those in the indirect CF ones would only have access to those forms they had been able to correct. This is of particular relevance when applied to the SLA notions of noticing (Schmidt, 1990, 2001) and Ellis'

(1994) conceptualization of acquisition as acquiring a partially internalized form. Direct CF alone can provide opportunities to correctly notice new forms and enhances the probability of acquiring partially acquired ones. Indirect CF, on the other hand, allows only for acquiring forms that have already been partially acquired. Another possible reason for these group differences could reside in the type of revision. It could be argued that having the participants rewrite their original version without access to the CF yet immediately after attending to the CF could have provided a greater incentive to remember the CF when writing a second piece of writing. That is, it may have increased the retention of the target forms.

3.12 Conclusion

The purpose of the pilot study was to investigate the research questions and importantly assess that validity of the pilot's design, instruments, procedures and analyses. The pilot study revealed a number of limitations.

1. The design
    a. Independent variables

        There were, in fact, three independent variables being examined: corrective feedback, student responses to the CF and revision. As the direct CF and indirect CF had different responses, it was impossible to make any claims about the effectiveness of CF alone.

    b. Dependent variables

        An issue surrounding the dependent variables (writing tasks) was the failure

to provide delayed post-tests. Thus, no claims could be made about the long term effects of the independent variables.

c. The Control Group

The Control Group was not required to complete a revised version of Task 2. This meant that the effectiveness of revision without CF was not investigated.

d. Counterbalancing

There was no counterbalancing of the tasks. Counterbalancing is used to ensure that varying degrees of task difficulty do not influence the results. This is achieved by dividing the learners in each group by the number of tasks, and then giving each of these sub-groups different tasks at each time yet they complete all of the tasks across time.

2. Instruments

a. The background questionnaire

The questionnaire asked some superfluous questions. There were five additional questions which asked about the amount of time learners spent studying English a week currently, when at high school, junior high school, primary school and at cram schools. It was decided that such information did not greatly enhance contextualizing the learners' English language learning experiences.

3. The procedures

a. The writing tasks

The design of the tasks required the teacher to read out the text to the learners

while they jotted down key words (dictogloss). This meant that the tasks could not be effectively counterbalanced as each of the texts for each of the tasks would have to be read out in the same class. Reading out each of the texts in one class could have increased the chances of students becoming aware of the target structures.

b. Revision

For both the direct CF and indirect CF groups, the decision was made to deny access to their corrections during the revision sessions. This meant the findings could not directly be compared to those of other studies (e.g. Fathman & Whalley, 1990; Chandler, 2003) that allowed the participants access to their corrections.

4. Data analysis

a. Past tense forms

The past tense was scored as one category: however, the three different forms of the simple past could provide insights into the effect of CF on a treatable structure (regular past tense) and a possibly untreatable structure (irregular past tense). The failure to examine how these different forms responded to the treatments is a further limitation to the study.

b. Degree of focus

No attempt was made to investigate how focused or unfocused the CF actually was.

3.13 Changes to the main study

The limitations of the pilot study serve to inform the changes needed to ensure validity of the main study. These changes comprise alterations to the pilot's design, instruments, procedures and data analysis.

1. The design

    a. Independent variables

    Based on the strength of the findings demonstrated in research question 3 whereby direct CF that was attended to outperformed indirect CF that was self-corrected, all learners in all four CF conditions will be required to attend to the CF by studying it for five minutes. That is, the learners in the indirect CF groups will not be required to self-correct their CF. By incorporating this CF response into the CF procedures, one of the independent variables present in the pilot will be removed leaving only CF and CF plus revision conditions. In such a way, CF rather than CF plus response can be examined.

    b. Dependent variables

    To examine the effect of CF on acquisition and CF plus revision on acquisition, both immediate post-tests and delayed post-tests will be incorporated into the design.

    c. Control Group

    The Control Group will be required to complete a revised version of a text.

2. Instruments

    a. The background questionnaire

    The extra questions deemed to be unnecessary will be removed from the questionnaire.

    b. The exit questionnaire

    An exit questionnaire will be designed to examine whether the learners became aware of the purpose of the study and to assess their views as to what they learned from the experience.

3. The procedures

    a. The writing tasks

    The design and quantity of the tasks will be changed. In order to achieve effective counterbalancing and reduce the possibility of the participants becoming aware of the target structure, the procedure for the writing tasks will be changed. The procedure requiring a teacher to read out the texts while the participants write down key words will be removed. Instead, the learners will be required to read the texts silently. Also, it will be necessary to provide an additional task for the delayed test. This will be used with the four tasks used when initially trialling the tasks.

    b. Revision

    In the revision session, the participants will have access to the CF.

4. Data analysis

   a. Past tense forms

      In addition to analysing the past tense as a whole, the regular past tense verb forms, irregular past tense verb forms and past tense copula will be analysed separately so as to investigate whether or not treatable and untreatable errors respond differently to the types of CF as well as to revision.

   b. Degree of focus

      An analysis of the degree of focus will be undertaken by comparing the focused and unfocused conditions. This will comprise combining the focused groups' scores (direct and indirect) and the unfocused groups' scores (direct and indirect) and comparing these amalgamated groups with each other and the Control Group.

   c. Degree of directness

      An analysis of the degree of directness will be conducted. This will involve combining the direct groups (focused and unfocused) and indirect groups (focused and unfocused). A group analysis will be completed between these two groups and the Control Group.

CHAPTER FOUR

METHODOLOGY

4.1 Scope of the research

The aim of this research is to build on the study by Ellis et al. (2008). They investigated the effectiveness of focused and unfocused direct CF on the acquisition of two functions of the English article system. It aims also to investigate the effectiveness of focused and unfocused CF on the acquisition of grammatical forms. However, this study adds to that of Ellis et al. in a number of important ways. Namely, CF distinguished in terms of both focus and directness was examined. Also, the CF was investigated with and without opportunities for the learners to revise their texts. The grammatical feature investigated in the study was the past tense.

As this research is an extension of that conducted by Ellis et al. (2008), a similar type of methodology to that used in their study was employed. This study, then, utilized an experimental approach to answer the research questions.

4.2 Research Questions

Six research questions were used to investigate three areas.

4.2.1 The effect of corrective feedback on new pieces of writing

RQ1. What effect did CF have on the learners' accurate use of past tense structures in new

pieces of writing?

RQ2. Were there any differences in the effects of the different types of CF on the learners' accurate use of past tense structures in new pieces of writing?

4.2.2 The effect of corrective feedback corrective feedback on a revised version of a text

RQ3. What effect did CF have on the learners' accurate use of past tense structures in a revised version of a text?

RQ4. Were there any differences in the effects of the different types of CF on the learners' accurate use of past tense structures in a revised version of a text?

4.2.3 The effect of corrective feedback plus revision on new pieces of writing

RQ5. What effect did CF followed by revision have on the learners' accurate use of past tense structures in new pieces of writing?

RQ6. Were there any differences in the effects of the different types of CF followed by revision on the learners' accurate use of past tense structures in new pieces of writing?

4.3 Target structure

The target structure investigated was the simple past tense. It was exactly the same as that described in the pilot study. It therefore comprised the active voice forms of regular past

tense verbs (e.g. *walked* and *talked*), irregular past tense verbs (e.g. *went* and *did*) and past tense copula verbs (*was* or *were*). Feedback was given on all three forms. However, during the scoring of the data, only sufficient obligatory occasions were found for the regular past tense and the irregular past tense, which were examined separately.

4.4 Population and sample

The population for the study consisted of university-level Chinese learners of English in Taiwan. The sample was taken from freshman and sophomore-year students at a Taiwanese junior college. The college made available English majors in the Department of Applied Foreign Languages. Upon the successful completion of their sophomore year, students at this junior college are awarded an associate bachelor's diploma. If they desire to attain a bachelor's degree, they need to complete an additional two years study at a university. These would be the equivalent of their junior and senior years in a four-year programme.

In their freshman and sophomore years, English majors undertake both required and elective courses. The required ones in the freshman year include English Conversation, International Trade, English Reading and English Writing, and those in the sophomore year are Speech Communication, Business English, Business English Conversation and News English. In both years, students can choose elective courses relating to business, translation, English literature, English conversation and English writing. Typically, each class meets once a week for two hours.

The writing classes tend to focus on different genres such as description, and conversation classes can be broadly described as task-based. According to the teachers of these classes, the emphasis was placed on the practical use of both spoken and written English.

The participants, on average, had been studying English for approximately 9 years and their ages ranged from nineteen to twenty-one years of age. About 90 % of the learners were female and 10% male.

The participants in the study belonged to the same population as those in the pilot study, so the level of proficiency between the two samples can be assumed to be broadly the same. That is, the participants of this study, like those in the pilot study, would not yet have fully acquired the simple past tense. Indeed, when the teachers of the participants used in the main study were consulted about the level of proficiency of the learners, they confirmed that the participants had not completely acquired the simple past tense.

While the population for the pilot study and the main study was the same, the samples were slightly different. Both samples comprised freshman and sophomore-level students and their ages and the length of time they had studied English were the same. There was, however, a difference in the distributions of gender between the pilot study and the main study. Also the participants in the main study were English majors whereas those in the pilot study were not. One could well argue that there would be individual differences that affected the proficiency level of the samples based, in particular, on the participants'

subject major. Namely, one could assume that learners, whose major was English, would have greater motivation to learn English and perhaps greater aptitude. Indeed, a comparison of mean pre-test scores of the pilot study and the main study revealed that those of the main study were higher than those of the pilot.

A total of 106 students were originally included in the study. Throughout the study, however, the number of students reduced considerably. There were a number of reasons for this mortality rate. Due to the duration of the study and multiple data collection episodes, a number of participants failed to participate in part of the data collection because they were absent from or late to class. These participants' data were removed from the study. Several students failed to provide at least two obligatory occasions for the regular simple past or the irregular simple past. The data for these participants were also removed. A sizable number of learners achieved a score above 90 on the pre-tests. This is the commonly recognized score at which learners are deemed to have acquired a particular structure (Brown, 1973). Such participants then were removed from the relevant section of the study. All in all, this meant that there were different numbers of students used to address each of the research questions. The number of participants included in the study for each of the research questions for the regular past tense and irregular past tense is presented in Table 12.

Table 12: Number of Participants for each Research Question and each Structure

Research questions	Regular past tense	Irregular past tense
1 and 2	65	56
3 and 4	57	45
5 and 6	51	42

4.5 Design

The quasi-experimental design involved four experimental conditions of focused direct CF, unfocused direct CF, focused indirect CF, unfocused indirect CF and one control group. One freshman and four sophomore classes were designated as one of these five groups.

As shown in Table 13, the design of the research comprised five writing tasks, two CF sessions and one revision session. It included the steps outlined below. The whole sample first completed writing task 1. The texts for the four experimental groups were corrected and returned in CF session 1 one week later. The participants were required to attend to their corrections. The control group received no CF and did not therefore participate in CF session 1. All five groups completed writing task 2 on the same day as the CF session and then two weeks after this they completed writing task 3. The texts for the four experimental conditions were corrected and returned to the participants in CF session 2 one week after writing task 3 was completed. The participants were required to attend to their corrections. The control group again had no CF. All the five groups then completed a

revision task. This involved writing a second draft for task 3. Following this the groups undertook writing task 4. Two weeks later writing task 5 was completed.

Table 13: Design of the Research

Step 1	Writing task 1	Writing task 1 completed
Step 2	Corrective feedback 1	Corrective feedback on writing task 1 attended to
Step 3	Writing task 2	Writing task 2 completed
Step 4	Writing task 3	Writing task 3 completed
Step 5	Corrective feedback 2	Corrective feedback on writing task 3 attended to
Step 6	Revision task	A revised version of writing task 3 undertaken
Step 7	Writing task 4	Writing task 4 completed
Step 8	Writing task 5	Writing task 5 completed

In order to address possible differences in task difficulty, the writing tasks were counterbalanced. The counterbalancing involved the division of writing tasks equally among the students in a group at different times so as to ensure that all participants completed a different writing task for each time. Writing tasks 1 (Task 1), 2 (Task 2) and 3 (Task 3) were counterbalanced as shown in Table 14. For Task 1 (time 1), each group was divided into three sub-groups. Each of the sub-groups completed a different writing task. While at Task 2 (time 2), each of these groups had completed a different writing task from that they completed in Task 1. Finally, for Task 3 (time 3), they undertook the final task of the three they had yet to complete.

Table 14: Counterbalancing of Tasks 1, 2 and 3

Group	Sub-group	Task 1 (time 1)	Task 2 (time 2)	Task 3 (time 3)
E.g. focused direct CF or control	1	Writing task (a)	Writing task (b)	Writing task (c)
	2	Writing task (b)	Writing task (c)	Writing task (a)
	3	Writing task (c)	Writing task (a)	Writing task (b)

Key: Writing task (a) = 'Lucky Dog'
Writing task (b) = 'KTV'
Writing task (c) = 'Lost Bag'

Writing tasks 4 (Task 4) and 5 (Task 5) were also counterbalanced as can be seen in Table 15. This involved each of the treatment groups first being divided into two sub-groups. For Task 4 (time 4), both of these groups completed a different writing task. When undertaking Task 5 (time 5), the members of the two sub-groups completed the writing task they had not yet done.

Table 15: Counterbalancing of Writing Tasks 4 and 5

Group	Sub-group	Task 4 (time 4)	Task 5 (time 5)
E.g. focused direct CF or control	1	Writing task (d)	Writing task (e)
	2	Writing task (e)	Writing task (d)

Key: Writing task (d) = 'A Landslide in Nantou'
Writing task (e) = 'Jade Mountain'

Table 16 illustrates how the design of the research was used to answer the research questions. Research questions 1 and 2 investigated the effects of CF on new pieces of writing. Task 1 operated as a pre-test, Task 2 as an immediate post-test and Task 3 a delayed post-test. The experimental groups received CF following Task 1 while the Control Group received no CF. Research questions 3 and 4 asked whether CF affected the accuracy of a revised version of a text. Task 3 was used as a pre-test and the revision task the post-test. The experimental groups received CF following Task 3. The Control Group received no CF. Research questions 4 and 5 examined the effects of CF and revision on new pieces of writing. Task 3, in this case, was the pre-test, Task 4 an immediate post-test and Task 5 a delayed post-test. The experimental groups received CF following Task 3 and wrote a revised draft of Task 3. The Control Group did not receive CF and wrote a revised version of Task 3.

Table 16: The Design, Research Questions and Variables

	Task 1	Corrective Feedback 1	Task 2	Task 3	Corrective Feedback 2	Revision Task	Task 4	Task 5
Research Questions 1 and 2	√	√	√	√				
Research Questions 3 and 4				√	√	√		
Research Questions 5 and 6				√	√	√	√	√

Key: √ The dependent variables are the writing tasks (tests) and revision task.

√ The independent variables are the CF and CF plus revision.

4.6 Data collection

4.6.1 Instruments

There were four types of instruments used for the collection of the data. These were a background questionnaire, five writing tasks, a revision task and an exit questionnaire (see Appendix A).

4.6.1.1 Background questionnaire

The background questionnaire was designed to collect information about the participants' age, gender, major and years of studying English. The actual questions asked were as follows:

1. How old are you?
2. Are you male of female?
3. What is your major (e.g. marketing, IT)?
4. How many years have you been studying English?

4.6.1.2 Writing tasks

The writing tasks were used to elicit writing samples. Each task was given to the participants as a task package that included a text, a series of six pictures, corresponding to the content of the text, and a lined piece of paper with the first sentence of the text written at the top of the page. This sentence included a time expression (e.g. yesterday) to signal that the past tense was needed. The instructions were in English. 'The Lucky Dog', below, is an example of one of the writing tasks.

The texts were fictional newspaper reports and a written report of a radio newscast about events that occurred in Taiwan. All the verbs used in the texts were drawn from the General Service List and therefore should have been familiar to the Taiwanese junior college students as the list generally forms the basis of vocabulary taught in high schools and universities. In addition to the four texts used in the pilot study ('The Lucky Dog', 'A Landslide in Nantou', 'The Lost Bag' and 'Jade Mountain'), one more text was written by the researcher. 'KTV' reported the events of two Taiwanese university students who participated in overseas trips through winning some Karaoke Television (KTV) singing competitions. The story was reported by a radio DJ. As this was a new text, it was evaluated by two experienced ESOL teachers. The feedback from these teachers was used to revise the text. This revision comprised a few minor changes to the vocabulary.

The instructions required the learners first to read the text and underline any unknown vocabulary. They were then required to get into groups of four and ask the members of that group about any unknown words. Next the participants were asked to read the text again and look at the pictures. When they were finished, the text was removed and the participants were required to rewrite the story using only the pictures.

Below is an example of one of the writing tasks.

The Lucky Dog

Instructions:

1. Read the newspaper story and underline any new words you do not know.
2. When you have finished, get into groups of four and ask the people in your group what the unknown words mean.
3. If there are some words you do not know, ask your teacher.
4. Look at the pictures and read the story again.

The Lucky Dog

Last Tuesday, a car accident happened on Dunwha South Road. It involved two cars and a lucky dog.

A policeman explained the story. Around 10am, a driver of a black car turned left into the traffic in front of a small red car. The driver then started to drive north when a dog walked onto the road.

The driver stopped suddenly, and he avoided the dog. The red car, however, crashed into the black car, and it damaged the back of the car.

The man in the red car opened his door, jumped out and walked quickly to the other car. He shouted at the man. The man in the black car opened his door and shouted at the man. Then they both laughed because they knew each other. They were old classmates from university, but they had lost contact twenty years ago.

The two friends next decided to look for the lucky dog. They searched under cars, and they looked around some trees. The dog, finally, walked up to the men and barked. The friends laughed again and returned the lucky dog to its owners.

Chinese Name:_____English Name:_____

Student ID:_____

Instructions:

1. Using the pictures, try and rewrite the story.
2 You may continue writing on the back of the page.
3. The start of the story has been given to help you.

Last Tuesday, a car accident happened on Dunwha South Road…

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____


1
2
3
4
5
6

4.6.1.3 Revision task

The revision task comprised a sheet of lined paper with the first line of the relevant text written at the top of the page and a series of instructions written in English. These required the participants to write an improved version of the story. The participants in the CF groups had access to their corrected version of Task 3 and those in the Control Group had access to their uncorrected version of Task 3. None of the groups were given a copy of the pictures when they completed their revised versions.

4.6.1.4 Exit questionnaire

The exit questionnaire was used to establish what the participants thought the purpose of the tasks was and what they believed they had learnt from completing the tasks. The questionnaire was based upon that in Ellis et al. (2008). The questionnaire first asked students whether they thought the tasks were about writing, grammar, reading or vocabulary through a multiple choice question. The participants were told they could choose only one of the four options. The questionnaire also contained an open-ended question asking what the students thought they had learned from the tasks.

The exact questions asked in the exit questionnaire were as follows:

1. What do you think the tasks were about?
    a. Practicing writing
    b. Practicing grammar
    c. Practicing reading
    d. Practicing vocabulary

2. What do you think you have learned from doing all the tasks?

4.7 Procedures

This section describes the procedures for the writing tasks, CF sessions and the revision session.

4.7.1 The procedure for the writing tasks

All five groups completed the writing tasks, and these all followed exactly the same procedure. This comprised a sequence of seven separate stages.

Stage 1

The writing task package was distributed to the participants.

Stage 2

The learners were told that they would have to rewrite a text based on the text provided.

Stage 3

They were asked to read the text once and underline any unknown vocabulary.

Stage 4

The students were told to get into groups of four and ask the members of the group about any unknown vocabulary. The students were allowed to write notes about the vocabulary on the text only. If the students did not know any remaining vocabulary, they could then ask the teacher. The students were given as much time as they wanted to read the text.

Stage 5

The text was then collected by the teacher.

Stage 6

The learners were subsequently instructed to use the pictures to rewrite the text.

Stage 7

When a student had finished, the writing was collected by the teacher.

4.7.2 The procedure for the corrective feedback sessions

CF was provided by the researcher on the experimental groups' writing. The procedure for these sessions followed a sequence of stages.

Stage 1

The tasks were returned to the learners with the relevant CF. The CF was one of the following four types.

1. Focused direct CF

   Only past tense errors were corrected by crossing out the errors and writing corrections above them. Missing words related to the simple past were written above the location where they should have been as indicated by a cursor. These verbs were chosen on the basis of what verb form would most likely be used in such a context.

	was	jogged	ate
Because	^ very hungry yesterdy, I ~~jog~~ home and then I ~~eat~~ a a sandwiches.		

2. Unfocused direct CF

All the linguistic errors were crossed out and the correct forms were written above them. Any extra words were crossed out while any missing words were written above the location where they should have been. The exact location was indicated by a cursor (^). For any missing past tense verbs, a verb that fitted the context was chosen and inserted above the place where a verb was missing. The choice of such verbs was made on the basis of which verb would most likely be used in such a context. In the example below, the irregular past tense verb 'ate' is inserted as it agrees with the notion of hunger presented in the dependent clause.

I was yesterday jogged ate sandwich Because ^ ^ very hungry ~~yesterdy,~~ I ~~jog~~ home and then I ^ a ~~a sandwiches~~.

3. Focused indirect CF

Errors relating solely to the simple past were underlined. Any missing past tense words were indicated by a cursor and a line.

Because ^ very hungry yesterdy , I <u>jog</u> home and then I <u>eat</u> a a sandwiches.

4. Unfocused indirect CF

All errors were underlined in a text. Missing words were indicated by a cursor and

a line (‾^‾).

Because ‾^‾^‾very hungry yesterdy , I jog_ home and then I eat a a sandwiches.

Stage 2

The students were instructed to study the corrections for five minutes. The teacher gave them no additional information, and they were not allowed to confer with fellow students. The instructions differed slightly for the direct and indirect CF.

1. Direct CF

   Study your corrections for five minutes. Do NOT change anything.

2. Indirect CF

   Study all the underlined errors for five minutes. A red line (‾^‾) above your writing means something is missing. Do NOT change anything.

Stage 3

The texts were collected by the teacher.

4.7.3 The procedure for the revision task

The procedures for the revision task were slightly different for the CF groups and the Control Group.

1. Corrective feedback groups

   Stage 1

The students were given the revision task.

Stage 2

Task 3 was returned with corrections written on it.

Stage 3

The learners were instructed to rewrite an improved version of Task 3. They had access to their corrections, and there was no time pressure on the students.

Stage 4

When they were finished, both Task 3 and the revised version of Task 3 were collected by the teacher.

2. Control Group

Stage 1

The students were given the revision task.

Stage 2

Task 3 was returned to the participants. This task had no corrections written on it.

Stage 3

The learners were instructed to write an improved version of Task 3. They had no access to any corrections, and there was no time pressure on the students.

Stage 4

When they were finished, both Task 3 and the revised version of Task 3 were collected by the teacher.

4.7.4 Procedure for the exit questionnaire

Following the completion of the final task, Task 5, the learners were asked to complete the exit questionnaire. The teacher responded only to questions relating to the instructions.

4.8 Data collection schedule

Table 17 shows the schedule for the data collection. The study was conducted over a seven week period. In the first week, The University of Auckland Human Participants Ethics Committee requirements were discussed with the participants and the consent forms signed in accordance with University of Auckland's regulations pertaining to research with human subjects (see Appendix B for copies of the relevant documents). Following this, the background questionnaire was completed before Task 1 was undertaken. The first CF session was undertaken on Task 1 in week two then Task 2 was completed. The participants had only to complete Task 3 in week four. The following week, week five, began with a CF session on Task 3. Then a revision session of the same task was completed. After this, Task 4 was undertaken. Two weeks after this in week 7 the participants completed Task 5. Following this, the learners finally completed the exit questionnaire.

Table 17: Data collection schedule for main study

Week 1	Week 2	Week 4	Week 5	Week 7
Ethics forms completed	CF session		CF session	
Background questionnaire			Revision of task 3	
Task 1	Task 2	Task 3	Task 4	Task 5
				Exit questionnaire

4.9 Analysis

4.9.1 Number of corrections received

The number of past tense corrections received by each participant in each CF group was recorded, and a mean score for that group generated. This was completed separately for the regular past tense verbs and irregular past tense verbs for Task 1 to address research questions 1 and 2, Task 3 for research questions 3 and 4 and Task 3 for research questions 5 and 6.

4.9.2 Scoring of the writing tasks

The writing tasks and revision task were scored using obligatory occasion analysis. As with the pilot study, Pica's (1983) Target-Like Use Analysis was used. The regular past tense forms and irregular ones were scored separately.

The same scoring issue identified in the pilot study was applicable to the main study. This related to obligatory contexts. When a verb was missing and the context established a past tense form was required, it was not coded as an obligatory context as it was impossible to determine whether it was a context for a regular or irregular past tense form. Take, for example, the following sentence.

(ran / jogged)

^

Because I was very hungry yesterday, I home and then I ate a sandwich.

The context demands a simple past tense verb but as both an irregular (ran) and regular (jogged) past tense form could be used, it cannot be scored as an obligatory occasion.

There was another scoring issue. If the context required a past tense verb form but more than one type of past tense could be used such as the past progressive, this was not coded as an obligatory occasion. Consider the example below.

(was jogging / jogged / ran)

^

When I was very hungry yesterday, I home. Later, I ate a sandwich.

One can see here that it is possible to use the past progressive tense or the simple past

tense in this context. Therefore the occasion cannot be scored as obligatory for the regular past tense or irregular past tense.

4.9.3 Reliability

The reliability of the scoring procedure was ascertained by calculating an intra-rater reliability score using Spearman Rank Order Correlation (SPSS version 17). Using Task 1 (the initial pre-test), twenty-five texts were randomly selected from all five treatment groups four months after the initial scoring. These were scored again by the researcher. An intra-rater reliability score of *rho* = .99 was calculated for the group as a whole.

4.9.4 Statistical analysis

The scores from the Target-Like Use Analysis were subject to a series of statistical analyses using SPSS version 17. Two separate analyses were completed to assess the distribution of the scores. For each separate group, Kolmorgorov-Smirnov Tests were undertaken to test for normality. On the whole, these tests found that the samples did not meet the assumption of normality required for parametric tests to be performed (refer Appendix C). Non-parametric tests were therefore used. For each of the tasks that doubled as pre-tests (Task 1 for RQ 1 and RQ2, Task 3 for RQ 3 and RQ 4 and Task 3 for RQ 5 and RQ6), Kruskal-Wallis Tests were used to test for group differences. Generally there were group differences so gain scores were generated to answer RQ 2, RQ 4 and RQ6. For RQ 2, gain scores were calculated by subtracting the scores in Task 1 from the scores in Task 2 (gain score 1) and the scores in Task 1 from those in Task 3 (gain score 2). The

gain score for RQ4 was calculated by subtracting the scores in Task 3 from the scores in the revised version of Task 3 (gain score). To answer RQ 6, two gain scores were required. The scores for Task 3 were subtracted for Task 4 (gain score 1) and the scores from Task 3 were also subtracted from Task 5 (gain score 2).

An examination of the number of corrections received was also completed. This involved tallying the number of corrections received in the corrective feedback sessions for each group - that is, the corrections written on Task 1 and Task 3. Group differences in the number of corrections in the four CF groups were assessed by means of Kruskal-Wallis Tests conducted separately for the regular and irregular past tense. The number of corrections was also correlated with the two gain scores generated for RQ 2, the one for RQ4 and the two for RQ6. These analyses were conducted using Spearman Rank Order Correlations.

Research question 1 was addressed by analysing the scores across Task 1, Task 2 and Task 3. Each of the five groups was subject to a separate Friedman Test. If a significant result was found, post-hoc Wilcoxon Signed Rank Tests with a Bonferonni adjustment was completed. This involved assessing changes from Task 1 to Task 2, Task 1 to Task 3 and Task 2 to Task 3.

Research question 2 involved an assessment between the groups for gain score 1 and gain score 2. Kruskal-Wallis Tests were conducted. If the results were significant, post-hoc

Mann-Whitney U Tests were undertaken. These required Bonferonni adjustments.

Answering research question 3 involved the comparison of the scores across Task 3 and a revised version of Task 3. Wilcoxon Signed Rank Tests were employed.

Kruskal-Wallis Tests were undertaken on the gain score to answer research question 4. Post-hoc Mann-Whitney U Tests with Bonferonni adjustments were computed if the Kruskal-Wallis Tests indicated there was statistical significance.

To answer research question 5, the scores across Task 3, Task 4 and Task 5 were investigated. For each of the groups, individual Friedman Tests were conducted and for those that reached significance, post-hoc Wilcoxon Signed Rank Tests employing Bonferonni adjustments were completed. These were undertaken from Task 3 to Task 4, Task 3 to Task 5 and Task 4 to Task 5.

Research question 6 was answered by conducting Kruskal-Wallis Tests on gain score 1 and gain score 2. A post-hoc analysis using Mann-Whitney U Tests was carried out if appropriate. Bonferroni adjustments were made in such instances.

4.9.5 Data from the exit questionnaire

The exit questionnaire was completed by all the participants after Task 5 including those subsequently removed from the study as explained above. There were two questions in

the exit questionnaire. The first was a multiple choice question, and this needed a simple analysis of percentages. If the participants supplied more than one answer to the multiple choice question, their data was removed. Only two participants out of the 101 who completed the questionnaire provided more than one answer.

The second question was an open-ended question. Written responses to the question were analyzed qualitatively using a content analysis. This content analysis comprised a deductive closed-coding as outlined in Ellis and Barkhuizen (2005). It was deductive in that only the responses that directly answered the question were coded. That is, responses that corresponded to learning writing, grammar, reading or vocabulary were tallied.

CHAPTER FIVE

THE EFFECT OF CORRECTIVE FEEDBACK ON NEW PIECES OF WRITING

5.1 Overview

Chapter five addresses research questions 1 and 2.

RQ1. What effect did CF have on the learners' accurate use of past tense structures in new pieces of writing?

To answer this research question, I will investigate the effect of CF over time (i.e. by examining whether there were any statistically significant differences in the five groups' scores between Task 1 and Task 2, between Task 1 and Task 3, and between Task 2 and Task 3).

RQ2. Were there any differences in the effects of the different types of CF on the learners' accurate use of past tense structures in new pieces of writing?

To answer this research question, I will investigate whether there were any statistically significant differences in the five groups' scores. I will also undertake group comparisons on the basis of focus and directness.

In answering both research questions, I will report separate results for regular past tense

scores and irregular past tense scores. The reason for investigating these structures is that the regular past tense is a rule-based grammatical feature. It therefore represents system-learning. The irregular past tense, on the other hand, involves learning the past forms of individual verbs and involves item-learning. System-learning and item-learning follow different paths to acquisition (Ellis, 2009b, pp. 110-111). As such, it is expected that these two structures may respond differently to CF.

This chapter will be organized according to the type of structure. That is, I will first report results for regular past tense scores and then irregular past tense scores. In the final section of this chapter I will provide a discussion of the results.

5.2 Investigating the distribution of the past tense scores

There were concerns with regard to the distribution of the past tense scores. First of all, it is commonly recognized that a score of 90 represents a level at which a learner can be considered to have acquired a particular structure (Brown, 1973). Therefore, participants that scored above 90 in Task 1, which operated as a pre-test, were removed from this part of the study. The remaining scores were then examined to see if they were normally distributed. The scores for the separate groups were subject to a series of Kolmorgorov-Smirnov Tests, and these found that a number of their distributions were in fact not normal (see Appendix C). Therefore, non-parametric rather than parametric tests were performed. Another concern with the distribution was that there were group differences evident in Task 1. To address this problem, gain scores were calculated and used in the group

comparisons. These were generated by calculating the difference in scores between Task 1 and Task 2 (gain score 1) and Task 1 and Task 3 (gain score 2).

5.3 Number of corrections received

For each of the four CF groups, the number of past tense corrections received in corrective feedback session 1 was tallied for the regular past tense scores and irregular past tense ones. These were then subject to Kruskal-Wallis Tests to test for any group differences. Table 18 presents the descriptive statistics for the regular past tense and Table 19 for the irregular past tense. The tests failed to find any significant differences between the four groups for the regular past tense (χ^2 (3, 49) = 4.54, p = .21) or the irregular past tense (χ^2 (3, 44) = 7.53, p = .06). The latter score, however, was approaching significance.

Table 18: Descriptive Statistics for Regular Past Tense Corrections Received during Corrective Feedback Session 1

Corrective Feedback	*n*	Number	*M*	*SD*	Range
Focused Direct	15	46.00	3.07	1.51	1.00 - 6.00
Unfocused Direct	9	19.00	2.11	1.17	1.00 - 4.00
Focused Indirect	11	34.00	3.09	2.12	1.00 - 8.00
Unfocused Indirect	14	32.00	2.29	2.09	1.00 - 8.00

Table 19: Descriptive Statistics for Irregular Past Tense Corrections Received during Corrective Feedback Session 1

Corrective Feedback	*N*	Number	*M*	*SD*	Range
Focused Direct	11	29.00	2.64	1.38	1.00 -5.00
Unfocused Direct	11	15.00	1.36	0.73	1.00 - 3.00
Focused Indirect	9	26.00	2.89	3.44	1.00 - 10.00
Unfocused Indirect	14	29.00	2.07	2.36	1.00 - 10.00

Spearman Rank Order Correlations for the learners in all the experimental groups were also undertaken on the number of corrections and the gain scores for both structures. For the regular past tense, there was a significant correlation between the number of corrections and both gain score 1 (*rho* = .43, $p < .01$) and gain score 2 (*rho* = .40, $p < .01$). The analysis of the irregular past tense found a significant relationship between the number of corrections and gain score 1 (*rho* = .32, $p = .03$) but not for gain score 2 (*rho* = .24, $p = .12$).

5.4 Results for regular past tense scores

To examine the effect of CF on new pieces of writing for the regular past tense verb forms, their scores were analyzed across Task 1, Task 2 and Task 3 as well as between the groups. Descriptive statistics are presented in Table 20.

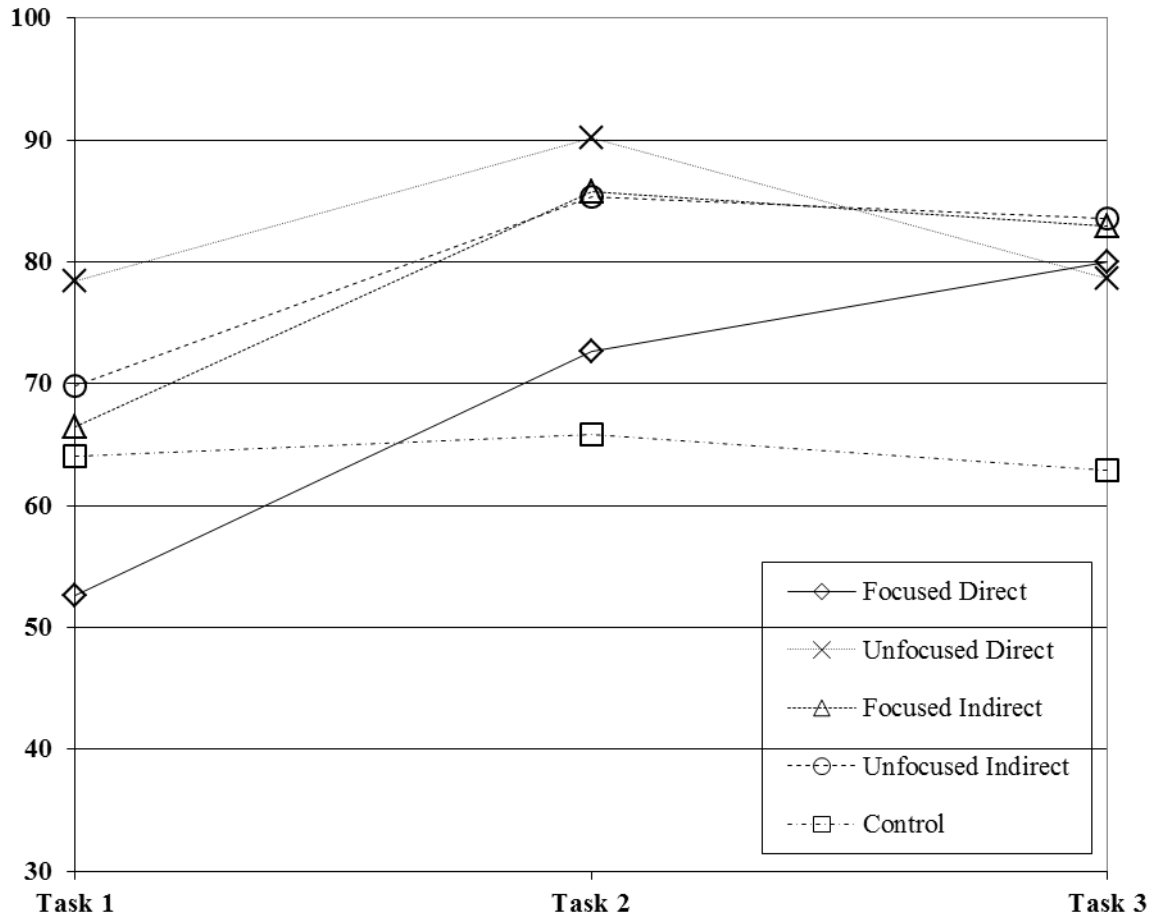
Table 20: Descriptive Statistics for the Effect of Corrective Feedback on Regular Past Scores

Corrective Feedback	*n*	Task 1			Task 2			Task 3		
		M %	*SD*	Range %	*M* %	*SD*	Range %	*M* %	*SD*	Range %
Focused Direct	15	52.61	23.96	0.00 – 81.82	72.69	29.28	0.00 – 100.00	79.99	18.24	60.00 – 100.00
Unfocused Direct	9	78.40	11.47	55.56 – 90.00	90.14	12.37	66.67 – 100.00	78.62	25.25	25.00 – 100.00
Focused Indirect	11	66.52	22.36	14.29 – 85.71	85.73	12.45	62.50 – 100.00	82.98	9.24	66.67 – 100.00
Unfocused Indirect	14	69.84	16.66	33.33 – 88.89	85.37	11.67	66.67 – 100.00	83.55	13.18	60.00 – 100.00
Control	16	64.10	22.20	25.00 – 90.00	65.86	31.09	0.00 – 100.00	62.94	29.32	0.00 – 100.00

5.4.1 Effects of corrective feedback over time

Figure 5 presents the mean scores for the five separate groups across the three tasks. Individual Freidman Tests found there were significant improvements over time for the Focused Direct CF Group (χ^2 (2, 15) = 12.86, $p < .01$), the Unfocused Direct CF Group (χ^2 (2, 9) = 6.73, $p = .04$), the Focused Indirect CF Group (χ^2 (2, 11) = 6.00, $p = .05$) and the Unfocused Indirect CF Group (χ^2 (2, 14) = 8.14, $p = .02$). There was no significant change over time for the Control Group (χ^2 (2, 16) = .90, $p = .64$).

Figure 5: Mean Regular Past Scores for the effect of corrective feedback on new pieces of writing



Post-hoc Wilcoxon Signed Rank Tests with Bonferroni adjustments were undertaken for comparisons involving three tasks (p = .017). The Focused Direct CF Group approached significance from Task 1 to Task 2 (z = -2.17, p = .030)[1], and there was significant improvement from Task 1 to Task 3 (z = -3.35, p < .01). The increase in scores from Task

[1] This change approached significance so a paired samples t-test was undertaken to further investigate whether the Focused Direct CF Group improved between Task 1 and Task 2. The group did significantly improve in accuracy (t (15) = -2.05, p = .03).

1 to Task 2 for the Unfocused Direct CF Group was significant (z = -2.20, p = .015). The Focused Indirect CF Group improved between Task 1 and Task 2 (z = -2.36, p = .018) and approached significance between Task 1 and Task 3 (z = -2.36, p = .037).[2] The Unfocused Indirect CF Group showed improvement from Task 1 to Task 2 (z = -2.61, p < .01) and from Task 1 to Task 3 (z = -2.61, p < .01).

5.4.2. Effects of different types of corrective feedback

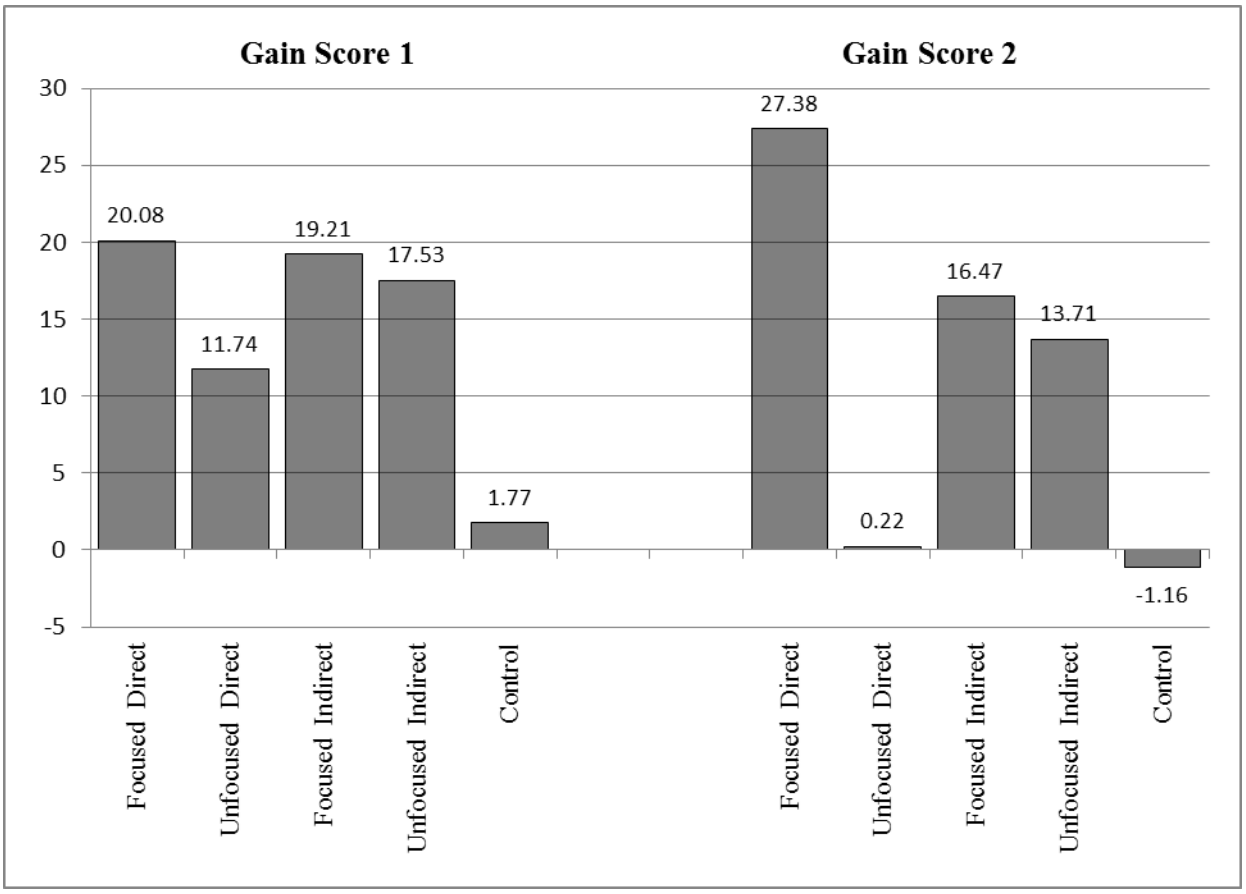
A Kruskal-Wallis Test was performed on Task 1. The result was significant (χ^2 (4, 65) = 9.71, p = .05). Gain scores were therefore calculated. The descriptive statistics are presented in Table 21 and mean gain scores in Figure 6.

Table 21: Descriptive Statistics for Regular Past Tense Gain Scores

Corrective Feedback	*N*	Gain Score 1			Gain Score 2		
		M %	*SD*	Range %	*M* %	*SD*	Range %
Focused Direct	15	20.08	31.07	-35.71 - 71.43	27.38	23.17	-5.00 - 66.67
Unfocused Direct	9	11.74	8.91	-5.00 - 23.08	0.22	22.08	-50.00 - 16.41
Focused Indirect	11	19.21	22.10	-20.83 - 46.67	16.47	23.76	-13.33 - 73.21
Unfocused Indirect	14	17.53	19.03	-23.33 - 43.59	13.71	16.98	-8.33 - 33.33
Control	16	1.77	27.57	-33.33 - 44.44	-1.16	30.89	-35.56 - 41.66

[2] As this analysis only approached significance, a paired samples t-test was undertaken to further assess whether the Focused Indirect CF Group significantly increased from Task 1 to Task 3. It did (t (13) = -3.02, p = .01).

Figure 6: Mean Regular Past Tense Gain Scores for the Effect of Corrective Feedback on New Pieces of Writing



Kruskal-Wallis Tests applied to the gain scores found that while there were no group differences in gain score 1 (χ^2 (4, 65) = 3.86, p = .43), they approached significance in gain score 2 (χ^2 (4, 65) = 8.66, p = .07).[3] As such, a series of 10 post-hoc Mann-Whitney U Tests (one-tailed) were undertaken comparing all possible combinations of the five

[3] As this score only approached significance, a one-way between groups ANOVA was undertaken on gain score 2. The result was significant (F (4, 64) = 3.42, p = .02).

groups.[4] Bonferroni adjustments were required for ten comparisons (p = .005). The results are presented in Table 22. One can see that the Focused Direct CF Group was more accurate than the Control Group (U = 63.00, z = -2.25, p = .01). However, this only approached significance.[5] The Focused Direct CF Group significantly outperformed the Unfocused Direct CF Group (U = 23.00, z = -2.65, p = .003).

Table 22: Results from Mann-Whitney U Tests on Gain Score 2 for the Regular Past Tense

Group	Control	Unfocused Indirect	Focused Indirect	Unfocused Direct	Focused Direct
Focused Direct	p = .01†	p = .07	p = .09	p = .003*	-
Unfocused Direct	p = .44	p = .15	p = .23	-	-
Focused Indirect	p = .14	p = .43	-	-	-
Unfocused Indirect	p = .11	-	-	-	-
Control	-	-	-	-	-

* significant
† approaching significance
- not applicable

[4] To compare the five groups, one-tailed Mann-Whitney U Tests were computed on the grounds that clear directional hypotheses could be made. For example, there is warrant in the literature for claiming that focused direct CF would result in a higher level of learning than unfocused direct CF and focused indirect CF would prove superior to unfocused indirect CF.

[5] As this group approached significance, a further analysis was undertaken using an independent samples t-test. The Focused Direct CF Group significantly outperformed the Control Group (t (29) = -2.89, p < .01).

A further analysis was carried out to compare the two focused groups, the two unfocused groups and the Control Group as well as the two direct groups, the two indirect groups and the Control Group. For the analysis of focus, Kruskal-Wallis Tests failed to find any significant differences between the groups in gain score 1 (χ^2 (2, 65) = 3.20, p = .20); however, the differences approached significance in gain score 2 (χ^2 (2, 65) = 5.71, p = .06).[6] Using the same tests, the analysis of directness failed to find any significant differences in either gain score 1 (χ^2 (2, 65) = 3.37, p = .19) or gain score 2 (χ^2 (2, 65) = 2.95, p = .23).

As the differences approached significance in gain score 2 for the analysis of focus, post-hoc one-tailed Mann-Whitney U-Tests were undertaken.[7] These required a Bonferroni adjustment for three comparisons (p = .017). The tests found that that the difference between the focused groups and the Control Group approached significance (U =128.00, z = -2.07, p = .019) as it did between the focused groups and the unfocused groups (U = 207.00, z = -1.84, p = .03).[8] There was no such difference between the unfocused groups and the Control Group (U = 156.50, z = -.79, p = .22).

---

[6] In this instance, the differences approached significance, so an additional analysis was conducted. A one-way between groups ANOVA found there were significant differences between the three groups (F (2, 64) = 5.20, p < .01).

[7] Directional hypotheses are warranted by the findings of previous studies (i.e. there is evidence that focused CF benefits learning to a greater extent than unfocused CF, that direct CF is superior to indirect CF and both focused CF and direct CF are better than no CF). Therefore, for these analyses one-tailed Mann-Whitney U Tests were employed.

[8] As these two analyses approached significance, additional comparisons were undertaken using independent samples t-tests. They found that the focused groups significantly outperformed the Control Group (t (40) = -2.83, p < .01) and the focused groups were significantly superior to the unfocused groups (t (47) = -2.29, p = .03).

5.5 Results for irregular past tense scores

The effect of CF on irregular past tense structures in new pieces of writing was examined by comparing the irregular past tense scores for Tasks 1, 2 and 3. Descriptive statistics are presented in Table 23.

Table 23: Descriptive Statistics for the Effect of Corrective Feedback on Irregular Past Tense Scores

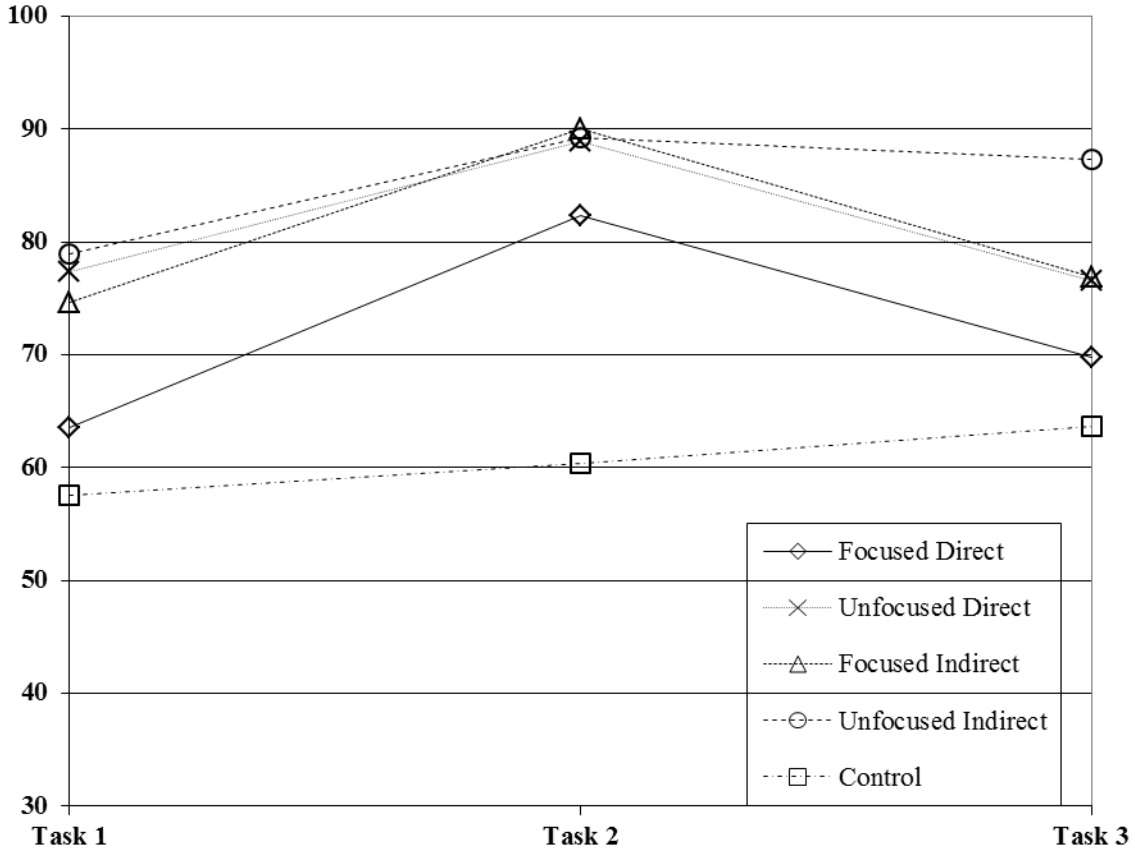
Corrective Feedback	*n*	Task 1			Task 2			Task 3		
		M %	*SD*	Range %	*M* %	*SD*	Range %	*M* %	*SD*	Range %
Focused Direct	11	63.57	16.04	28.57-83.33	82.37	14.52	50.00-100.00	69.71	20.96	25.00-90.91
Unfocused Direct	11	77.33	10.76	60.00-88.89	88.92	8.75	80.00-100.00	76.52	28.20	25.00-100.00
Focused Indirect	9	74.66	18.38	31.25-88.89	90.03	8.50	50.00-100.00	76.88	32.45	60.00-100.00
Unfocused Indirect	14	78.96	7.64	62.96-85.71	89.21	12.74	70.00-100.00	87.31	16.04	50.00-100.00
Control	11	57.52	29.02	0.00-88.89	60.41	37.10	0.00-100.00	63.65	24.24	20.00-100.00

5.5.1 Effects of corrective feedback over time

The changes in mean scores over time are presented in Figure 7. Freidman Tests were conducted on the separate groups and these demonstrated there was a significant change over time for the Focused Direct CF Group (χ^2 (2, 11) = 7.58, p = .02) and the Unfocused

Direct CF Group approached significance (χ^2 (2, 11) = 5.42, p = .07).[9] There were no such significant changes for the Focused Indirect CF Group (χ^2 (2, 9) = 2.12, p = .35), the Unfocused Indirect CF Group (χ^2 (2, 14) = 4.31, p = .17) or the Control Group (χ^2 (2, 11) = .47, p = .79).

Figure 7: Mean Irregular Past Tense Scores for the Effect of Corrective Feedback on New Pieces of Writing



[9] As the changes over time approached significance an additional one-way within groups ANOVA was undertaken. This demonstrated that there were significant changes over time. (F (2, 11) = 7.96, p = .01).

Post-hoc Wilcoxon Signed Rank Tests with Bonferroni adjustments were conducted for three comparisons (p = .017). The improvement between Task 1 and Task 2 approached significance for the Focused Direct CF Group (z = -2.13, p = .03)[10] and reached significance for the Unfocused Direct CF Group (z = -2.58, p = .01).

5.5.2 Effects of different types of corrective feedback

A Kruskal-Wallis Test was computed on the scores for Task 1. The result was significant (χ^2 (4, 56) = 10.67, p = .03). Gain scores were therefore generated. Descriptive statistics are presented in Table 24 and the mean gain scores in Figure 8.

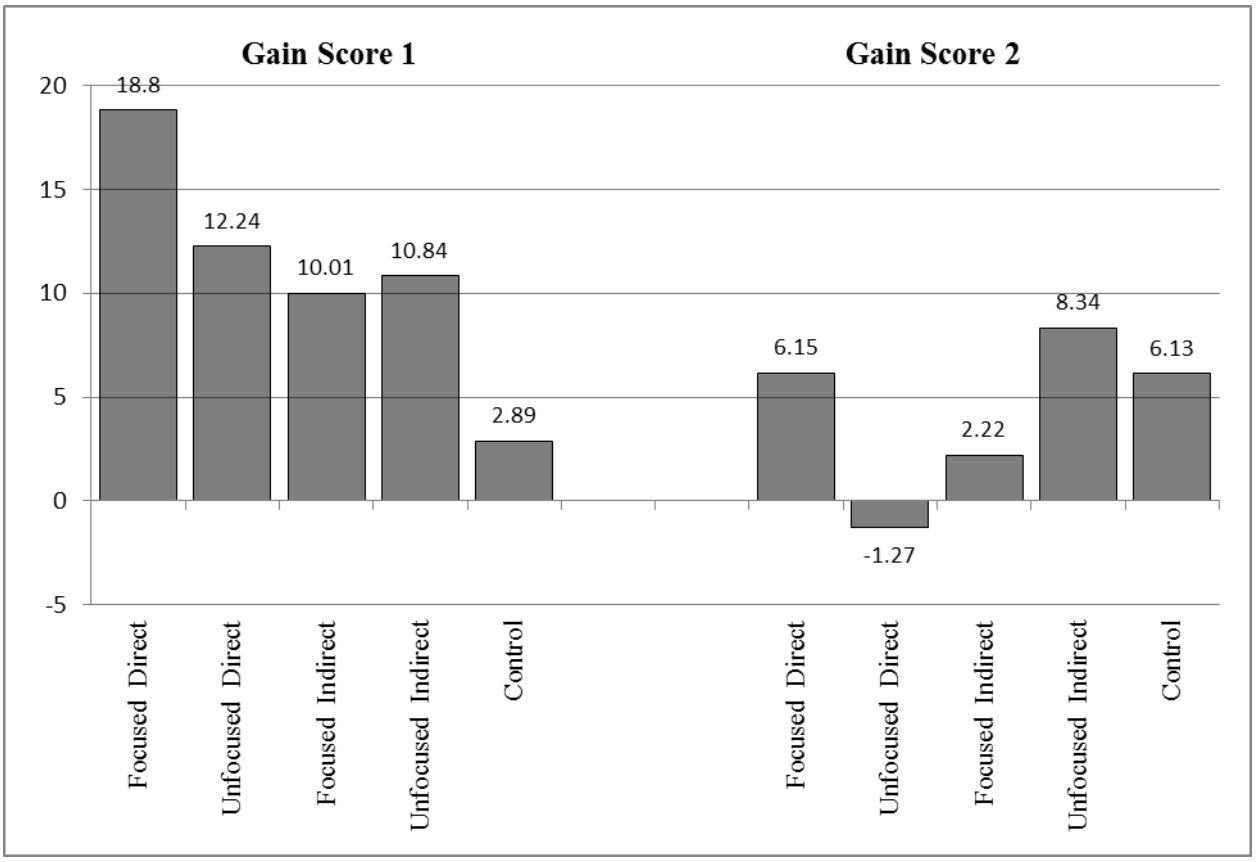
Table 24: Descriptive Statistics for Irregular Past Tense Gain Scores

Corrective Feedback	*N*	Gain Score 1			Gain Score 2		
		M %	*SD*	Range %	*M* %	*SD*	Range %
Focused Direct	11	18.80	27.17	-27.78 - 71.43	6.15	30.29	-47.73 - 62.34
Unfocused Direct	11	12.24	11.87	-8.89 - 33.34	-1.27	30.57	-63.89 - 40.00
Focused Indirect	9	10.01	19.69	-25.00 - 38.75	2.22	36.09	-86.67 - 28.75
Unfocused Indirect	14	10.84	12.87	-10.00 - 33.33	8.34	19.58	-30.00 - 37.04
Control	11	2.89	24.41	-33.34 - 38.01	6.13	31.90	-30.96 - 77.78

[10] The difference approached significance so an additional analysis was undertaken. A paired samples t-test demonstrated that there was a significant improvement between Task 1 and Task 2 for the Focused Direct CF Groups (t (11) = -2.30, p = .05)

Kruskal-Wallis Tests were conducted, and these found there were no significant differences between the groups in gain score 1 (χ^2 (4, 56) = 2.57, p = .63) or in gain score 2 (χ^2 (4, 56) = .67, p = .96).

Figure 8: Mean Irregular Past Tense Gain scores for the Effect of Corrective Feedback on New Pieces of Writing



As with the regular past tense, additional analyses were undertaken by comparing the combined focused groups, unfocused groups and the Control Group as well as the

combined direct groups, indirect groups and Control. There were no significant differences in gain score 1 for focus (χ^2 (2, 56) = 1.65, p = .44) or directness (χ^2 (2, 56) = .72, p = .70), nor were there any in gain score 2 for focus (χ^2 (2, 56) = .18, p = .91) or directness (χ^2 (2, 56) = .52, p = .77).

5.6 Summary of statistically significant results

Table 25 presents a summary of the statistically significant effects of CF over time demonstrated by the Friedman Tests with post-hoc Wilcoxon Signed Rank Tests.

Table 25: Summary of Statistically Significant Results over Time

Structures	Corrective Feedback	Task 1 – Task 2	Task 1 – Task 3	Task 2 – Task 3
Regular Past	Focused Direct	Yes ↑†	Yes ↑	No
	Unfocused Direct	Yes ↑	No	No
	Focused Indirect	Yes ↑	Yes ↑†	No
	Unfocused Indirect	Yes ↑	Yes ↑	No
	Control	No	No	No
Irregular Past	Focused Direct	Yes ↑†	No	No
	Unfocused Direct	Yes ↑	No	No
	Focused Indirect	No	No	No
	Unfocused Indirect	No	No	No
	Control	No	No	No

† approaching significance
↑ significantly increasing scores

There were other near significant results for the investigation of group differences. One - tailed Mann-Whitney U Tests conducted on gain score 2 found that for the regular past tense, the Focused Direct CF Group outperformed the Control Group and this approached significance. The Focused Direct CF Group was also more accurate than the Unfocused Direct CF Group. The differences in this instance reached significance. Again for the regular past tense, one-tailed Mann-Whitney U Tests found that the focused groups outperformed the unfocused groups as well as the Control Group. These results approached significance.

5.7 Discussion

I will begin by discussing the results for the number of corrections of past tense errors that the learners in the four experimental groups received. This will be followed by a discussion of the findings of the exit questionnaire. I will then consider the findings for the regular past and irregular past separately. For each of the structures I will discuss the two research questions that informed the study. In so doing, I will compare the results of my own study to those of other studies on written CF.

5.7.1 Number of corrections received

One of the main differences between this study and previous studies is the extent of the CF the learners received. In most other studies corrective feedback was provided on a number of different pieces of writing. For example, in Ellis et al. (2008) the learners received CF on three separate pieces of writing. In the study reported in this chapter, the learners

received CF on only one piece of writing. In many ways the approach to CF adopted in this study has greater ecological validity. It is doubtful whether writing teachers can afford the luxury of continuously focusing correction on a single grammatical structure. More likely they will need to shift the focus of correction from one problem feature to another over time. Thus, it is important to ask what effect CF directed at a single feature can have on subsequent accuracy in writing when it is provided on a single piece of writing. In fact, the number of corrections that each learner received was quite small. The mean number of corrections only exceeded three in the case of two of the groups for both structures (Focused Direct and Focused Indirect). A number of learners in each group received only one correction.

The crucial issue for the design of the study was whether there were any differences in the number of corrections in the four experimental groups. However, the Kruskal-Wallis Tests showed that the group differences were not statistically significant for either regular or irregular past tense corrections. Thus, differing quantities of the CF received by each group can largely be excluded as an explanation for the gains evident from Task 1 to Tasks 2 and 3 or for the differences in gains among the groups.

It should also be noted that for the regular past tense the correlations between the number of corrections for the four CF groups as a whole were statistically significant for gain score 1 (*rho* = .43, $p < .01$) and gain score 2 (*rho* = .43, $p < .01$); however, this was only attained with gain score 1 (*rho* = .32, $p = .03$) for the irregular past tense. The result was not

significant with gain score 2 (*rho* = .24, *p* = .12). It would appear that the number of corrections is important for improving these learners' accuracy in the use of the regular past tense while the number of corrections would only appear to be important for the irregular past tense immediately following a CF session.

5.7.2 The exit questionnaire

The results for the exit questionnaire are presented in Appendix D. The first question, a multiple choice question, asked what the learners believed the tasks were about - their purpose. Overwhelmingly, the participants thought the aim of the tasks was to develop writing skills (85%). The second question, an open-ended question, asked what the participants learned from the tasks. The majority of the students responded that they had learned to develop their writing skills (60%). These findings are remarkably similar to those presented by Ellis et al. (2008). As such, they provide some evidence against the claims of Xu (2009) that the design of studies such as Ellis et al.'s (2008) one and that presented here make the focus of the studies evident to the learners. For the sample as a whole, it would appear then that by and large the learners were unaware that the primary focus was on past tense verbs.

5.7.3 Regular past tense

5.7.3.1 The effects of corrective feedback on the learners' accurate use of the regular past tense in new pieces of writing

The Control Group demonstrated no significant changes over time. That is, writing

practice alone had no effect on the accuracy of the regular past tense over time. The learners' level of accuracy remained essentially the same over the three pieces of writing. This result then does not support Truscott's (1996) contention that grammatical accuracy will improve simply as a result of writing practice although it is possible that improvement would occur if the practice is more extensive.

Other studies have also reported that writing practice alone does not lead to greater grammatical accuracy. Bitchener and Knoch (2008), for example, found that it had no effect on learners' use of articles. However, one corrective feedback study (Sheen et al. 2009), which also investigated the effects of CF on English articles, did report significant improvements over time for a writing practice group. There are four major differences between Sheen et al.'s study and the study reported here. First, Sheen et al. required the students to complete five writing tasks compared to the three in this study. Second, their study was conducted over a much longer period of time (i.e. nine weeks rather than the four weeks in this study). Third, Sheen et al. measured grammatical accuracy as the ratio of correct to incorrect forms whereas this study used Target-Like Use Analysis. Bitchener and Knoch used a variant of Target-Like Use Analysis, obligatory occasion analysis. It is always difficult to compare the results of studies that differ markedly in design as there is no way of telling which factor is responsible for the reported differences. Finally, their study was conducted in an ESL context whereas this one was conducted in an EFL environment. Arguably, ESL contexts allow for more additional input both inside and outside classrooms compared to EFL ones.

However, it is clearly premature to rule out Truscott's claim that writing practice alone can contribute to grammatical accuracy. Indeed, it would seem likely that this is the case if writing is viewed as a form of pushed output (Swain, 1995), which has been hypothesized to promote acquisition. However, the question arises as to how much of such practice is needed for any benefit to appear. Also, if it can be shown that improvement can be accelerated through the provision of CF, then there is a clear case for not relying on writing practice alone.

All the CF groups improved in the writing task immediately following the CF session. The differences between Task 1 and Task 2 were all statistically significant or approaching significance. However, the results were mixed for the differences between Task 1 and Task 3. In other words, the effect of the CF was much stronger in the writing task immediately following the CF than it was in the writing task completed two weeks later. In fact, the groups showed a decline in the accuracy between Task 2 and Task 3 - the only exception being the Focused Direct Group. None of the time differences between Task 2 and Task 3 reached statistical significance.

These results indicate, therefore, that whereas CF has an immediate effect on accuracy in the regular past tense use, this effect often atrophies over time. Other studies (e.g. Ellis et al., 2008; Bitchener, 2008; Bitchener, 2010b) have also found that CF has an immediate effect but some studies (e.g. Sheen et al., 2009; Bitchener, 2010a; Bitchener, 2010b) have shown that the effect is durable. As noted above, one of the main differences between these studies and the study reported in this chapter is the extent of the CF that was

provided. In this study the learners received CF on only one writing task (i.e. Task 1) and several of the learners received only a very limited number of corrections (i.e. only one correction directed at either regular or irregular past tense). A tentative conclusion, therefore, is that while such limited CF may lead to immediate gains it may not be sufficient to ensure that these are maintained over the long-term. This is an important finding. What is now needed is a study that examines the effect of different quantities of CF.

5.7.3.2 Differences in the effects of the different types of CF on the learners' accurate use of the regular past tense in new pieces of writing

Research question 2 asked whether there were any differences in the effects of different types of CF on learners' accurate use of the past tense in new pieces of writing. Group differences in Task 1 meant that in order to answer this question gain scores needed to be generated. Two sets of analyses were undertaken on these scores. Gain score 1 and gain score 2 were analyzed for between group differences. An analysis of group differences in the gain scores was also undertaken on the experimental groups based on whether the CF was focused as opposed to unfocused and direct as opposed to indirect.

The main finding was that focused CF was shown to be more effective than both unfocused CF and no CF (Control) in the long-term. Closer inspections of the findings provide a number of explanations as to why this was the case.

The difference between the focused and unfocused groups can be explained by a number of different factors. The regular past tense involves system-learning. It constitutes a

structure that can be easily learned explicitly as there is a clear rule-of-thumb. Sheen (2007) hypothesized, furthermore, that focused CF would place less 'attentional strain' on learners. That is, the focused nature of CF would allow learners to notice (Schmidt, 1990) and perhaps notice with metalinguistic understanding (Schmidt, 2001). The fact that the Focused Direct CF Group was the only group not to show a decline from Task 2 to Task 3 provides some support for the claim that focused CF provides less 'attentional strain' on learners in the case of system-learned structures. It could be argued that the learners in this group were able to store an explicit representation of the rule following the completion of Task 2 and then subsequently use this rule when completing Task 3. That is, it would appear they had access to the rule from their long-term memory.

Further support for the efficacy of focused CF over unfocused CF can be found in the fact that the Focused Direct CF Group outperformed the Unfocused Direct CF Group. The findings here lend support to the claim that the number of corrections a participant receives influences the effectiveness of CF. The Focused Direct CF Group in this study received an average of 3.07 corrections while the Unfocused Direct CF Group received 2.11. Also the range of corrections presented in Table 18 suggests that the learners in the Unfocused Direct Group tended to receive only a single correction. Indeed, a breakdown of these scores revealed that 66% of the errors required a single correction whereas this figure was only 18% for the Focused Direct CF Group. It would appear then that the number of corrections a learner receives effects their ability to notice the corrections. This claim is supported by the fact that there was a significant positive correlation between the number

of corrections the learners received and their gain scores.

Related to this issue of the number of corrections is the initial level of pre-existing knowledge of the regular past tense. The Unfocused Direct CF Group had by far the highest level of pre-existing knowledge in Task 1 with a pre-test score of 78.10, so they made fewer errors and hence had fewer corrections. Thus they did not have as many opportunities to notice the errors.

That the focused groups outperformed the Control Group in the long-term indicates the efficacy of focused CF over writing practice alone. Perhaps the number of corrections received by these groups was sufficient for the learners to develop explicit representations of the rule for the regular past.

The finding that the Focused Direct CF Group tended to outperform the Control Group provides evidence in support of this claim. As suggested above, the number of corrections received by the Focused Direct CF Group was sufficient for the learners to apply a rule for the regular past tense in gain score 2; however, the writing practice group had no such feedback. The three episodes of writing did not enable the learners to improve their accuracy in the use of the regular past tense.

How then do these findings reflect those in the literature? There have been several studies that have found a positive effect for focused CF. A number of studies have found a long-

term effect for focused direct CF compared to a control group (e.g. Sheen, 2007; Bitchener, 2008; Bitchener & Knoch, 2008; Sheen, 2010; Bitchener & Knoch, 2010a; Bitchener & Knoch, 2010b). Indeed, Bitchener & Knock 2010a demonstrated this not only eight weeks after the treatment but also six months and ten months after receiving the initial CF. The study presented here, then, lends support to the growing body of research that demonstrates the efficacy of focused direct CF compared to writing practice alone.

There were no short-term group differences between the Focused Indirect CF Group and the Control Group. This result contrasts with Bitchener and Knoch (2010b), who found their focused indirect CF group was more accurate than a writing practice group in an immediate post-test. Why then do these two studies differ? In this study, non-parametric tests were used while in Bitchener and Knoch's study they used parametric tests. It is commonly recognized that due to the less sensitive nature of non-parametric tests, they sometimes fail to find significant differences between groups that might appear with parametric tests, which are more powerful (Pallant, 2007, p. 210). Bitchener and Knoch's study was furthermore conducted in an ESOL context while this one was conducted in an EFL one. These factors may account for the different findings between this study and theirs.

Studies that have compared focused and unfocused CF have produced inconsistent results. Whereas this study found significant differences between a focused direct CF group and an unfocused direct CF group in the long-term, Sheen et al. (2009) found such a difference only in the short-term. That is, there were no such gains evident in their delayed post-test.

Ellis et al. (2008) failed to find any significant differences between their focused and unfocused groups in either the short-term or the long-term. Sheen et al.'s (2009) learners may have benefited in the short-term because they received a substantial number of corrections (more than in the present study). However, it is not clear why the effect was not durable – possibly it was due to the different populations investigated (i.e. ESOL learners in a community college as opposed to university EFL learners in this study). The reason why Ellis et al. (2008) found no difference between their focused and unfocused groups may have been because the focused and unfocused CF was not clearly different as they themselves suggested. In this study the two types were more clearly distinguished. The difference between the Focused Direct CF Group and the Unfocused Direct CF Group in the long-term may reflect this fact.

5.7.4 The irregular past tense

The results for research question 1 indicate that only the two direct CF groups improved in accuracy significantly or near significantly between Task 1 and Task 2. From Task 1 to Task 2 all the remaining groups' scores improved but not significantly. All of the groups' scores deteriorated between Task 2 and Task 3. However, these differences were not statistically significant. The results for research question 2 show that there were no significant group differences in the gain scores in either the short-term or the long-term. This was the case for both the individual groups and for the comparisons involving directness and focus.

It would appear on the face of it that that there is a short-term within-group effect in the case of focused and unfocused direct CF. However, in order to obtain a clearer picture of the effect of the different kinds of CF on accuracy in the use of irregular past tense forms, it is necessary to examine to what extent learners were able to use the specific irregular forms that they produced incorrectly and that were corrected in their subsequent writing. This is because, as pointed out before, the acquisition of irregular verb forms involves item- rather than system-learning.

For the four CF groups, the data were inspected to determine whether those irregular verbs that were corrected in Task 1 were not used, used correctly or used incorrectly in a subsequent piece of writing. This was undertaken for both Task 2 and Task 3 (see Appendix E). Table 26 shows the frequency of learners' correct and incorrect use of those irregular verb forms that had been corrected in Task 1. It is clear that by and large the learners were not able to make use of the corrections they had received in new pieces of writing. Less than 50% of the irregular verbs that had received correction were subsequently used correctly.

Table 26: Correct and Incorrect Use of Those Irregular Verb Forms Corrected in Task 1

Corrective Feedback	Task 2		Task 3	
	Correct	Incorrect	Correct	Incorrect
Focused Direct	9	17	5	22
Unfocused Direct	4	11	0	15
Focused Indirect	3	19	2	20
Unfocused Indirect	4	22	0	26

In order to assess whether there were any group differences in the correct and incorrect use of irregular verbs in Tasks 2 and 3, a series of Fisher Exact Tests were conducted. Six separate comparisons were carried out for each task. The results are presented in Table 27 and Table 28. The only significant group difference was between the Focused Direct CF Group and the Unfocused Indirect CF Group in Task 3. The results for the Focused Direct CF group suggest studying the irregular past tense forms did help the learners to use these forms correctly in new pieces of writing. This type of correction is both intensive and provides learners with the correction. In contrast, Unfocused Indirect CF is less likely to make the learners aware of the need to pay attention to irregular verb forms and does not provide them with the correction.

Table 27: Results from the Fisher Exact Tests on Task 2

Group	Unfocused Indirect	Focused Indirect	Unfocused Direct	Focused Direct
Focused Direct	$p = .20$	$p = .18$	$p = .73$	-
Unfocused Direct	$p = .43$	$p = .42$	-	-
Focused Indirect	$p = 1.0$	-	-	-
Unfocused Indirect	-	-	-	-

Table 28: Results from the Fisher Exact Tests on Task 3

Group	Unfocused Indirect	Focused Indirect	Unfocused Direct	Focused Direct
Focused Direct	$p = .05$*	$p = .44$	$p = .39$	-
Unfocused Direct	$p = .37$	$p = 1.0$	-	-
Focused Indirect	$p = .19$	-	-	-
Unfocused Indirect	-	-	-	-

* significant

5.8 Summary

There was no significant effect over time for writing practice alone. All four CF groups demonstrated greater accuracy in the use of the regular past tense in the short term but this

effect largely atrophied over time. The analyses of overall correct irregular verb forms produced in the new pieces of writing failed to show any short or long-term effect for any type of CF. This conclusion is borne out by the analysis of those specific irregular verbs that had received correction. In all four CF groups, less than 50% of the corrections were uptaken in new pieces of writing.

Truscott (2004, p. 337) asked "...whether correcting is better for the development of accuracy than not correcting..." The results for Research Question 1 suggest that the answer to this question is not straightforward. In the case of irregular past tense there is no clear evidence that the CF (of any type) helped. In the case of regular past tense, however, there is evidence to suggest that focused CF is effective in the long-term. This finding lends support to Ferris' (1999) assertion that some types of error are more treatable than others. System-learned structures would appear to respond better to CF than item-learned ones.

Research question 2 asked whether there were there any differences in the effects of the different types of CF on the learners' accurate use of past tense structures in new pieces of writing. The results showed that focused corrective feedback is more effective than unfocused corrective feedback in the long-term for the regular past tense. An explanation for this can be found Sheen's (2007) claim that focused CF places less 'attentional strain' on learners than unfocused CF and makes the nature of the linguistic problem salient to the learners. Once again, though, this is only applied to 'treatable' grammatical structures like

the regular past tense.

The post-hoc analyses of the irregular past tense verbs indicated that CF had little effect. Those irregular verbs that had been corrected were typically not used correctly in subsequent pieces of new writing. However, focused direct correction does seem to result in greater uptake of the corrected forms than other types of CF in particular unfocused indirect correction.

CHAPTER SIX

THE EFFECT OF CORRECTIVE FEEDBACK ON A REVISED VERSION OF A TEXT

6.1 Overview

This chapter deals with research questions 3 and 4.

RQ3. What effect did CF have on the learners' accurate use of past tense structures in a revised version of a text?

RQ4. Were there any differences in the effects of the different types of CF on the learners' accurate use of past tense structures in a revised version of a text?

In order to answer research question 3, I will examine the effect of CF over time by investigating whether there are any significant changes in the five groups' scores between Task 3 and a revised version of Task 3. I will address research question 4 by examining whether there are any significant differences between the five groups' scores. It should be noted, furthermore, that for the revised versions, the learners had access to the corrections while revising. To answer both research questions, I will detail individual results for regular past tense scores and irregular past tense ones. In the final section of the chapter, I will discuss the results.

6.2. Investigating the distribution of the past tense scores

In line with Brown's (1973) criteria, participants who scored above 90 % on Task 3, which operated as a pre-test, were removed from this part of the research as they were considered to have already acquired the structures. The distributions of the remaining scores were examined to see whether they were normally distributed. Kolmorgorov-Smirnov tests found that some groups were not normally distributed (refer Appendix C), so non-parametric rather than parametric tests were used. For the analysis of group differences, while there were no group differences evident in Task 3, it was decided to generate gain scores for the sake of a consistent analysis across all the research questions investigating group differences (i.e. RQ2, RQ4 and RQ6). The gain scores were generated by calculating the difference between the scores in Task 3 and the revised version of Task 3.

6.3 Number of corrections received

The number of corrections received in corrective feedback session 2 was tallied for each of the four groups for both the regular and irregular past tense. Table 29 presents the descriptive statistics for the regular past tense. A Kruskal-Wallis Test was computed to gauge whether there were any significant differences between the four CF groups in the number of corrections received. The result was not significant (χ^2 (3, 47) = 4.27, p = .23). The descriptive statistics for the number of corrections received by the groups for the irregular past tense are presented in Table 30. A Kruskal-Wallis Tests was also conducted on these scores to test whether there were any significant differences in the number of corrections received. Again the result was not significant (χ^2 (3, 35) = 3.84, p = .28).

Table 29: Descriptive Statistics for Regular Past Tense Corrections Received during Corrective Feedback Session 2

Corrective Feedback	*n*	Number	*M*	*SD*	Range
Focused Direct	10	28.00	2.80	1.99	1.00 - 7.00
Unfocused Direct	10	28.00	2.80	1.32	1.00 - 5.00
Focused Indirect	16	31.00	1.94	0.85	1.00 - 4.00
Unfocused Indirect	11	20.00	1.82	0.75	1.00 - 3.00

Table 30: Descriptive Statistics for Irregular Past Tense Corrections Received during Corrective Feedback Session 2

Corrective Feedback	*n*	Number	*M*	*SD*	Range
Focused Direct	10	21.00	2.10	1.10	1.00 - 4.00
Unfocused Direct	9	21.00	2.33	1.00	1.00 - 4.00
Focused Indirect	9	12.00	1.71	1.20	1.00 - 4.00
Unfocused Indirect	7	29.00	2.07	1.11	1.00 - 4.00

Spearman Rank Order Correlations were also calculated to assess whether there were any correlations between the combined number of corrections received by the four groups and the gain scores. These were carried out separately for both structures. There was a significant and positive correlation between the number of corrections received and gain

scores for the regular past tense (*rho* = .38, $p < .01$) and the irregular past tense (*rho* = .51, $p < .01$).

6.4 Results for regular past tense

The effect of corrective feedback on a revised version of a text for the regular past tense was investigated by comparing the scores for Task 3 and the revised version of Task 3. This was done over time and between groups. Descriptive statistics are displayed in Table 31.

Table 31: Descriptive Statistics for the Effect of Corrective Feedback on a Revised Version of Task 3 for the Regular Past Tense

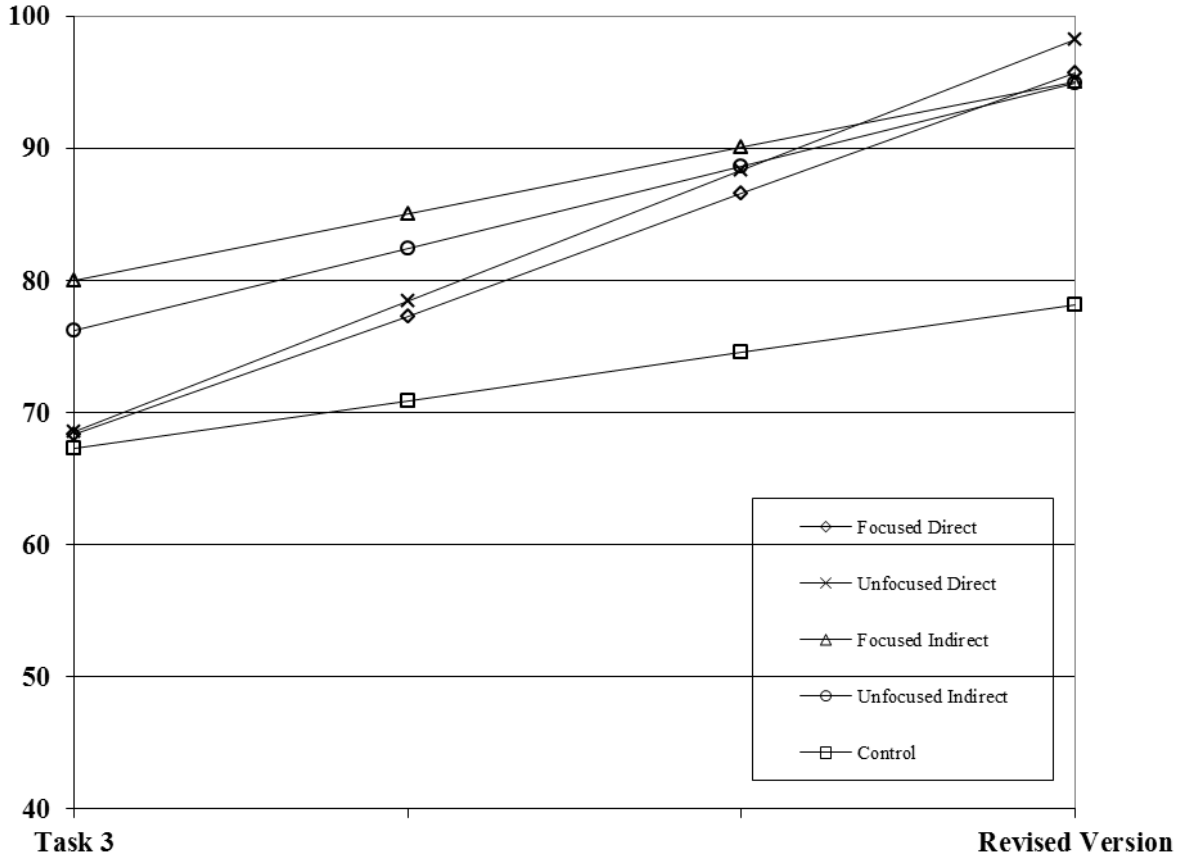
Corrective Feedback	*n*	Task 3			Revised Version		
		M %	*SD*	Range %	*M* %	*SD*	Range %
Focused Direct	10	68.29	15.97	33.33 - 90.00	95.68	9.24	75.00 - 100.00
Unfocused Direct	10	68.51	19.68	25.00 - 87.50	98.17	3.88	90.00 - 100.00
Focused Indirect	16	79.98	9.22	60.00 - 90.00	95.58	7.50	80.00 - 100.00
Unfocused Indirect	11	76.24	9.11	60.00 - 90.00	94.88	9.56	75.00 - 100.00
Control	10	67.25	13.01	44.44 - 87.50	78.16	20.20	33.33 - 100.00

6.4.1 Effects of corrective feedback over time

The mean scores for the effect of CF on a revised version of Task 3 are presented in Figure

9. Separate Wilcoxon Signed Rank tests were undertaken, and there were significant increases in the scores between Task 3 and the revised version for the Focused Direct CF Group (z = -2.81, p < .01), the Unfocused Direct CF Group (z = -2.81, p < .01), the Focused Indirect CF Group (z = -3.41, p < .01) and the Unfocused Indirect CF one (z = -2.94, p < .01). The Control Group failed to demonstrate any significant change over time (z = -1.54, p = .12).

Figure 9: Mean Regular Past Tense Scores for the Effect of Corrective Feedback on a Revised Version of Task 3



6.4.2 Effects of different types of corrective feedback

To test for group differences in Task 3, a Kruskal-Wallis Test was conducted. The result approached significance (χ^2 (4, 57) = 9.15, p = .06). Gain scores were generated. The descriptive statistics for the gain scores are presented in Table 32 and the mean gain scores in Figure 10.

Table 32: Descriptive Statistics for the Effect of Corrective Feedback on Regular Past Tense Gain Scores

Corrective Feedback	*n*	Gain Score		
		M %	*SD*	Range %
Focused Direct	10	27.40	13.18	6.82 - 44.44
Unfocused Direct	10	29.66	20.76	10.00 - 30.77
Focused Indirect	16	15.60	10.20	-8.89 - 33.33
Unfocused Indirect	11	18.64	8.37	2.78 - 33.33
Control	10	10.91	20.56	-33.33 - 44.45

A Kruskal-Wallis Test applied to the gain score found the differences among the groups approached significance (χ^2 (4, 57) = 8.54, p = .07).[11] A series of post-hoc Mann-Whitney

[11] As this score only approached significance, a one-way between groups ANOVA was undertaken on gain score 2. The result was significant (F (4, 56 = 2.82, p = .03).

U tests (one-tailed) were therefore undertaken.[12] These required a Bonferroni adjustment for 10 comparisons (p = .005). The results from the tests are presented in Table 33. Approaching significance, the Focused Direct CF Group outperformed the Control Group U = 25.50, z = -1.85, p = .03) and the Focused Indirect CF Group (U = 39.00, z = -2.16, p = .03). The Unfocused Direct CF Group wrote more accurate revised versions than the Control Group (U = 24.50, z = -1.93, p = .03) and the Focused Indirect CF Group (U = 48.50, z = -1.66, p = .03). It should be noted that hese differences only approached significance, however.[13]

[12] To compare the five groups, one-tailed Mann-Whitney U Tests were computed on the grounds that clear directional hypotheses could be made. There is, for example, justification in the literature for claiming that unfocused direct CF would result in more accurate revisions than unfocused indirect CF.

[13] As these group differences approached significance, a series of additional independent t-tests were conducted. The Focused Direct CF Group had significantly more accurate revisions compared to the Focused Indirect CF Group (t (24) = -2.34, p = .03). Approaching significance, the Focused Direct CF Group outperformed the Control Group (t (18) = -1.94, p = .07), the Unfocused Direct CF Group did better than the Control Group (t (18) = -2.03, p = .06) and the Unfocused Direct CF Group had more accurate revisions compared to the Focused Indirect CF Group (t (24) = -2.39, p = .06).

Figure 10: Mean Regular Past Tense Gain Scores for the Effect of Corrective Feedback on a Revised Version of Task 3

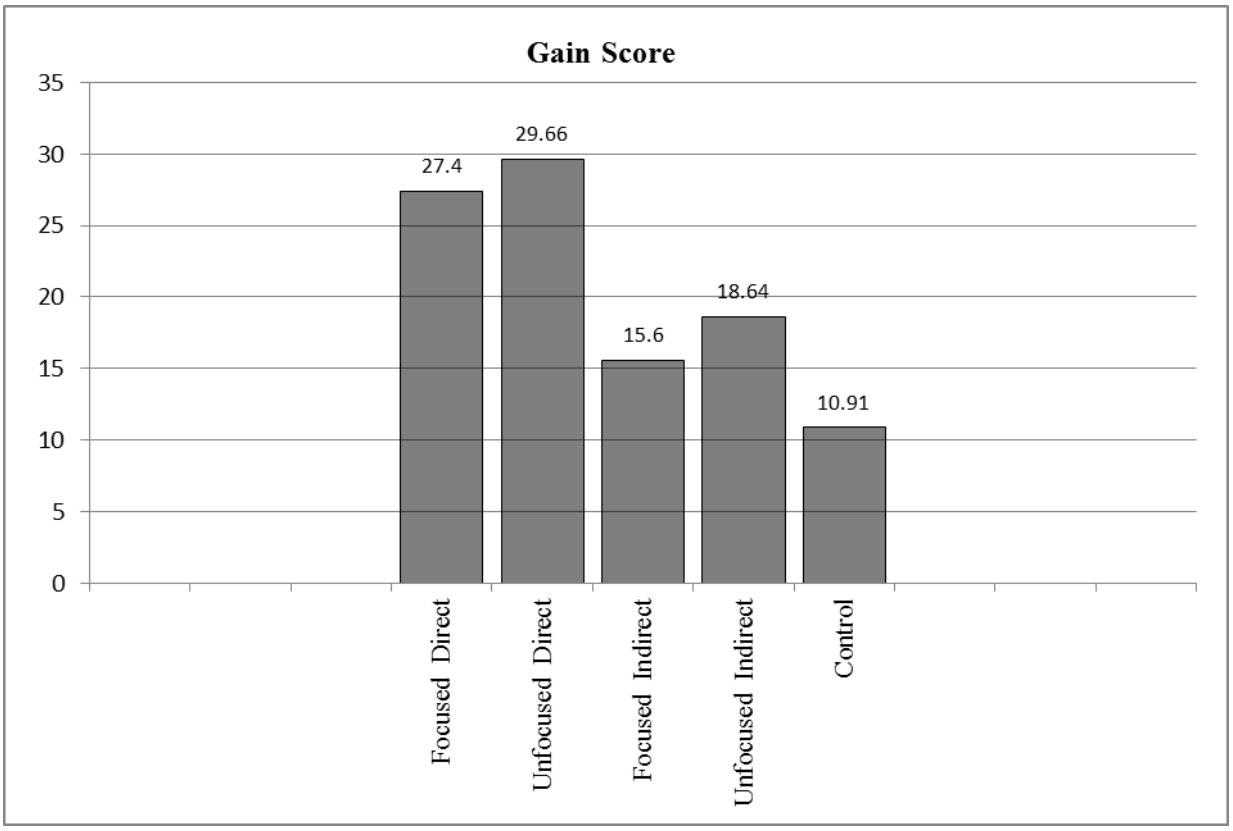


Table 33: Results from the Mann-Whitney U Tests on the Revised Version of Task 3

Group	Control	Unfocused Indirect	Focused Indirect	Unfocused Direct	Focused Direct
Focused Direct	$p = .03$†	$p = .10$	$p = .02$†	$p = .45$	-
Unfocused Direct	$p = .03$†	$p = .15$	$p = .05$†	-	-
Focused Indirect	$p = .21$	$p = .16$	-	-	-
Unfocused Indirect	$p = .11$	-	-	-	-
Control	-	-	-	-	-

† approaching significance
- not applicable

A further analysis was carried out to compare the two focused groups, the two unfocused groups and the Control Group as well as the two direct groups, the two indirect groups and the Control Group. Kruskal-Wallis Tests found that for the gain score there were no significant differences in focus (χ^2 (2, 57) = 3.63, $p = .16$); however, there were in directness (χ^2 (2, 57) = 6.67, $p = .02$).

As there were significant differences in the analysis of directness, one-tailed Mann-Whitney U Tests were undertaken.[14] These required a Bonferroni adjustment for 3

[14] Directional hypotheses are warranted by the findings of previous studies (i.e. there is evidence that focused CF benefits learning to a greater extent than unfocused CF, that direct CF is superior to indirect CF and that CF that is focused and direct outperforms control groups). Therefore, for these analyses one-tailed Mann-Whitney U Tests were employed.

comparisons (p = .017). The direct groups significantly outperformed the Control Group (U = 50.00, z = -2.20, p = .01) and the indirect groups (U = 164.00, z = -2.28, p = .01). There were no such differences between the indirect groups and the Control Group (U = 101.50, z = -1.15, p = .13)

6.5 Results for the irregular past tense

The effect of CF on a revised version of a text for the irregular past tense was examined by comparing the scores for Task 3 and a revised version of Task 3. Comparisons were undertaken across time and between groups. Descriptive statistics are presented in Table 34.

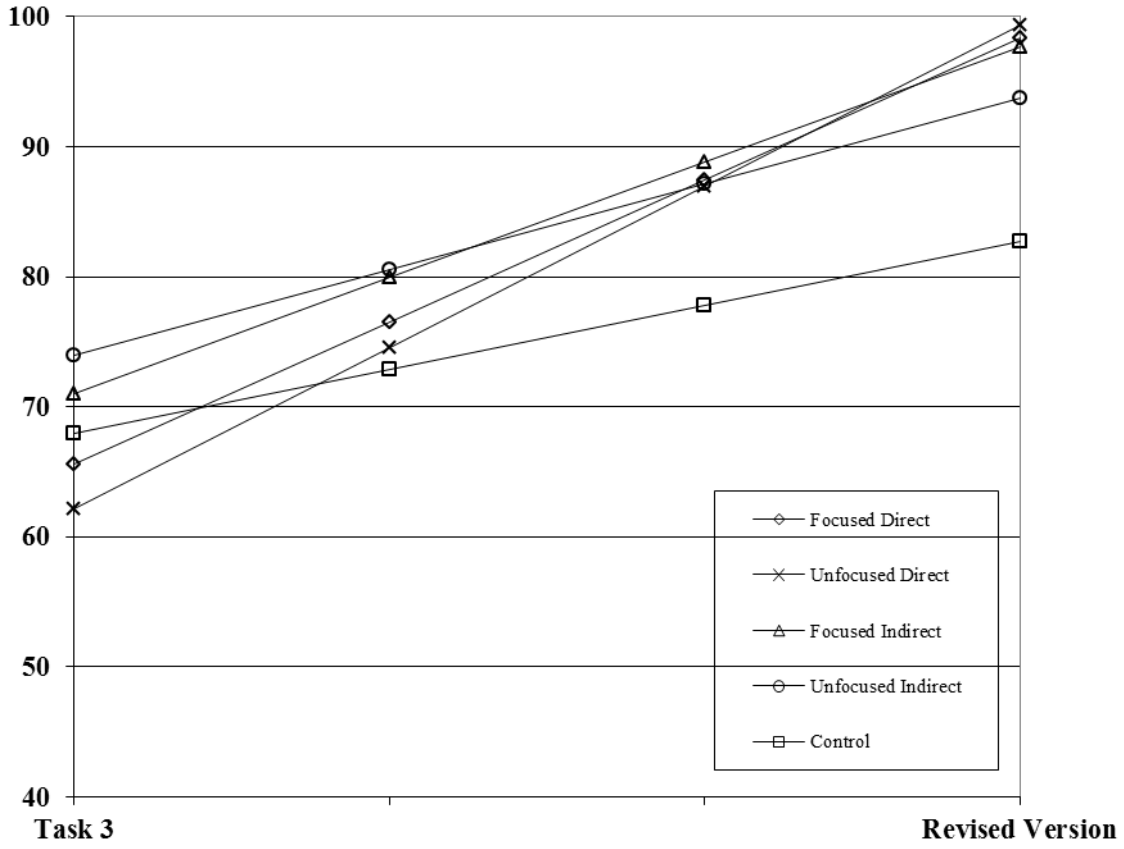
Table 34: Descriptive Statistics for the Effect of Corrective Feedback on a Revised Version of Task 3 for the Irregular Past Tense

Corrective Feedback	*n*	Task 3			Revised Version		
		M %	*SD*	Range %	*M* %	*SD*	Range %
Focused Direct	10	65.60	18.14	25.00 - 90.00	98.33	5.27	83.33 - 100.00
Unfocused Direct	9	62.15	20.62	25.00 - 87.50	99.31	2.08	93.75 - 100.00
Focused Indirect	9	71.03	28.10	0.00 – 88.89	97.69	4.71	87.50 - 100.00
Unfocused Indirect	7	73.92	17.09	40.00 – 90.00	93.72	11.15	72.73 - 100.00
Control	10	67.93	10.89	50.00 – 83.33	82.75	16.40	50.00 - 100.00

6.5.1 Effects of corrective feedback over time

Figure 11 graphically shows the mean scores for Task 3 and the revised version of Task 3. For each of the five groups, Wilcoxon Signed Rank tests were conducted. Significant differences across time were found for the Focused Direct Group (z = -2.67, p < .01), the Unfocused Direct Group (z = -2.67, p < .01), the Focused Indirect Group (z = -2.67, p < .01), the Unfocused Indirect Group (z = -2.20, p = .03) and the Control Group (z = -2.38, p = .02). The all improved over time.

Figure 11: Mean Irregular Past Tense Scores for the Effect of Corrective Feedback on a Revised Version of Task 3



6.5.2 Effects of different types of corrective feedback

To check for group differences in Task 3, a Kruskal-Wallis Test was undertaken. The result was not significant (χ^2 (4, 45) = 4.60, p = .33). Gain scores were generated nonetheless. Descriptive statistics for the gain score is presented in Table 35 and the mean gain scores in Figure 12.

Table 35: Descriptive Statistics for the Effect of Corrective Feedback on the Irregular Past Tense Gain Score

Corrective Feedback	*n*	Gain Score		
		M %	*SD*	Range %
Focused Direct	10	32.74	20.56	0.00 - 75.00
Unfocused Direct	9	37.93	22.51	6.25 - 75.00
Focused Indirect	9	26.35	28.20	11.11 - 100.00
Unfocused Indirect	7	19.81	21.78	-9.09 - 60.00
Control	10	13.82	14.64	-2.78 - 40.00

A Kruskal-Wallis Test was conducted on the gain scores. The result approached significance (χ^2 (4, 45) = 8.87, p = .07).[15] A series of post-hoc Mann-Whitney U Tests

[15] This score approached significance, so a one-way between groups ANOVA was conducted on the gain score. In this instance, there were no significant differences found between the groups (F (4, 45 = 1.84, p = .14). Despite there being no group differences, post-hoc analyses were undertaken for the sake of consistency between the regular and irregular past tense.

(one-tailed) were undertaken.[16] These required a Bonferroni adjustment for 10 comparisons (p = .005), and the results of these tests are presented in Table 36. The Focused Direct CF Group outperformed the Control Group (U = 22.00, z = -2.13, p = .02). The difference approached significance.[17] The Unfocused Direct CF Group also outperformed the Control Group (U = 14.50, z = -2.49, p = .005). In this instance, the difference reached significance. An additional analysis was undertaken to discern whether there were any group differences based on focus or directness. Kuskal-Wallis Tests were used to test for group differences between the focused groups, unfocused groups and the Control Group as well as the direct groups, indirect groups and the Control Group. The differences in focus were not significant (χ^2 (2, 45) = 4.68, p = .10) while those of directness were (χ^2 (2, 45) = 8.54, p = .01).

[16] Comparing these five groups, one-tailed Mann-Whitney U tests were undertaken as there is a clear justification in the literature for directional hypotheses. For instance, one could expect unfocused direct CF to outperform unfocused indirect CF.

[17] An additional independent samples t-test was undertaken as the difference between the groups that approached significance. The Focused Direct CF Group significantly outperformed the Control Group (t (18) = -2.37, p = .03).

Figure 12: Mean Irregular Past Tense Gain Scores for the Effect of Corrective Feedback on a Revised Version of Task 3

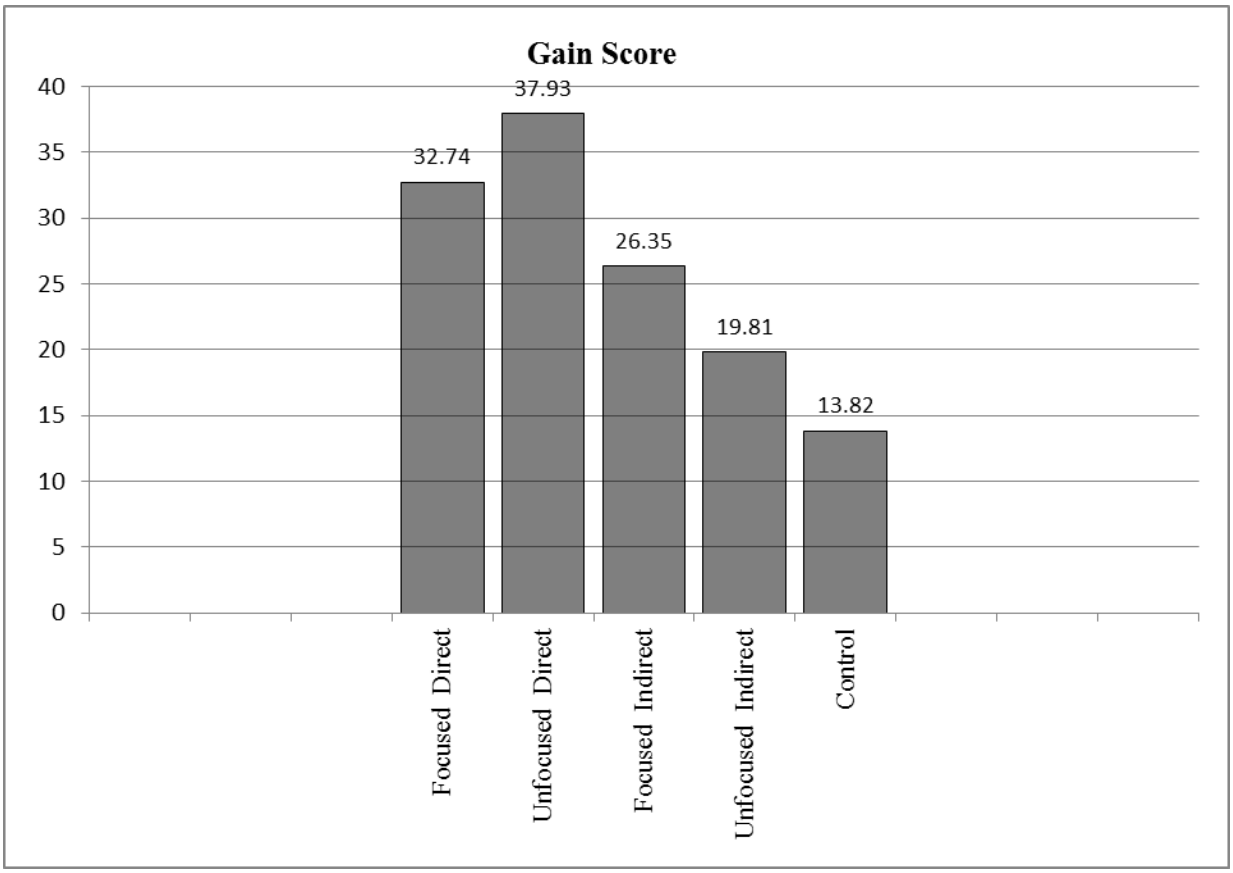


Table 36: Results from the Mann-Whitney U Tests on Gain Score

Group	Control	Unfocused Indirect	Focused Indirect	Unfocused Direct	Focused Direct
Focused Direct	*p* = .02†	*p* = .11	*p* = .10	*p* = .31	-
Unfocused Direct	*p* = .005*	*p* = .06	*p* = .08	-	-
Focused Indirect	*p* = .11	*p* = .27	-	-	-
Unfocused Indirect	*p* = .33	-	-	-	-
Control	-	-	-	-	-

* significant
† approaching significance
- not applicable

As with the analysis of the regular past tense, a series of one-way Mann-Whitney U Tests were undertaken. These required a Bonferroni adjustment for three comparisons ($p = .017$). The investigation of directness found the direct groups outperformed the Control Group ($U = 36.50$, $z = -2.70$, $p < .01$) and the indirect groups ($U = 91.50$, $z = -2.01$, $p = .02$). This latter difference only approached significance, however. There were no significant differences between the indirect groups and the Control Group ($U = 59.50$, $z = -1.08$, $p = .15$).

6.6 Summary of statistically significant results

The statistically significant results over time and between groups are summarized in this

section. Table 37 presents the statistically significant differences over time.

Table 37: Summary of Statistically Significant Results over Time

Structures	Corrective Feedback	Task 3 – Revised Version
Regular Past	Focused Direct	Yes ↑
	Unfocused Direct	Yes ↑
	Focused Indirect	Yes ↑
	Unfocused Indirect	Yes ↑
Irregular Past	Focused Direct	Yes ↑
	Unfocused Direct	Yes ↑
	Focused Indirect	Yes ↑
	Unfocused Indirect	Yes ↑
	Control	Yes ↑

↑ significantly increasing scores

For the regular past tense, differences in the gain scores approached significance. Post-hoc Mann-Whitney U-Tests (one-tailed) found the Focused Direct CF Group outperformed the Control Group and the Focused Indirect CF Group. The Unfocused Direct CF Group also produced more accurate revised versions than the Control Group and the Focused Indirect CF Group. Group differences in the gain scores for the irregular past tense were also demonstrated. Post-hoc Mann Whitney U Tests (one-tailed) found the Focused Direct CF Group outperformed the Control Group. This difference approached significance. The Unfocused Direct CF Group also demonstrated superior accuracy in a revised version of a text compared to the Control Group. In this instance, the difference reached significance.

Table 38 summarizes the statistically significant findings for the analyses of focus and directness using one-tailed Mann-Whitney U Tests.

Table 38: Summary of Statistically Significant Differences in Focus or Directness

Structures	Corrective Feedback	Gain Score
Regular Past	Direct	> Control > Indirect
Irregular Past	Direct	> Control > Indirect†

†approaching significance

6.7 Discussion

I will first discuss the findings of the analysis of the number of corrections received before I address the two research questions informing this chapter. In the process of discussing these questions, I will also compare the results with those in the literature on written CF.

6.7.1 Number of corrections received

A significant issue is whether there were any differences in the number of corrections in the four experimental groups. The Kruskal-Wallis Tests showed that there were no significant group differences for either the regular past tense or the irregular past tense. Differences in the amount of CF provided can therefore be largely eliminated as a cause for any group differences in gain scores.

Spearman Rank Order Correlations were also undertaken between the number of corrections received and gain scores for the regular past tense and the irregular past tense. There were significant positive correlations for the regular past tense (*rho* = .38, $p < .01$) and the irregular past tense (*rho* = .51, $p < .01$). It would seem that the number of corrections is important for improving the accuracy of a revised version of a text for both tenses.

6.7.2 The effects of corrective feedback on the learners' accurate use of past tense structures in a revised version of a text

Research question 3 asked whether CF had any effect on learners' accurate use of past tense structures in a revised version of a text. It was answered by comparing the scores from Task 3 to a revised version of Task 3 for both the regular past tense and the irregular past tense.

The Control Group failed to demonstrate greater accuracy in the use of regular past tense verb forms in a revised version of Task 3. Previous research has produced the same result (e.g. Fathman & Whalley, 1990; Shintani & Ellis, mimeograph). For the regular past tense, the lack of CF meant the learners were unable to repair their regular past tense errors.

However, in the case of irregular past tense verbs, the Control Group reached significance. A possible explanation for why the learners were able to notice errors in the irregular past tense may lie in the fact that the irregular past tense forms represented high frequency

lexical forms. This being the case, the learners were able to identify and repair the errors during the revision process or alternatively used different irregular verb forms in place of those they had made initial errors in.

All four CF groups significantly improved over time for both the regular past tense and the irregular past tense. This is not a surprising finding as all four CF groups had access to the CF when they revised their texts. That there was a significant improvement in accuracy in a revised version of a text has also been demonstrated for unfocused direct CF (Chandler, 2003) and unfocused indirect CF (Fathman & Whalley, 1990). These studies together with this one indicate that, irrespective of directness or focus, CF is effective in improving the accuracy of a revised version of a text.

In contrast, a recent study by Shintani and Ellis (mimeograph) failed to demonstrate that focused direct CF on system-learned articles led to improved accuracy over time. However, unlike previous research into the effect of CF on revision, the participants did not have access to their corrections or to their original uncorrected version. One could argue then that such a procedure was simply testing what they had learned or could remember from studying the CF. Also, in this study there was only a single instance of CF and, given the level of the learners (low-intermediate) this was insufficient for them to internalise the structures they had noticed.

It is also worth noting that all the CF groups achieved over the 90% threshold for both the

regular and irregular forms in the revised text. This threshold is considered to represent acquisition of a particular structure (Brown, 1973). However, as Truscott (1996) notes, demonstrating improvement in the accuracy of a revised version of a text does not necessarily mean that such improvement will translate into the acquisition of particular structures. In other words, it does not mean learners will be able to accurately produce these target structures in new pieces of writing. This is examined in the next chapter.

It is noteworthy that all the CF groups still made quite a lot of errors. One might have expected some errors in the indirect CF groups but perhaps not in the direct CF groups. In the direct groups all the participants had to do was copy the corrected past tense forms when they revised their text. To gain a better understanding of why learners were not successfully correcting errors in their revised texts, the types of responses to the CF were identified and amounts tallied for those participants that did not achieve 100% accuracy in the revised version of Task 3. The results are presented in Table 39. What is immediately evident is that all of the CF was attended to. In other words, none of the feedback was ignored. Interestingly, despite the provision of CF, a few learners made additional new errors in past forms of verbs. There were quite a few instances when there was no uptake. Of these, one can see that there were more instances with the indirect CF groups (i.e. 12) compared to the direct CF ones (only 2) for both structures.

Table 39: Types and Quantities of Incorrect Responses to the Corrective Feedback in the Revisions

Structure	Corrective Feedback	Responses		
		No Uptake	New Error	Feedback Ignored
		n	*n*	*n*
Regular Past	Focused Direct	1	2	0
	Unfocused Direct	1	1	0
	Focused Indirect	4	3	0
	Unfocused Indirect	2	1	0
Irregular Past	Focused Direct	0	1	0
	Unfocused Direct	0	1	0
	Focused Indirect	2	0	0
	Unfocused Indirect	4	0	0

6.7.3 Differences in the effects of the different types of corrective feedback on the learners' accurate use of past tense structures in a revised version of a text

Research question 4 asked whether there were any differences between the different types of CF on learners' accurate use of past tense structures in a revised version of a text. For both the regular and the irregular past tense, this was answered by comparing the different groups' gain scores generated from the scores in Task 3 and the revised version of Task 3. An analysis of group differences was also undertaken on the basis of focus and directness.

The major finding for both the regular past tense and irregular past tense is that the direct groups outperformed the indirect ones. One would expect that direct CF would outperform

indirect CF as the learners had access to the corrections when they revised Task 3 whereas those in the indirect CF groups did not. Those in the indirect CF groups had to draw on their existing linguistic knowledge to successfully correct their errors. It would appear that in some instances this knowledge was insufficient to successfully correct them.

A closer inspection of the findings for the regular past tense provides support for the differences between the direct and indirect CF. Both the Focused Direct CF Group and Unfocused Direct CF Group tended to outperform the Focused Indirect CF Group. It would appear that when revision is concerned it is not focus that is important but rather directness. This is due to the same reasons as discussed above, namely, for indirect CF learners have to know how to correct their errors and were not often successful in doing so.

Another finding of note is that the direct CF groups tended to outperform the control groups for both structures. This is not so surprising when one considers that the control groups had neither the corrections provided nor errors indicated. They could only draw on their existing knowledge to identify where an error had occurred and then correct it. The learners' knowledge then was not at a high enough level for them to both identify errors in the regular past and irregular past and correct them.

Support for the superiority of direct CF over no CF can be found by looking at the results for the group differences. Both the Focused Direct CF Group and Unfocused Direct CF Group, for the regular past tense, tended to produce more accurate revisions than the

Control Group. Analyses of the irregular past tense demonstrated the Focused Direct CF Group showed a tendency to outperform the Control Group and the Unfocused Direct CF Group actually did write more accurate revisions than the Control Group. Other studies (Van Beuningen, De Jong & Kuiken, 2012) have reported a similar finding (i.e. unfocused direct CF resulted in more effective revisions than both self-editing and no CF. Sachs and Polio (2007) also found that unfocused direct CF group assisted revision. Only Ellis and Shintani (mimeograph) reported no effect for CF on revision but, as already explained, that may have been because CF was only provided once and the students did not have access to their corrected text when revising.

6.8 Summary

This chapter investigated two research questions. The first asked whether CF had any effect on the accuracy of a revised version of a text. The answer to this question is that accuracy did improve as a response to the CF provided. Irrespective of the type of structure, all four CF groups' accuracy improved between Task 3 and the revised version of Task 3. This is to be expected as the learners had access to the corrections during the revision process. However, as Truscott (1996) points out, demonstrating that learners can write accurate drafts with the assistance of teachers' CF provides no evidence as to the effectiveness of such feedback for the acquisition of grammatical forms.

The second research question asked whether there were any differences in the effects of the CF groups on the accuracy of revised version of Task 3. For both structures, group

differences were found between the direct groups on the one hand and the indirect groups and control groups on the other. The provision of corrections during the revision process clearly results in more accurate revisions than indicating and locating errors or providing no CF at all.

CHAPTER SEVEN

THE EFFECT OF CORRECTIVE FEEDBACK PLUS REVISION ON NEW PIECES OF WRITING

7.1 Overview

Research questions 5 and 6 are addressed in this chapter.

RQ5. What effect did CF followed by revision have on the learners' accurate use of past tense structures in new pieces of writing?

To answer research question 5, I will explore the effect of CF plus revision over time. This will be achieved by examining whether there are any statistically significant differences in the five groups' scores by comparing the scores for Task 3 and Task 4, Task 3 and Task 5 and Task 4 and Task 5.

Task 3 represented the delayed post-test in the analysis of research questions 1 and 2; however, it was used as a pre-test for this analysis. Following the completion of Task 3, the learners' texts were corrected using the designated types of CF for the groups to which they belonged. The Control Group received no CF. The corrected text for Task 3 was made available to the learners after the CF session so they could complete a revised version of Task 3. The participants then completed Task 4 after finishing the revised version of Task 3. Task 4 then represents an immediate post-test. Task 5 was completed two weeks later and is used as a

delayed post-test. Tasks, 3, 4 and 5 all represented new pieces of writing.

RQ6. Were there any differences in the effects of the different types of CF followed by revision on the learners' accurate use of past tense structures in new pieces of writing?

I will answer this question by testing whether there are any significant differences between the five groups' scores. For both questions, I will report results for the regular past tense and the irregular past tense seperately, and the chapter will be organized as such. I will present a discussion of the results in the final section of the chapter.

7.2 Investigating the distribution of the past tense scores

A number of issues surrounding the distribution of the past tense scores need to be addressed. Task 3, which operated as a pre-test, was inspected for any scores above 90. Participants with such scores were removed from this part of the sample.[18] The remaining scores were checked to see if they had a normal distribution by undertaking Kolmorgorv-Smirnov tests on each of the five groups for each of the three tasks (see Appendix C). Some of the groups failed to demonstrate a normal distribution, so non-parametric rather than parametric tests were required. An analysis of the five groups' scores for Task 3 also showed that there were group

[18] As Task 3 was also used as a delayed post-test to answer research questions 1 and 2, it is obvious that those groups that performed well under the experimental conditions necessarily had fewer members available for the experimental treatments examined here. That is, a number of learners in these groups were deemed to have already acquired the target structures. For example, the number of participants in the Focused Direct CF Group was quite high for research questions 1 and 2 (n = 15); however, for the research questions investigated in this chapter, the number was more than halved (n = 7).

differences evident for the regular past tense but not for the irregular past tense. For the sake of consistency with the research questions that addressed group differences (i.e. RQ2, RQ4 and RQ6), gain scores were generated by calculating the difference between Task 3 and Task 4 (gain score 1) and between Task 3 and Task 5 (gain score 2).

7.3 Number of corrections received

The number of corrections received by the four experimental groups in CF session 2 was counted separately for the regular past tense and irregular past tense (see Tables 40 and 41). The number of regular past tense corrections was subject to a Kruskal-Wallis Tests to test for any group differences. There were no significant differences between the four groups for the regular past tense (χ^2 (3, 42) = 4.45, p = .22). A Kruskal-Wallis Test found no significant differences in the amount of corrections received by the irregular past tense experimental groups (χ^2 (3, 32) = 4.60, p = .18).

Table 40: Descriptive Statistics for Regular Past Tense Corrections Received during Corrective Feedback Session 2

Corrective Feedback	*n*	Number	*M*	*SD*	Range
Focused Direct	7	25.00	3.57	3.21	1.00 - 10.00
Unfocused Direct	9	24.00	2.67	1.11	1.00 - 4.00
Focused Indirect	15	29.00	1.93	0.80	1.00 - 4.00
Unfocused Indirect	11	20.00	1.82	0.98	1.00 - 4.00

Table 41: Descriptive Statistics for Irregular Past Tense Corrections Received during Corrective Feedback Session 2

Corrective Feedback	*n*	Number	*M*	*SD*	Range
Focused Direct	7	12.00	1.71	0.76	1.00 - 3.00
Unfocused Direct	9	25.00	2.78	1.56	1.00 - 6.00
Focused Indirect	9	25.00	2.78	1.20	1.00 - 4.00
Unfocused Indirect	7	13.00	1.86	1.07	1.00 - 4.00

Spearman Rank Order Correlations showed that, for the regular past tense, there were positive and significant correlations between the number of corrections received and both gain score 1 (*rho* = .43, $p < .01$) and gain score 2 (*rho* = .43, $p < .01$). The correlations for the irregular past tense were also positive and significant for both gain score 1 (*rho* = .45, $p = .01$) and gain score 2 (*rho* = .45, $p = .01$).

7.4 Results for the regular past tense

The regular past tense was examined by comparing the scores for the three tasks. Analyses were undertaken across time and between groups. Descriptive statistics are presented in Table 42.

Table 42: Descriptive Statistics for the Effect of Corrective Feedback plus Revision on Regular Past Scores

Corrective Feedback	*n*	Task 3			Task 4			Task 5		
		M %	*SD*	Range %	*M* %	*SD*	Range %	*M* %	*SD*	Range %
Focused Direct	7	62.11	12.85	55.56 - 90.00	88.22	6.90	80.00 - 100.00	88.63	17.54	60.00 - 100.00
Unfocused Direct	9	69.45	20.64	25.00 - 90.00	88.43	11.13	66.67 - 100.00	89.80	10.24	70.00 - 100.00
Focused Indirect	15	79.83	9.52	60.00 - 90.00	86.87	14.93	50.00 - 100.00	84.41	15.40	40.00 - 100.00
Unfocused Indirect	11	76.24	9.11	60.00 - 90.00	86.35	18.24	42.86 - 100.00	84.60	12.58	57.14 - 100.00
Control	9	65.00	11.56	44.44 - 80.00	77.57	21.19	37.50 - 100.00	63.45	24.30	25.00 - 93.33

7.4.1 Effects of corrective feedback over time

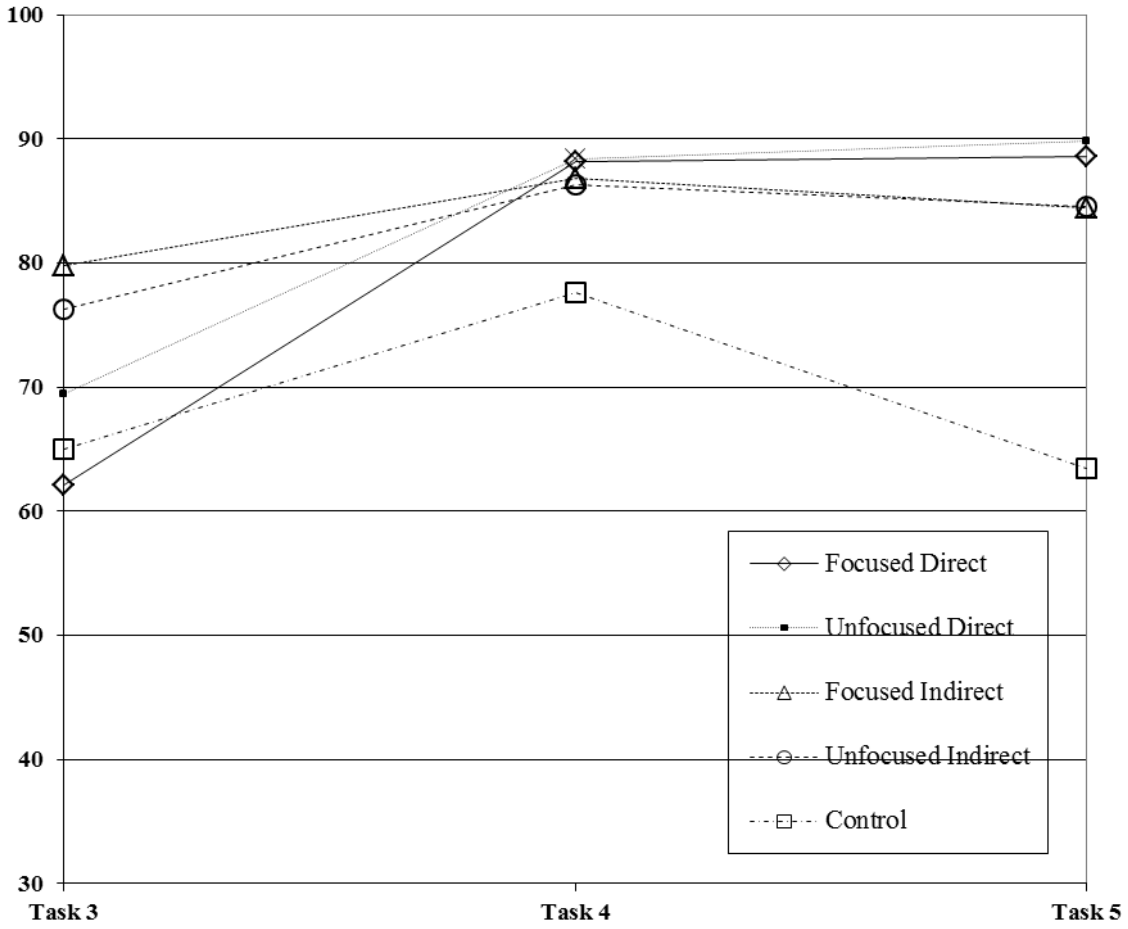
The mean scores for the five groups across the three tasks are presented in Figure 13. Individual Friedman tests found that there were significant changes over time for the Focused Direct CF Group (χ^2 (2, 7) = 11.19, $p < .01$) and the Unfocused Direct CF Group (χ^2 (2, 9) = 8.40, $p = .02$). There were no such changes for the Focused Indirect CF Group (χ^2 (2, 15) = 3.96, $p = .14$), the Unfocused Indirect CF Group (χ^2 (2, 11) = 3.12, $p = .21$) or the Control Group (χ^2 (2, 9) = 4.67, $p = .10$).

Post-hoc Wilcoxon Signed Rank Tests with a Bonferroni adjustment for three comparisons were conducted ($p = .017$). For the Focused Direct CF Group, the improvement was significant

between Task 3 and Task 4 (z = -2.37, p = .018) as well as from Task 3 to Task 5 (z = -2.37, p = .018). The Unfocused Direct CF Group approached significance between Task 3 and Task 4 (z = -2.07, p = .04).[19] It reached significance between Task 3 and Task 5 (z = -2.43, p = .015).

[19] As the changes over time only approached significance, an additional paired samples t-test was completed. The improvement between Task 3 and Task 4 was in this instance significant (t (9) = -2.68, p = .03).

Figure 13: Mean Regular Past Scores for the Effect of Corrective Feedback plus Revision on New Pieces of Writing



7.4.2 Effects of different types of corrective feedback

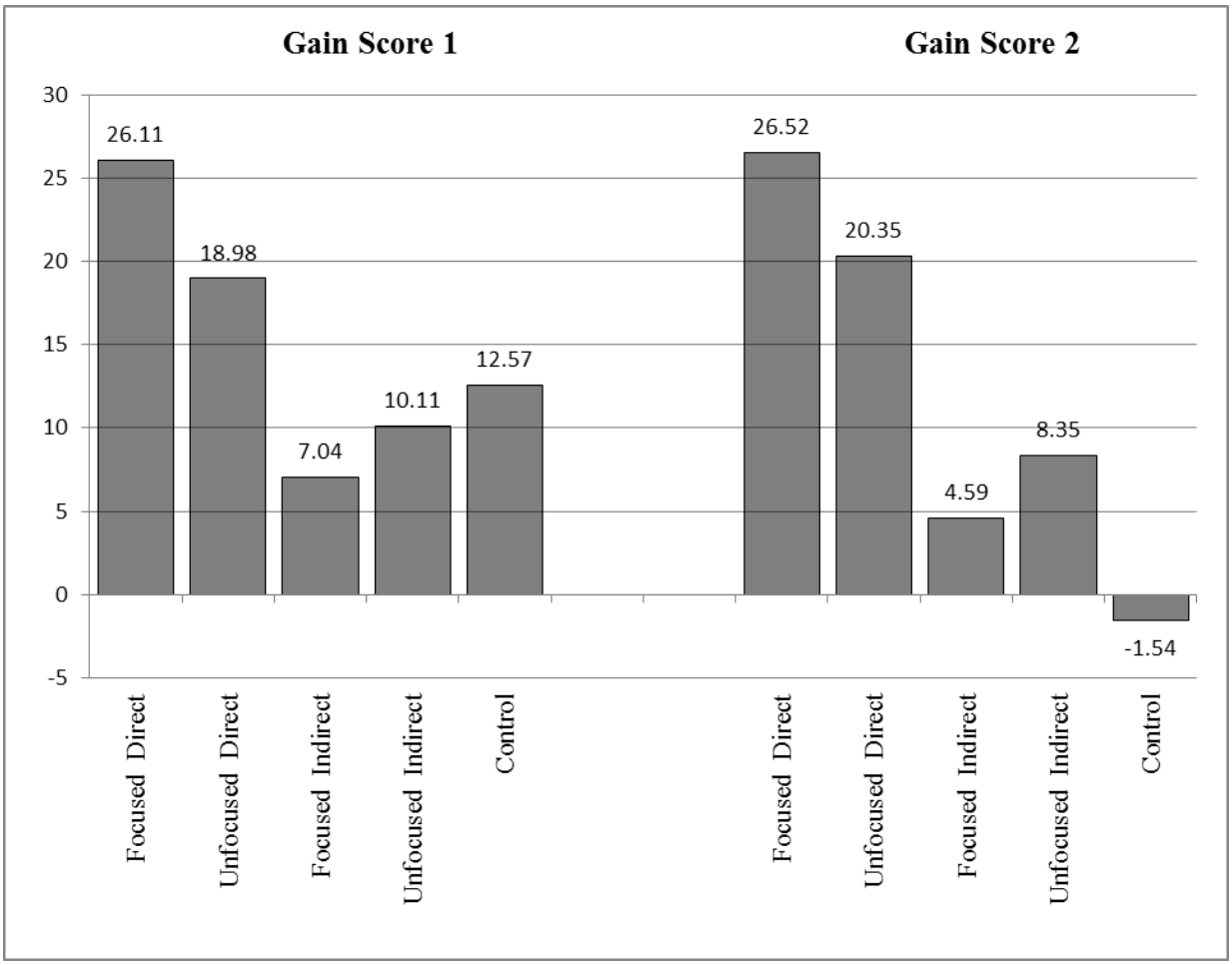
A Kruskal-Wallis Test applied to Task 3 found there were significant differences between the groups (χ^2 (4, 51) = 12.95, p = .01). Gain scores were generated. The descriptive statistics are presented in Table 43 and the mean gain scores in Figure 14.

Table 43: Descriptive Statistics for Regular Past Tense Gain Scores

Corrective Feedback	*n*	Gain Score 1			Gain Score 2		
		M %	*SD*	Range %	*M* %	*SD*	Range %
Focused Direct	7	26.11	14.83	2.22 – 49.17	26.52	11.14	4.44 – 41.67
Unfocused Direct	9	18.98	21.27	-13.33 - 53.57	20.35	20.07	-1.79 – 55.00
Focused Indirect	15	7.04	16.89	-37.50 - 40.00	4.59	19.24	-37.78 - 33.33
Unfocused Indirect	11	10.11	23.70	-37.14 – 40.00	8.35	13.07	-17.86 - 25.00
Control	9	12.57	28.09	-37.50 - 41.27	-1.54	27.56	-39.39 - 38.89

Kruskal-Wallis Tests were applied to the gain scores. For gain score 1, there were no significant differences between the groups (χ^2 (4, 51) = 5.97, p = .20); however, there were for gain score 2 (χ^2 (4, 51) = 10.49, p = .03). As the differences between the groups reached significance in gain score 2, a series of one-tailed Mann-Whitney U Tests were undertaken to

Figure 14: Mean Regular Past Tense Gain Scores for the Effect of CF plus Revision on New Pieces of Writing



ascertain where the differences lay.[20] These required a Bonferroni adjustment for 10 comparisons (p = .005). The results are presented in Table 44. It is evident that the Focused Direct CF Group outperformed the Control Group (U = 10.00, z = -2.28, p = .01), the Unfocused Direct CF Group had superior gain scores compared to both the Control Group (U =

[20] One-tailed Mann-Whitney U Tests were undertaken because there is justification in the literature that directional hypotheses can be made. For example, focused direct CF should lead to superior results compared to unfocused direct CF as should focused indirect CF compared to unfocused indirect CF.

22.50, z = -1.59, p = .06) and the Focused Indirect CF Group (U = 39.50, z = -1.67, p = .05). However, these differences only approached significance.[21] The Focused Direct CF Group significantly outperformed both the Focused Indirect CF Group (U = 17.00, z = -2.50, p = .005) and the Unfocused Indirect CF Group (U = 9.50, z = -2.63, p = .001).

Table 44: Results from the Mann-Whitney U Tests on Gain Score 2

Group	Control	Unfocused Indirect	Focused Indirect	Unfocused Direct	Focused Direct
Focused Direct	p = .01†	p = .003*	p = .005*	p = .13	-
Unfocused Direct	p = .06†	p = .15	p = .05†	-	-
Focused Indirect	p = .39	p = .31	-	-	-
Unfocused Indirect	p = .23	-	-	-	-
Control	-	-	-	-	-

* significant
† approaching significance
- not applicable

An additional analysis was undertaken to investigate whether there were any significant differences between the focused groups, unfocused groups and the Control Group as well as the

[21] The changes here approached significance, so additional analyses were conducted using independent samples t-tests. The Focused Direct CF Group significantly outperformed the Control Group (t (14) = -2.76, p = .02). The Unfocused Direct CF Group was more accurate than both the Control Group (t (16) = -1.93, p = .07) and the Focused Indirect CF Group (t (22) = -1.91, p = .07). These latter two t-tests approached significance only.

direct groups, the indirect groups and the Control Group. Kruskal-Wallis Tests found there were no significant differences in gain score 1 for focus (χ^2 (2, 51) = .39, p = .82) or directness (χ^2 (2, 51) = 4.41, p = .11). In gain score 2, there were also none for focus (χ^2 (2, 51) = 1.88, p = .39); however, the differences were significant for directness (χ^2 (2, 51) = 9.00, p = .01). Post-hoc one-tailed Mann-Whitney U Tests were required to establish where the differences resided.[22] These required a Bonferroni adjustment for three comparisons (p = .017). The direct groups significantly outperformed the Control Group (U = 32.50, z = -2.24, p = .01) and the indirect groups (U = 101.50, z = -2.76, p < .01). The indirect groups did not outperform the Control Group (U = 101.00, z = -.60, p = .28).

7.5 Results for the irregular past tense

The analysis of the irregular past involved comparing the scores for the five groups across the three tasks. Descriptive statistics are presented in Table 45.

[22] Directional hypotheses, in this instance, were warranted by the findings of previous studies (i.e. there is evidence that direct CF is superior to indirect CF as well as control groups). Therefore, one-tailed Mann-Whitney U Tests were employed.

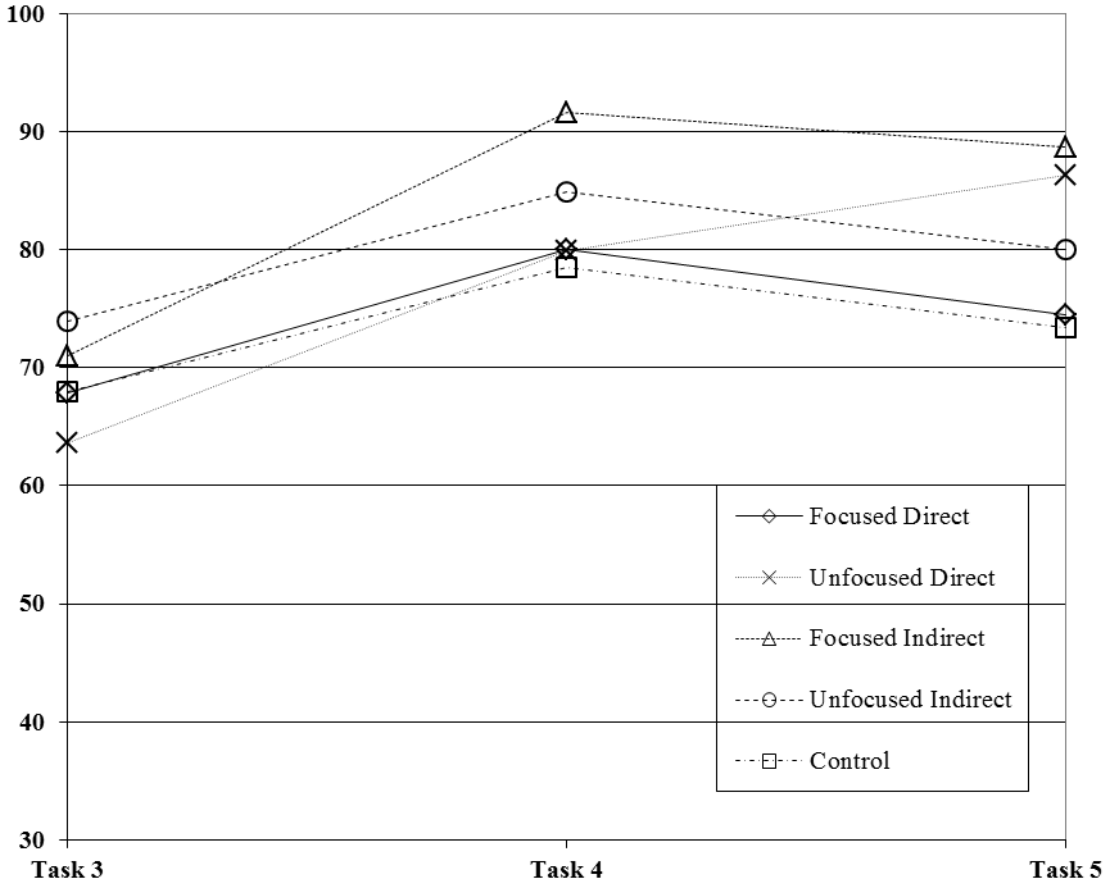
Table 45: Descriptive Statistics for the Effect of Corrective Feedback plus Revision on Irregular Past Scores

Corrective Feedback	*N*	Task 3			Task 4			Task 5		
		M %	*SD*	Range %	*M* %	*SD*	Range %	*M* %	*SD*	Range %
Focused Direct	7	67.86	21.53	25.00 - 90.00	80.05	18.45	50.00 - 100.00	74.44	17.13	50.00 - 100.00
Unfocused Direct	9	63.64	21.04	25.00 - 86.67	79.86	17.54	45.46 - 100.00	86.33	14.82	66.67 - 100.00
Focused Indirect	9	71.03	28.10	0.00 - 88.89	91.62	11.64	66.67 - 100.00	88.75	13.82	72.73 - 100.00
Unfocused Indirect	7	73.92	17.09	40.00 - 90.00	84.95	22.01	50.00 - 100.00	80.02	10.71	66.67 - 100.00
Control	10	67.93	10.89	50.00 - 83.33	78.50	19.41	40.00 - 100.00	73.39	20.27	42.86 - 100.00

7.5.1 Effects of corrective feedback over time

Figure 15 presents the mean irregular past tense scores over time. Separate Friedman Tests were undertaken on the scores for the five groups. There were no significant changes over time for any of the groups: the Focused Direct Group (χ^2 (2, 7) = .29, p = .87), the Unfocused Direct Group (χ^2 (2, 9) = 4.67, p = .10), the Focused Indirect Group (χ^2 (2, 9) = 4.06, p = .13), the Unfocused Indirect Group (χ^2 (2, 7) = 1.14, p = .57) and the Control Group (χ^2 (2, 10) = 2.39, p = .30).

Figure 15: Mean Irregular Past Scores for the Effect of Corrective Feedback and Revision on New Pieces of Writing



7.5.2 Effects of different types of corrective feedback

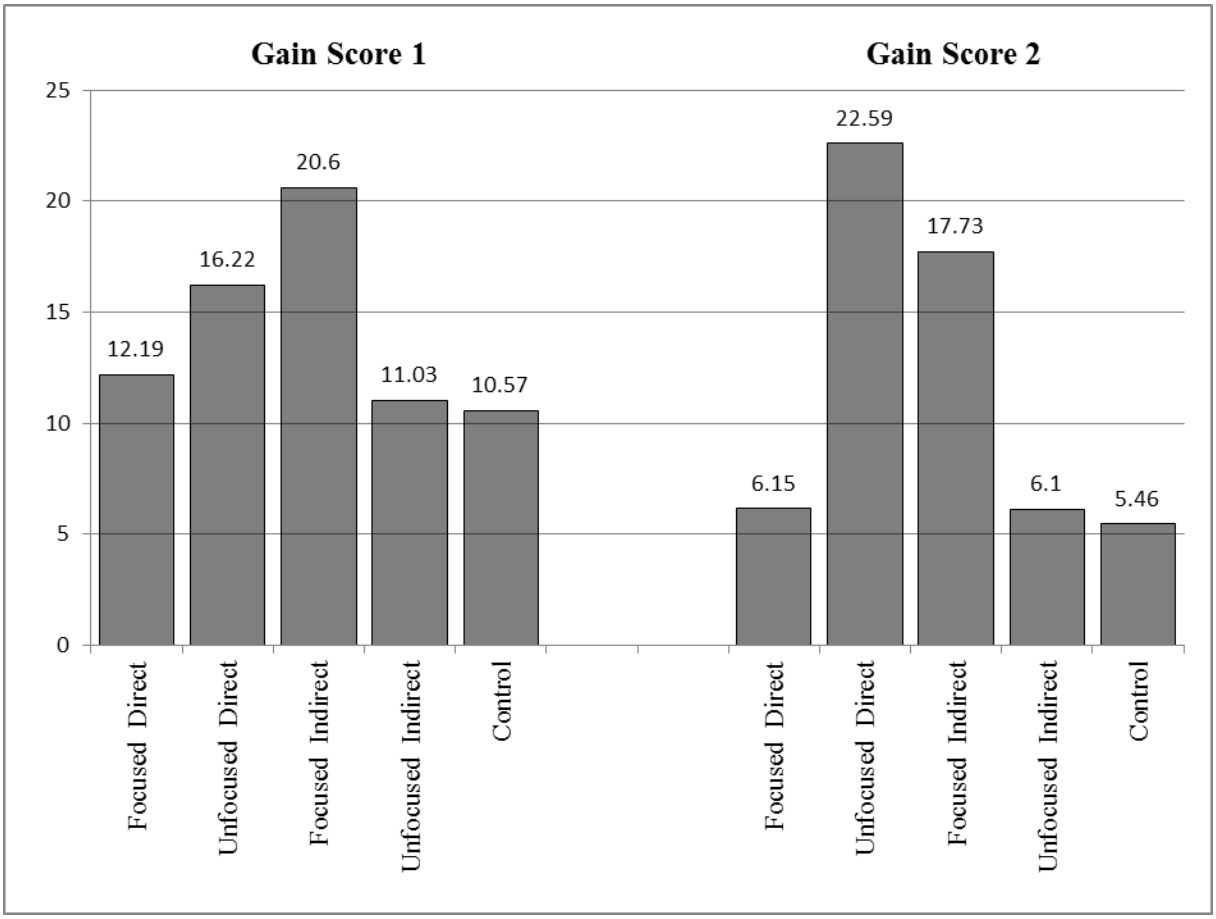
As applied to Task 3, a Kruskal-Wallis test found there were no significant differences between the groups (χ^2 (4, 45) = 1.96, p = .74). However, gain scores were generated. The descriptive statistics are presented in Table 46 and the mean gain scores in Figure 16.

Table 46: Descriptive Statistics for Irregular Past Tense Gain Scores

Corrective Feedback	*n*	Gain Score 1			Gain Score 2		
		M %	*SD*	Range %	*M* %	*SD*	Range %
Focused Direct	7	12.19	35.51	-33.33 - 63.89	6.15	35.87	-35.86 - 75.00
Unfocused Direct	9	16.22	21.64	-4.54 - 60.71	22.59	24.77	-13.33 - 75.00
Focused Indirect	9	20.60	26.94	-8.33 - 83.33	17.73	34.02	-18.89 - 100.00
Unfocused Indirect	7	11.03	30.57	-32.86 – 60.00	6.10	21.85	-23.33 – 33.33
Control	10	10.57	24.58	-43.33 – 38.46	5.46	17.73	-18.68 - 30.00

Kruskal-Wallis Tests conducted on the gain scores failed to find any significant differences between the groups in either gain score 1 (χ^2 (4, 42) = .12, p = .10) or gain score 2 (χ^2 (4, 42) = 3.57, p = .47).

Figure 16: Mean Irregular Past Tense Gain Scores for the Effect of Corrective Feedback plus Revision on New Pieces of Writing



An additional set of analyses were conducted to investigate whether there were any differences between the groups on the basis of focus or directness. As with the regular past tense, a comparison was made between the focused groups, unfocused groups and Control Group as well as the direct groups, indirect groups and Control Group. Kruskal-Wallis Tests found there were no significant differences in gain score 1 for focus (χ^2 (2, 42) = .07, p = .97) or directness

(χ^2 (2, 42) = .08, p = .96). Similarly, there were no significant differences in gain score 2 for focus (χ^2 (2, 42) = 1.13, p = .57) or directness (χ^2 (2, 42) = .72, p = .70).

7.6 Summary of statistically significant results

This section presents a summary of the statistically significant results over time and between groups. Table 6 summarizes those differences across time for the five groups and for the two structures.

Table 47: Summary of Statistically Significant Differences across Time

Structures	Corrective Feedback	Task 3 – Task 4	Task 3 – Task 5	Task 4 – Task 5
Regular Past	Focused Direct	Yes ↑	Yes ↑	No
	Unfocused Direct	Yes ↑†	Yes ↑	No

†approaching significance
↑ significantly increasing scores

Table 48 presents the statistically significant results between the five groups for the two structures using one-tailed Mann-Whitney U Tests.

Table 48: Summary of Statistically Significant Differences between Groups

Structures	Corrective Feedback	Gain Score 1	Gain Score 2
Regular Past	Focused Direct	None	> Control†
			> Focused Indirect
			> Unfocused Indirect
	Unfocused Direct	None	> Control†
			> Focused Indirect†

† approaching significance

Table 49 summarizes the significant findings for the comparisons based on focus and directness.

Table 49: Summary of Statistically Significant Differences in Focus or Directness

Structures	Corrective Feedback	Gain Score 1	Gain Score 2
Regular Past	Direct	None	> Control
			> Indirect

7.7 Discussion

The discussion will first consider the findings of the analysis of number of corrections received before addressing the two research questions addressed in this chapter. This will involve discussing the results for the regular past and irregular past separately. In so doing, I will compare the results with those in the literature.

7.7.1 Number of corrections received

It was important to ascertain whether the number of corrections received by each experimental group was not significantly different; otherwise, it could be argued that it was the quantity of CF not the CF itself that was responsible for any improvements evident over time or between separate groups. Kruskal-Wallis Tests conducted confirmed there were no such significant group differences for either the regular or irregular past tense.

Spearman Rank Order Correlations were also conducted on the combined number of corrections received for the experimental groups and gain scores for both tenses. For the regular past tense there were positive and significant correlations with both gain score 1 (*rho* = .43, $p < .01$) and gain score 2 (*rho* = .43, $p < .01$). The same was the case for the irregular past tense, that is, with gain score 1 (*rho* = .45, $p = .01$) and gain score 2 (*rho* = .45, $p = .01$). Thus, the greater the number of corrections received the better the gains for both past tense structures. In Chapter 5, however, there was no positive correlation for in gain score 2 for the irregular past tense. It would appear, on the face of it, that the additional opportunity to complete a revised version of a piece of writing is responsible for the positive correlation in gain score 2 presented here.

7.7.2 The regular past tense

7.7.2.1 The effects of corrective feedback followed by revision on learners' accurate use of the regular past tense in new pieces of writing

What is at first evident is the Control Group failed to significantly improve across time. That is,

writing practice with an opportunity to revise did not affect the accuracy of learners' writing over time. Writing practice with the opportunity to revise can be considered a type of pushed output (Swain, 1995). While research using think-aloud-protocols has identified that learners do attend to issues of accuracy when they revise (Cumming, 1990; Swain & Lapkin, 1995), the findings presented here suggest that a single episode of revision is not sufficient to have an effect on even a system-learned grammatical feature. This finding then supports Krashen's (1998) claim that there is no real evidence to support the effectiveness of this type of pushed output for the acquisition of linguistic structures. It is possible, of course, that multiple revisions of the same paper would yield different results. However, if CF followed by revision can be shown to enhance the accuracy of students' writing, then clearly this is preferable to relying on just writing practice with revision.

All of the CF groups improved between Task 3 and Task 4; however, only the two direct groups approached or reached significance. There were no significant changes over time for the two indirect CF groups. Overall then the results show that there were short-term improvements for CF followed by revision but only for the direct groups. Chandler (2003) was able to demonstrate significant short-term improvements for two of her groups - an unfocused direct CF group and an unfocused indirect CF one. Why then did the indirect ones in this study fail to significantly improve in the short-term? There are a number of possible explanations why the findings of Chandler's study and this one differ. One factor may have been the structures investigated. Chandler's study provided corrections on 16 different types of error of which the majority were not grammatical whereas the part of the study presented here investigated only

grammatical features. Another consideration may be the influence of the context. Chandler's study was conducted in a college-level ESOL environment whereas this study was completed in a college-level EFL one. Chandler also provided additional direct CF on the errors that were not initially revised correctly. A final consideration is that the two studies used different measures of accuracy. Chandler used error rate, calculated as total number of errors divided by the total number of words times 100. Target-like Use Analysis was used in this study. Parallels are difficult to draw between studies that use different measures for calculating accuracy.

Between Task 3 and Task 5, the results varied. Both the Focused Direct CF Group and Unfocused Direct CF Group reached significance for the regular past tense while all the remaining groups did not. In fact, these two groups continued to demonstrate improvement between Task 4 and Task 5 whereas the two indirect CF groups' accuracy deteriorated. None of the changes between Tasks 4 and 5 reached significance. As few studies to date have provided a long-term measure of the effectiveness of CF followed by revision, these results are noteworthy.

There would appear to be differences over time dependent on directness. This is of particular interest when one considers Truscott's (1996) claim about CF and revision. He noted that showing CF is effective when revising does not mean that it will be effective in subsequent pieces of writing. The findings suggest here that he is partially correct. All of the CF groups demonstrated significant improvements between Task 3 and the revised version of Task 3; however, both indirect CF groups failed to significantly improve over time across Task 3, Task

4 and Task 5. It would appear, then, that in the case of indirect CF followed by revision, Truscott is correct. However, the two direct CF groups either approached or reached significance over the three tasks. Direct CF plus revision was effective. This runs counter to Truscott's claims.

A recent study has, on the other hand, presented results that contrast with the findings for the Focused Direct CF Group. Shintani and Ellis (mimeograph) investigated CF directed at the indefinite article and found a focused direct CF group given first an opportunity to study their errors and then incorporate them into a revised version of a text failed to improve in either the short- or the long-term. In fact, in their study the focused direct CF group's accuracy deteriorated. Why then do the results between these two studies differ? One can perhaps consider the Taiwanese EFL context in which this study was conducted and the American ESOL one where Shintani and Ellis' one was completed. In Taiwan, the teaching of grammar is emphasised at all levels of education and, as a result, learners in this context may have well-developed metalinguistic knowledge to draw on whereas the learners in Shintani and Ellis' study may not have. Shintani and Ellis provide evidence that would appear to support this claim. The findings of a stimulated recall showed that these learners were unable to form or draw on a metalinguistic rule for the function of the article investigated. An alternative explanation may lie in the difference of the structures investigated in this study and Shintani and Ellis' research. Whereas the regular past tense has a single form-function relationship requiring the addition of the regular past tense –ed morpheme (a morphological change), the indefinite article system has multiple functions for a single form.

7.7.2.2 Differences in the effects of the different types of corrective feedback followed by revision on learners' accurate use of the regular past tense in new pieces of writing

The major finding was that the direct CF groups outperformed the indirect CF groups in the long-term. Explanations for why these results occurred can be found through a closer inspection of the findings.

The first explanation relates to the additional input that CF affords and the output that revision allows - output which varies in function according to the type of CF. The provision of direct corrections meant the learners had the opportunity to notice (Schmidt, 1990) and maybe notice with metalinguistic understanding (Schmidt, 2001). Revision following such feedback comprises a mechanical process of copying the corrections. It does not represent pushed output as defined in the Output Hypothesis (Swain, 1995), so more than likely it was just the opportunity for the learners to notice their errors and the correction that was responsible for the superior performance of the direct CF groups. That is, the learners had two opportunities to notice the errors and corrections - one during the CF session and one when they were asked to revise. Indirect CF only provides opportunities for the learners to notice they have made an error; it does not supply them with the correct form. Clearly, if they are not capable of correcting their errors using their own linguistic resources, they are not able to benefit from the correction. The fact that learners believed they were primarily engaged in practicing writing (see results of the exit questionnaire in Appendix D) may also have lead to the students in the indirect groups paying only limited attention to the linguistic corrections.

A further possible reason for the failure of the indirect CF plus revision to have any effect on acquisition of the regular past tense is the starting levels for the indirect groups may have been approaching a ceiling (i.e. they were notably higher than the starting levels of the direct CF groups).

There is then evidence that the effect of CF followed by revision differed depending on whether the CF was direct or indirect. It is not known, however, to what extent it was the additional input that direct CF affords or the additional input that revision allows that was responsible for these changes. To address this issue, there would have had to have been groups that just had direct CF without any opportunities to revise. This was the case in the results reported in Chapter 5. These results showed only that the focused CF groups outperformed the unfocused ones. In this chapter, as discussed above, the two direct groups outperformed the indirect ones. Thus it would appear that the effectiveness of direct CF is enhanced if learners also have the opportunity to revise.

A further finding was that the combined direct CF groups outperformed the control group in the long-term. This is perhaps not surprising as both the focused direct CF group and the unfocused direct CF group approached significance at this time, compared to the self-editing control group. However, Van Beuningen, et al. (2012) reported both short-term and long-term superiority in grammatical accuracy for unfocused direct CF compared to a writing practice control group but not a self-editing control group. What can explain these different findings? One possibility is the way the target structures were measured. Van Beuningen, et al. (2012) used an error ratio while this study used Target-Like-Use Analysis. Van Beuningen, et al. also

investigated eight identifiable error categories while this study only examined one - the regular past tense.

7.7.3 Irregular past tense

The results for research questions 5 and 6 showed that there were no statistically significant differences over time and also no differences between the groups. On the whole, the gain scores tended to worsen in the long-term - the one exception being the Unfocused Direct CF Group. Thus, it would appear that there were no significant effects for CF plus revision on the accuracy of irregular past tense forms in new pieces of writing. However, to better understand the effect of CF plus revision on the accuracy of irregular verbs in new pieces of writing, it was necessary to look at the specific irregular verbs that were corrected in Task 3 and the learners' subsequent use of these in Tasks 4 and 5.

A post-hoc analysis of the four CF groups was undertaken to see whether the verbs that were corrected in Task 3 were not used, used correctly or used incorrectly in the revised version of Task 3, Task 4 and Task 5 (see Appendix F). Table 50 shows the frequency of correct and incorrect verbs in Task 4 and Task 5 for those irregular verbs that were corrected in Task 3. It is evident that although the learners were able to correct the irregular verb errors in their revised texts, they were not able to use the corrected forms in the new pieces of writing. Less than 25% of the corrected verbs were actually used correctly in the subsequent pieces of writing.

Table 50: Correct and Incorrect Use of Those Irregular Verb Forms Corrected in Task 3

Corrective Feedback	Revised Version		Task 4		Task 5	
	Correct	Incorrect	Correct	Incorrect	Correct	Incorrect
Focused Direct	11	0	1	10	1	10
Unfocused Direct	18	0	1	17	0	10
Focused Indirect	10	3	1	13	2	12
Unfocused Indirect	9	2	2	9	0	11

To test whether there were any significant differences between the four CF groups in the amount of correct and incorrect use of the irregular past tense verbs in Task 4 and Task 5, a series of Fisher Exact Tests were used to compare the frequency of correct and incorrect uses in these two tasks. Six comparisons were undertaken for each of the two tasks. The results are presented in Table 51 and Table 52. It is clear that there were no statistically significant differences in the frequency of correct and incorrect use of the irregular past tense in either of the two tasks.

Table 51: Results from the Fisher Exact Tests on Task 4

Group	Unfocused Indirect	Focused Indirect	Unfocused Direct	Focused Direct
Focused Direct	p = 1.0	p = 1.0	p = 1.0	-
Unfocused Direct	p = .54	p = 1.0	-	-
Focused Indirect	p = .57	-	-	-
Unfocused Indirect	-	-	-	-

Table 52: Results from the Fisher Exact Tests on Task 5

Group	Unfocused Indirect	Focused Indirect	Unfocused Direct	Focused Direct
Focused Direct	p = 1.0	p = 1.0	p = .38	-
Unfocused Direct	p = N/A	p = .18	-	-
Focused Indirect	p = .49	-	-	-
Unfocused Indirect	-	-	-	-

N/A No Fisher Exact Test generated as both groups had constant variables (i.e. all incorrect)

In short, for a structure like irregular verbs there does not appear to be any grounds for expecting corrective feedback followed by revision to have any noticeable impact on learning. This supports Ferris' (1999) comment regarding the relative effects of CF on treatable and

untreatable grammatical features. Irregular verbs are untreatable.

7.8 Summary

Research question 5 asked whether CF followed by revision had any effect on the accurate use of the past tense structures in new pieces of writing. The writing practice only group (i.e. the Control Group) did not improve over time in either past tense structure even though they were given the opportunity to revise the writing they produced for Task 3. For the regular past tense, the two direct CF groups' accuracy improved significantly (or near significantly) between Task 3 and Task 4 and also from Task 3 to Task 5. In fact, their accuracy continued to improve from Task 4 to Task 5 although these improvements did not reach significance. The two indirect CF groups, on the other hand, failed to demonstrate any significant changes over time. The results for the irregular past tense show there were no significant changes in the accuracy of the learners' writing across the three tasks. This is supported by the findings of the post-hoc analysis which found that for all four CF groups, less than 25% of irregular past tense errors that were corrected were subsequently used correctly in the new pieces of writing even though they did use these verbs correctly in their revised text.

Thus, as the results for the previous research questions have shown, CF appears to have a different effect depending on the target structure. That is, there is evidence that direct CF in combination with the opportunity to revise has an effect on accuracy in subsequent writing in the case a rule-based feature (regular past tense) but no effect in the case of an item-based feature (irregular past tense). These findings then both support and refute Truscott's (1996)

assertion that CF may be effective in revisions but is not in new pieces of writing. CF assists revision for both rule-based and item-based features. However, CF plus revision is only effective in new pieces of writing if (1) the CF is direct and (2) the target feature is 'treatable'.

CHAPTER EIGHT

CONCLUSION

8.1 Aims of the study

Building on Ellis et al. (2008), the aims of this study were to investigate the effectiveness of written CF, with and without opportunities to revise, on the acquisition of two past tense structures - the regular past tense and the irregular past tense. CF was, furthermore, distinguished in terms of focus and directness. From a theoretical perspective, this allowed for an examination of the major claims made by Truscott (1996) and Ferris (1999). Pedagogically, the study aimed to investigate which type or types of CF were the most effective in the context where the study was undertaken and whether requiring learners to revise contributed to this effectiveness in any way.

8.2 Design of the study

The sample comprised freshman and sophomore English language majors at a junior college in Taipei, Taiwan. Intact classes were assigned to one of five groups: a Focused Direct CF Group, an Unfocused Direct CF Group, a Focused Indirect CF Group, an Unfocused Indirect CF Group and a Control Group. Using a quasi-experimental and step-up design, data were collected over a seven week period using various instruments. These included a background questionnaire, five writing tasks (Task 1, Task 2, Task 3, Task 4 and Task 5), a revision task and an exit questionnaire. The five writing tasks were counterbalanced to account for the possibility of varying degrees of task difficulty.

In week one, all participants completed a background questionnaire. To answer research questions 1 and 2, Task 1 was completed by all five groups. This task served as a pre-test. One week later, the participants in the four CF treatment groups had their writing returned with all their errors in the use of past tense corrected in accordance with their allocated type of CF. The learners in these groups were required to attend to the CF by studying it for five minutes. The Control Group did not have their writing returned to them. All five groups next completed Task 2, which operated as an immediate post-test. Two weeks later the learners in all five groups completed Task 3. This served as the delayed post-test for the first two research questions. Task 3 also represented the pre-test for research questions 4 and 5. One week after the completion of Task 3, the four treatment groups received CF on Task 3. They were required to study it for five minutes and then completed the revision task. The revision task requested the learners to write an improved draft of Task 3. These learners had access to their corrections while completing the revision task. The Control Group also completed Task 3 and the revision task. They were not however required to study their writing for five minutes. In answering research questions 3 and 4, the revision task represented a post-test. Task 3 also represented the pre-test for research questions 5 and 6. The learners in the four CF groups had their writing returned to them one week after the completion of Task 3 and duly corrected in accordance with their allocated type of CF. They were required to study their feedback for five minutes, and then asked to complete the revision task. The Control Group had Task 3 returned to them. It had no CF and the learners in this group were not required to study their writing for five minutes and were only required to complete the revision task. Following the completion of the revision task, all five groups completed Task 4. This operated as an immediate post-test. Two weeks later all

five groups completed Task 5 (delayed post-test). Following the completion of Task 5, all participants completed the exit questionnaire.

The five tasks were scored using Pica's (1983) Target-Like-Use Analysis. This is a type of obligatory occasion analysis that takes into account overgeneralization of one type of structure onto another (e.g. the overuse of the regular past tense -ed morpheme with irregular verbs). Intra-rater reliability was obtained by the researcher scoring a sample of the Task 1 a second time two months after the initial scoring of the data. The distribution of these scores were then examined and it was decided to use non-parametric rather than parametric tests. The number of corrections received by individual learners was also tallied. Using SPSS version 17, a series of statistical procedures was used to answer the research questions. A multiple choice question and an open-ended question was used for the exit questionnaire.

8.3 Summary of the main findings

8.3.1 The exit questionnaire

The exit questionnaire found that the participants as a whole considered the purpose of the writing tasks was to develop writing skills (85%), and they considered that they had practiced writing skills as a result of completing the tasks (60%). These findings suggest then that the focus of the study was not evident to the learners. That is, for the sample as a whole, it would appear that by and large the learners were unaware that the primary focus was on past tense verbs.

8.3.2 Research question 1

What effect did CF have on the learners' accurate use of past tense structures in new pieces of writing?

8.3.2.1 The regular past tense

Three episodes of writing practice without CF failed to demonstrate any improvement in accuracy over time. In contrast, all the CF groups demonstrated short-term improvement in accuracy. However, only the Focused Direct CF Group showed a pattern of increased accuracy from the short to the long-term while all the other groups tended to show a decrease in accuracy. Requiring learners to attend to a single episode of focused direct CF led to sustained improvement in accuracy.

8.3.2.2 The irregular past tense

There was a short-term improvement for the two direct CF groups. This accuracy deteriorated from the short- to the long-term. There was no improvement in accuracy for the two indirect CF groups or the writing practice group.

8.3.3 Research question 2

Were there any differences in the effects of the different types of CF on the learners' accurate use of past tense structures in new pieces of writing?

8.3.3.1 The regular past tense

There were group differences in the long-term on the basis of focus but not of directness. In particular, the Focused Direct CF Group outperformed the Unfocused Direct CF Group. There were also positive correlations between the number of corrections received and gain scores 1 and 2, suggesting that the number of corrections received contributed to accuracy. However, it should be noted that groups that had higher initial past tense scores (e.g. the Unfocused Direct CF Group) received fewer corrections and this may explain why the CF they received was not found to be effective.

8.3.3.2 The irregular past tense

There were no group differences in the initial analysis of the gains scores using ANOVAs, but there was a positive correlation between the number of corrections received and gain score 1. The effect of CF on those specific irregular forms initially produced incorrectly and then used in subsequent pieces of writing was investigated. Post-hoc analyses found that less than 50% of the irregular verbs that had received correction were subsequently used correctly. Additional post-hoc analyses found there were group differences in the long-term. The Focused Direct CF Group outperformed the Unfocused Indirect CF Group. It would appear that when CF is both direct and focused, it assists the acquisition of irregular past tense structures.

8.3.4 Research question 3

> What effect did CF have on the learners' accurate use of past tense structures in a revised version of a text?

Learners in the Control Group were unable to improve in accuracy between Task 3 and the revised version of Task 3 for the regular past tense. However, for the irregular past tense, the learners' accuracy improved from Task 3 to the revised version, but this may have been because they used different irregular verbs in their revised text. All of the CF groups demonstrated greater accuracy in the revised version of Task 3 for both tenses.

8.3.5 Research question 4

> Were there any differences in the effects of the different types of CF on the learners' accurate use of past tense structures in a revised version of a text?

For both the regular past tense and the irregular past tense, the direct groups tended to outperform the indirect ones as well as the control groups. The Focused Direct CF Group outperformed the Focused Indirect CF Group in the regular past tense and the Focused Direct CF Group and Unfocused Direct CF Group tended to outperform the Control Group in the irregular past tense. There were also significant, positive correlations between the number of corrections received and the gain scores for the regular past tense and the irregular past tense.

8.3.6 Research question 5

> What effect did CF followed by revision have on the learners' accurate use of past tense structures in new pieces of writing?

8.3.6.1 The regular past tense

Writing practice with an opportunity to revise failed to result in improvement in the learners' written accuracy. That is, writing practice with an opportunity to self-edit was not successful. Only the Focused Direct CF Group and the Unfocused Direct CF Group improved in the short-term for the regular past tense. This improvement was maintained in the long-term. In fact these groups continued to improve over time. The indirect CF groups failed to show any short-term or long-term improvement in accuracy. Their accuracy actually deteriorated from the short- to the long-term.

8.3.6.2 The irregular past tense

The results suggest that there was no improvement over time for any of the five groups.

8.3.7 Research question 6

> Were there any differences in the effects of the different types of CF followed by revision on the learners' accurate use of past tense structures in new pieces of writing?

8.3.7.1 The regular past tense

The direct CF groups outperformed the indirect CF groups. The Focused Direct CF Group tended to outperform the Focused Indirect CF Group and the Unfocused Indirect CF Group. There were positive correlations between the number of corrections received and both short-term and long-term gain scores.

8.3.7.2 The irregular past tense

Again, as the irregular past tense is item-learned, it was necessary to examine the effect of CF plus revision on incorrectly produced verb forms that were subsequently used in new pieces of writing. Less than 25% of irregular verbs that received correction were used in new pieces of writing. A further post-hoc analysis of correct and incorrect verb forms failed to find any group differences in either the short-term or the long-term.

8.4 Theoretical implications

8.4.1 Truscott's arguments against written corrective feedback

Truscott (1996) presented a number of arguments against the value of CF. He suggested that writing practice alone can lead to improved accuracy in writing, but CF has no effect on accuracy in new pieces of writing. He argued that while CF can assist in improving the accuracy of a revised version of a text, such an effect is not transferred to new pieces of writing. He also claimed that CF has a negative effect on fluency and induces avoidance. Each of these claims will be discussed separately.

8.4.1.1 Writing practice alone can lead to improved accuracy

The claim that writing practice alone can lead to improved accuracy is clearly refuted by the findings. The results for research question 1 show there was no improvement in accuracy over time in the Control Group for both the regular past tense and the irregular past tense. The Control Group was only required to complete three writing tasks. As mentioned previously, much of Truscott's theoretical position is informed by Krashen's (1985) Input Hypothesis;

however, his assertion as to the effectiveness of writing practice alone also draws on Swain's (1985) Output Hypothesis - originally proposed as a complement to the Input Hypothesis. There is no evidence then that three episodes of writing provided sufficient opportunities for pushed output to be successful. In other words, writing practice alone failed to draw learners' attention to gaps in their knowledge and so to enable them to move from semantic processing to syntactic processing.

8.4.1.2 Written corrective feedback has no effect on grammatical accuracy in new pieces of writing

The findings of research question 1 for the regular past tense fail to support the assertion that CF has no effect on grammatical accuracy in new pieces of writing. There was a short-term improvement for all four CF groups. As the learners were required to study their feedback for five minutes, this would appear to suggest that the learners were able to store a representation of the structure in their short-term memory. That is, the CF facilitated noticing-the-gap and the surface features of the structures were noticed (Schmidt, 1990). This was also demonstrated by Ellis et al. (2008), where the Focused Direct CF Group continued to show improvement in accuracy from the short- to the long-term. Learners in this group were perhaps able to draw on an explicit representation of the rule for the regular past tense as they had no additional input by which to again notice the structure in question. That is to say there is a case to be made for correction helping to develop metalinguistic understanding and thus to promote long-term learning.

The findings for research question 2 further support the effectiveness of CF. It was found that the combined focused CF groups tended to outperform the control group in the long-term as did the Focused Direct CF Group. This research then contributes to the growing body of research that has shown the long-term benefits of focused direct CF (Ellis, et al., 2008, Bitchener & Knoch, 2008; Sheen et al., 2009; Sheen, 2010, Bitchener & Knoch, 2010a)

8.4.1.3 Written corrective feedback can assist in the accuracy of a revised version of a text

The findings for research question 3 provide evidence in support of Truscott's position that CF can assist in improving the accuracy of a revised version of a text. Indeed, previous research has shown this is the case (e.g. Fathman & Whalley, 1990, Chandler, 2003). All four CF groups in this study improved in accuracy between Task 3 and the revised version of Task 3 for both the regular and irregular past tense. The learners were first required to study the feedback for five minutes and had access to the feedback when completing the revised version. When learners are required to complete a revised version of a text, the CF can be viewed as output-prompting CF, leading to all the CF groups demonstrating uptake of the structures receiving CF.

8.4.1.4 The effect of written corrective feedback on a revised version of a text is not transferred to new pieces of writing

The argument that any effect demonstrated on a revised version of a text is not transferrable to new pieces of writing is both rejected and supported. Research question 3 shows that the two direct CF groups improved in accuracy between Task 3 and a revised version of Task 3 for the

regular past tense. As demonstrated in research question 5, these same two groups improved in accuracy in the short-term, and they also both continued to show improvement from the short-term to the long-term. Here then is evidence refuting Truscott's claim. The two indirect CF groups also improved in accuracy between Task 3 and a revised version of Task 3 for the regular past tense (research question 3); however, there was no improvement in the accuracy of new pieces of writing for the same two groups (research question 5). Here there is some evidence supporting Truscott's argument.

Additional evidence in support of the effectiveness of direct CF can be effective when revising a text and in subsequent new pieces of writing can be found in the group comparisons of the CF groups and a control group (self-editing group). The results for research question 4 demonstrated that the combined direct CF groups outperformed the control group in a revised version of a text. The results for research question 6 showed the same pattern in the long-term for new pieces of writing. Both the focused and unfocused direct CF groups outperformed the self-editing control group in the long term although the difference did not reach statistical significance. These findings contrast with those in the literature. Shintani and Ellis (mimeograph) failed to find long-term differences between a focused direct CF group and a control group (self-editing). Van Beuningen et al. (2012) failed to find any differences between an unfocused direct CF group and a self-editing control group. However, both these studies involved heterogeneous subjects in ESOL contexts in contrast to the homogenous subjects in the EFL context investigated in this study.

An interesting question is why was the effect on a revised version of a text transferred onto new pieces of writing in some instances and not others? This is explicable in terms of the cognitive process prompted by the different types of CF when learners are first required to study the CF and then incorporate the corrections into a revised version of a text. For the direct CF groups, the learners had two opportunities to notice their corrections. The first being when they attended to the corrections by studying them and the second being during the mechanical process of incorporating these corrections into their revised version of Task 3. The short-term improvement would appear to suggest that learners were able to successfully notice the surface features of the regular past tense (Schmidt, 1990). The continued improvement from the short- to the long-term suggests that the learners in these groups were able to develop metalinguistic understanding of the regular past tense (Schmidt, 2001). It could be argued that this is the case as the learners had no opportunities to notice the regular past tense between the short- and the long-term, so they may very well have drawn on a rule-of-thumb in that the CF had helped them to construct. The indirect CF groups also had an opportunity to notice their errors as they were required to study their corrections for five minutes, and they had an opportunity for pushed output in the revised version of a text. With indirect CF, learners have to draw on their existing linguistic knowledge and as there were no improvements in the short- or long-term, clearly this knowledge and the pushed output was insufficient for the learners written accuracy to improve over time.

8.4.1.5 Written corrective feedback has a negative effect on fluency and induces avoidance

Truscott's (1996) claim that CF has a negative effect on fluency and induces avoidance was not

directly investigated in this study.

8.4.2 Ferris' arguments in favour of written corrective feedback

Ferris (1999) produced a number of arguments in favour of CF. This study directly investigated one of them - this being the treatable and untreatable nature of grammatical errors.

8.4.2.1 Treatable versus untreatable grammatical features

Ferris (1999) introduced the notion of treatable versus untreatable errors. She claimed that errors following some kind of rule or regularity are treatable with CF whereas those that do not follow such a rule (e.g. lexis and idiomatic expressions) are untreatable by CF. From a SLA perspective, these notions of treatable and untreatable errors correspond to different ways in which structures are learned. Ellis (2009b) distinguished system-learned structures and item-learned ones. The former type corresponds to treatable errors and the latter untreatable ones. In the context of this study, the regular past tense is a system-learned structure and the irregular past tense and item-learned one. The findings across the whole study support Ferris' claim as to the treatable and untreatable nature of errors. While there is evidence that CF is effective under certain conditions with the regular past tense, there is almost no evidence that it is with the irregular past tense. The failure to find any effect for CF on the irregular past tense was largely demonstrated in the analysis of the effectiveness of CF on new pieces of writing. For those irregular past tense errors that were corrected, learners tended to not use them correctly in new pieces of writing (research question 1). One could argue that this was the case because learners had to write on a new topic involving different irregular forms. However, of those that were

subsequently used correctly, there was some evidence that some types of CF were more effective than others. Notably, focused direct CF was more effective in the long-term than unfocused indirect CF.

8.5 Pedagogical implications

Lee (2009) investigated teachers' beliefs about written feedback and teachers' written feedback practice, identifying ten mismatches between such beliefs and practice. Seven of these are relevant to the study presented here, so the pedagogical implications of the findings of the study will be discussed from the perspective of these seven problems with written CF.

8.5.1 "Teachers pay most attention to language form but they believe there's more to good writing than accuracy" (p. 15)

The first area of concern relates to teachers focusing on issues of accuracy when they believe that this is at the expense of providing feedback on content and organization. It was suggested that one of the reasons for teachers focusing on form at the expense of content and organization is that when learners have considerable problems with written accuracy, there is not enough room to also provide feedback on content and organization. She also notes that teachers are often concerned with focusing on accuracy to prepare learners for external examinations. However, this need not be the case. One study has demonstrated that learners would appear to be able to deal with issues of accuracy and content simultaneously (Ashwell, 2000). The findings of the study presented here demonstrated that when learners were required to attend to CF, a single episode of focused direct CF was effective in improving the accuracy of new

pieces of writing for the regular past tense (research question 1) and when learners were required to attend to CF and incorporate CF into a revised version of a text, focused direct CF was also effective in enhancing accuracy in subsequent pieces of writing for the same structure (research question 5). Providing focused direct CF would not compete with providing feedback on other equally important aspects of writing and as it enhances the acquisition of system-learned structures such as the regular past tense, one can argue that providing focused direct CF on such forms helps learners prepare for high stakes exams.

8.5.2 “Teachers mark errors comprehensively although selective marking is preferred” (p. 15)

Lee also found that while teachers thought that selective CF was preferable, comprehensive CF was the most common practice. Selective CF provides feedback on some predetermined feature while comprehensive CF provides CF on all errors. The study presented here found that focused CF outperformed unfocused CF in the long-term for the regular past tense (research question 2). A factor considered to have an influence on the effectiveness of focused over unfocused CF was the level of learners’ pre-existing knowledge about a target structure. It was shown that the Focused Direct CF Group outperformed the Unfocused Direct CF Group in the long-term (research question 2). The Unfocused Direct CF Group had the highest pre-existing levels of knowledge, received the fewest number of corrections and of these the majority were of a single correction. The Focused Direct CF Group, on the other hand, had the lowest pre-existing level of knowledge, received the greatest number of corrections with most learners receiving more than a single correction. Lower pre-existing levels of knowledge result in more corrections and the finding that there was a positive correlation between the number of

corrections received and gain score 2 suggests that learners require a certain amount of CF in order to successfully notice their corrections. Levels of pre-existing knowledge then may be another criteria for selecting CF. Ferris' (1999) notion of treatable versus untreatable errors is another possible selection criterion. It was demonstrated that there was no observable effect on acquisition for the item-learned irregular past tense whereas under certain conditions there was for the system-learned regular past tense (research questions 1 and 4).

8.5.3 "Teachers tend to correct and locate errors for students but believe that through teacher feedback students should learn to correct and locate their own errors" (p. 16)

Another issue for teachers was that they tend to correct and locate errors (i.e. direct CF); however, they believe that students should learn to correct and locate their own errors (i.e. self-edit). Learners can also be asked to correct errors that are located for them (e.g. indirect CF). There are two ways of interpreting this issue. The first of these is from the perspective of acquisition. The findings from the study showed that when learners were required to incorporate CF of the regular past tense into a revised version of a text, the direct CF groups improved over time while the indirect CF groups and the Control Group, which allowed for writing practice plus an opportunity to revise (i.e. self-edit), did not (research question 5). It was also demonstrated that direct CF outperformed indirect CF in the long-term and it outperformed the Control Group in the long-term (research question 6). Where acquisition is concerned, direct CF would appear to be superior to indirect CF and self-editing. Clearly learners will not be able to successfully correct located errors through indirect CF or correct and locate errors by self-editing unless they have the necessary linguistic resources. Learners

then cannot be expected to competently use indirect CF or self-edit until they have high levels of proficiency. The second perspective relates to training learners how to self-edit. Ferris (1995), for example, provides a breakdown of several activities including responding to CF designed to help learners self-edit. It is important, however, that if such training is given, learners are at the appropriate high level of proficiency to benefit from such training.

8.5.4 “Teachers respond mainly to weaknesses in student writing although they know that feedback should cover both strengths and weaknesses” (p. 17)

Another concern related to this study was the finding that nearly all feedback was directed at weaknesses despite the belief that Lee found that teachers felt they should also provide feedback on the strengths of student writing. This reflects the error focused approach to written feedback. Indeed, CF was used as negative evidence in this study in line with Long’s (1996) interpretation of negative evidence. Long also argued that there was a need for positive evidence about what is grammatical. While beyond the scope of this study, it is possible to provide such positive evidence through CF. Bitchener and Knoch (2010a), for example, combined negative evidence in the form of direct CF on learners’ errors and positive evidence in the form of ticks (√) above learners’ correct forms. Feedback could provide positive evidence on issues of written accuracy in writing. However, clearly there is a need to investigate the effects of providing such positive evidence in addition to or as an alternative to providing negative evidence in written CF.

8.5.5 “Teachers ask students to do one-shot writing although they think process writing is beneficial” (p. 18)

Students are often required to undertake writing as a product (one-shot) versus writing as a process even though teachers consider process writing to have more benefits for learners. Lee noted that due to the pressure of preparing learners for examinations teachers need to focus on multiple topics and texts types so they do not have the time for asking students to prepare multiple drafts. The study here does present an alternative. As mentioned above, Ashwell (2000) was able to demonstrate that learners are equally able to attend to issues of content and form at the same time, and it was argued and shown that the provision of focused direct CF with the opportunity to incorporate the CF into a revised version of a text is effective in enhancing accuracy in new pieces of writing (research question 5), and as such this type of feedback would not only help learners develop the accuracy required for exams but also provide enough room for teachers to provide written feedback on issues of content and structure. While not representing a typical process approach to writing, having learners incorporate CF into a revised version of a text represents part of a process approach to writing. In contexts where teachers cannot effectively use the established process approach, having learners complete a revised version of a text is a tenable option to incorporate a process element into a writing program.

8.5.6 “Teachers continue to focus on student written errors although they know that mistakes will recur” (p. 18)

Teachers also feel that despite concentrating on written errors, mistakes will continue. That is,

the same errors recur. This in all likelihood is the result of teachers providing comprehensive feedback and the wrong type of CF. The study presented here demonstrated that when learners were required to attend to CF, a single episode of focused direct CF was enough for the learners' accuracy to continue to improve from the short- to the long-term (research question 1) and that focused versus unfocused CF was more effective in the long term (research question 2). It was also demonstrated that when learners were required to attend to and incorporate focused direct CF into a revised version of a text, the accuracy of the learners' writing continued to improve from the short- to the long-term (research question 5), and, furthermore, direct CF was more effective than indirect CF in the long-term (research question 6). However, this was only the case for the system-learned and treatable regular past and not the case for item-learned and untreatable irregular past. When these findings are considered in relation to other studies that have found an effect for focused direct CF with system-learned structures (e.g. Ellis et al., 2008), it can be argued that CF will be effective if it is focused and direct and targeted at system rather than item-learned structures.

8.5.7 "Teachers continue to mark student writing in the ways they do although they think their effort does not pay off" (p. 18)

The last of the relevant mismatches identified by Lee is that teachers mark writing in ways they consider are not cost-effective. In other words, they consider that the amount of effort is not commensurate with perceived improvements in writing. The previous six areas of concern do provide some suggestions as to a more efficient and effective way of providing written feedback and CF in particular. Until further research suggests otherwise, there is a strong case

to be made for CF to be provided on system-learned and treatable structures and this CF should be of be both focused and direct. Teachers should require learners furthermore to first attend to the focused direct CF and then incorporate the CF, along with other issues of content and organization, into a revised version of a text. In such a way, it is suggested that the provision of feedback will be more cost-effective.

8.6 Limitations

In quasi-experimental studies such as the one presented here, it is necessary to use intact classes, and this brings with it a number of problems. Notably, the sample size is limited by the size of the classes. This in turn is affected by the mortality rate. It is to be expected in any classroom-based study that there will be a natural mortality rate as learners fail to attend one of several data collection episodes. With obligatory occasion analysis, a score of above 90 is often considered as representing a level at which a learner has acquired a particular structure (Brown, 1973). Learners who scored over 90 in any of the tasks that operated as a pre-test (Task 1 and Task 3), were removed from the study. This turned out to be the greatest contributor to the mortality rate and subsequent reduced sample sizes. Another issue related to using intact classes is that it is not possible to have control over learners pre-existing levels of knowledge of the target structures. This to some degree may have affected the results. Namely, those groups of learners scoring higher on the pre-test tasks would have received fewer corrections. It was demonstrated that the amount of corrections received by learners positively correlated with gain scores for the regular past tense.

An additional limitation with the study involves the control groups. The control groups were only required to complete the writing tasks or the writing tasks plus the revision task. The results for these groups may have been different if they had also been asked to study their papers looking for errors.

The final limitation of the study involves the past tense copula. It was intended that the study would not only investigate the regular and irregular past tense but also the past tense copula; however, the writing tasks were unable to elicit sufficient obligatory occasions of the copula to undertake an analysis of the effectiveness of CF for this structure.

8.7 Future research

Whilst undertaking this study, a number of areas for future research have been identified. First of all, there is a need to undertake research that examines the effectiveness of varying quantities of CF types on the accuracy of new pieces of writing. This could be undertaken with and without opportunities for revision. Such a study would perhaps need to be undertaken in both ESOL and EFL contexts. The value of such research would be to better understand the extent of CF required for learners to notice the CF or for pushed output to be effective.

The control groups were only required to complete the writing tasks and the revision. Future studies could investigate a writing practice group alone versus a writing practice plus reflection group. A writing practice plus reflection group would require learners to study their writing for errors before completing the next writing or revision task.

Bitchener and Knoch (2010a) demonstrated that when positive and negative evidence is combined it is effective in improving the accuracy of new pieces of writing. However, no study to date has investigated the comparative effectiveness of CF as positive and negative evidence. This is an area worthy of investigation.

There is now a growing body of research that supports the efficacy of CF particularly when it is focused and direct. Ferris (2004), however, identified a need for research investigating the effect of explicit instruction on the accuracy of writing. No study to date has compared explicit instruction and written CF. There is clearly a gap in the literature for such a study.

8.8 Conclusion

The research has gone some way to better understand the effects of CF. There were only effects for CF evident in the system-learned regular past. The findings found that when learners were required to study their CF for a period of time, a single episode of CF was effective in improving the accuracy of student writing - in particular for those learners that received focused direct CF. These learners alone demonstrated continued improvement from the short- to the long-term. There were no changes over time for writing practice. When learners had an opportunity to study the CF as well as incorporate corrections into a revised version of a text, both the focused and unfocused direct CF groups improved in the short-term and continued to demonstrate improvement from the short- to the long term. There is evidence that focused CF outperforms unfocused CF in the long-term when learners are required to study the CF, and when they are additionally asked to incorporate corrections into a revised version of a text,

direct CF outperforms indirect CF. From a theoretical standpoint the findings lend some support for the role of the cognitive processes of noticing and pushed output. The findings also have major implications for pedagogy by suggesting how the mismatches between writing teachers beliefs and practice can be addressed. There is a strong case to be made at this time for teachers to provide focused direct CF on system-learned structures and to require learners to first attend to the CF by reflecting on it and second incorporate the corrections and other types of content and organizational feedback into a revised version of a text. The findings are of notable relevance to the practice of teaching and learning of writing in the tertiary Chinese EFL context but are potentially of relevance to other pedagogic contexts as well.

LIST OF REFERENCES

Arndt, V. (1987). Six writers in search of texts: A protocol-based study of L1 and L2 writing. *ELT Journal*, *41*, 257-266.

Ashwell, T. (2000). Patterns of teacher response to student writing in a multiple-draft composition classroom: Is content feedback followed by form feedback the best method? *Journal of Second Language Writing*, *9*, 227-257.

Batstone, R. (1994). *Grammar*. Oxford: Oxford University Press.

Bitchener, J. (2008). Evidence in support of written corrective feedback. *Journal of Second Language Writing*, *17*, 102-118.

Bitchener, J. (2009). Measuring the effectiveness of written corrective feedback: A response to "Overgeneralization from a narrow focus: A response to Bitchener (2008)." *Journal of Second Language Writing*, *18*, 276-279.

Bitchener, J., & Knoch, U. (2008). The value of written corrective feedback for migrant and international students. *Language Teaching Research*, 12, 409-431.

Bitchener, J., & Knoch, U. (2010a). The contribution of written corrective feedback to language development: A ten month investigation. *Applied Linguistics*, 31, 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*, 207-217.

Bitchener, J., Young, S., & Cameron, D. (2005). The effect of different types of corrective feedback on ESL student writing. *Journal of Second Language Writing*, *9*, 227-258.

Brown, R. (1973). *A first language: The early stages*. Cambridge, Mass.: Harvard University Press.

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*, 267-296.

Chandler, J. (2004). A response to Truscott. *Journal of Second Language Writing*, *13*, 345-348.

Chandler, J. (2009). Response to Truscott. *Journal of Second Language Writing*, *18*, 57-58.

Cohen, A. (1987). Student processing of feedback on their compositions. In A.L. Wenden & J. Rubin (Eds.), *Learner strategies in language learning* (pp. 57-69). Englewood Cliffs, NJ: Prentice-Hall.

Cohen, A., & Cavalcanti, M. (1990). Feedback on compositions: Teacher and student verbal reports. In B. Kroll (Ed.). *Second language writing: Research insights for the classroom* (pp. 155-177). Cambridge: Cambridge University Press.

Cummings, A. (1990). Metalinguistic and ideational thinking in second language composing. *Written Communication*, *7*, pp. 482-511.

DeKeyser, R. (1998). Beyond focus on form: Cognitive perspectives 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.

Doughty, C. (1991). Second language instruction does make a difference. Evidence from an empirical study of SL relativization. *Studies in Second Language Acquisition*, *13*, 431-469.

Ellis, N. (1993). Rules and instances in foreign language learning: interactions of implicit and explicit knowledge. *European Journal of Cognitive Psychology, 5,* 289-318.

Ellis, R. (1994). A theory of instructed second language acquisition. In N. C. Ellis (Ed.). *Implicit and explicit language learning* (pp. 79-114). San Diego, CA: Academic Press.

Ellis, R. (2003). *Tasked-based language learning*. Oxford: Oxford University Press.

Ellis, R. (1998). Teaching and research: Options in grammar teaching. *TESOL Quarterly*, *32*, 39-60.

Ellis, R. (2009a). A typology of written corrective feedback types. *English Language Teaching Journal*, *63*, 97-107.

Ellis, R. (2009b). *The study of second language acquisition*. (2nd ed.). Oxford: Oxford University Press.

Ellis, R. (2010). A framework for investigating oral and written corrective feedback. *Studies in Second Language Acquisition*, *32*, 335-349.

Ellis, R., & Barkhuizen, G. (2005). *Analyzing learner language*. Oxford: Oxford University Press.

Ellis, R., Sheen, Y., Murakami, M., & Takashima, H. (2008). The effects of focused and unfocused written corrective feedback on Japanese university students' use of English articles in narratives. *System*, *36*, 353-371.

Fathman, A., & 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. (1995). Teaching students to self-edit. *TESOL Journal*, *4*, 18-22.

Ferris, D. (1999). The case for grammar correction in L2 writing classes: A response to Truscott (1996). *Journal of Second Language Writing*, *8*, 1-11.

Ferris, D. (2002). *Treatment of error in second language writing classes.* Ann Arbor, MI: University of Michigan Press.

Ferris, D. (2003). *Response to student writing: Implications for second language students*. Mahwah, NJ: Erlbaum.

Ferris, D. (2004). The grammar correction debate in L2 writing: Where are we and where do we go from here? (and what do we do in the meantime…?). *Journal of Second Language Writing*, *13*, 49-62.

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). Cambridge: Cambridge University Press.

Ferris, D., & Hedgecock, J. (1998). *Teaching ESL composition: Purpose, Process, and Practice.* Mahwah, New Jersey: Lawrence Erlbaum.

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. (1997). *Input, interaction and the second language learner*. Mahwah, NJ: Lawrence Erlbaum Associates.

Gass, S., Svetics, I., & Lemelin, S. (2003). Differential effects of attention. *Language Learning*, *53*, 497-545.

Guenette, D. (2007). Is feedback pedagogically correct? Research design issues in studies of feedback on writing. *Journal of Second Language Writing*, *16*, 40-53.

Hart, D., & Lapkin, S. (1989). *French immersion at the secondary/post-secondary interface*. (Final Report to the Ontario Ministry of Education.) Toronto: OISE, Modern Language Centre.

Hedgcock, J., & Lefkowitz, N. (1994). Feedback on feedback: Assessing learner receptivity to teacher response in L2 composing. *Journal of Second Language Writing*, *3*, 141-163.

Hedge, T. (1988). *Writing*. Oxford: Oxford University Press.

Hendrickson, J. M. (1978). Error correction in foreign language teaching: Recent theory, research, and practice. *Modern Language Journal*, *62*, 387-398.

Hendrickson, J. M. (1980). The treatment of error in written work. *Modern Language Journal*, *64*, 216-221.

Kepner, C. (1991). An experiment in the relationship of types of written feedback to the development of second-language writing skills. *Modern Language Journal*, *75*, 305-313.

Krashen, S. (1981). Second language acquisition and second language learning. Oxford: Pergamon.

Krashen, S. (1984). *Writing: Research, Theory, and Applications*. Oxford: Pergamon Press.

Krashen, S. (1985). *The Input Hypothesis: Issues and Implications*. London: Longman.

Krashen, S. (1998). Comprehensible Output? *System*, 26, 175-182.

Krashen, S., & Terrell, T. (1983). *The Natural Approach: Language Acquisition in the Classroom*. Oxford: Pergamon.

Lalande, J. (1982). Reducing composition errors: An experiment. *Modern Language Journal*, *66*, 140-149.

Lapkin, S., Hart, D., & Swain, M. (1991). Early and Middle French immersion programs: French language outcomes. *Canadian Modern Language Review, 48*, 11-40.

Lee, I. (1997). ESL learners' performance in error correction in writing: Some implications for college-level teaching. *System, 25*, 465-477.

Lee, I. (2009). Ten mismatches between teachers' beliefs and written feedback practice. *English Language Teaching Journal, 63*, 13-22.

Leki, I. (1991). The preferences of ESL students for error correction in college-level writing classes. *Foreign Language Annals, 24*, 203-218.

Leow, R. (1997). Attention, awareness and foreign language behaviour. *Language Learning, 47*, 467-505.

Leow, R. (2000). A study of awareness in foreign language behavior: Aware versus unaware learners. *Studies in Second Language Acquisition, 22*, 557-584.

Long, M. (1996). The role of the linguistic environment in second language acquisition. In W. Ritchie and T. Bhatia (Eds.). *Handbook of second language acquisition* (pp. 413-468). San Diego: Academic Press.

Lyster , R. (2004). Differential effects of prompts and recasts in form-focused instruction. *Studies in Second Language Acquisition*, *19*, 399-432.

Lyster, R., & Ranta , L. (1997). Corrective feedback and learner uptake. *Studies in Second Language Acquisition*, *19*, 37-66.

Lyster, R., & Saito, K. (2010). Oral feedback in classroom SLA: A meta-analysis. *Studies in Second Language Acquisition*, *32*, 265-302.

Milton, J. (2006). Resource-rich web-based feedback: Helping learners become independent writers. In K. Hyland and F. Hyland (Eds). *Feedback in Second Language Writing: Contexts and Issues*. (pp. 123-139). Cambridge: Cambridge University Press.

Oshima, A., & Hogue, A. (2006). *Writing Academic English* (4th ed.). New Jersey, NY: Longman.

Pallant, J. (2007). *SPSS Survival Manual*. Crows Nest: Allen and Unwin.

Pica, T. (1983). Adult acquisition of English as a second language under different conditions of exposure. *Language Learning*, *33*, 465-497.

Polio, C., Fleck, N., & 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.

Posner, M., & Petersen, M. (1990). The attention system of the human brain. *Annual Review of Neural Science*, *13*, 25-42.

Raimes, A. (1985). What unskilled ESL students do as they write: A classroom study of composing. *TESOL Quarterly*, *19*, 229-258.

Raimes, A. (1987). Language proficiency, writing ability, and composing strategies: A study of ESL college student writers. *Language Learning*, *37*, 439-457.

Richards, J., Hull, J., & Proctor, S. (1997). *New Interchange 1*. Cambridge: Cambridge University Press.

Robinson, P. (1996). Learning simple and complex rules under implicit, incidental, rule-search conditions, and instructed conditions. *Studies in Second Language Acquisition*, *18*, 27-67.

Robb, T., Ross, S., & Shortreed, I. (1986). Salience of feedback on error and its effect on EFL writing quality. *TESOL Quarterly*, *20*, 83-93.

Rosa, R., & O'Neill, M. D. (1999). Explicitness, intake, and the issue of awareness. *Studies in Second Language Acquisition*, *21*, 511-556.

Rutherford, W., & Sharwood Smith, M. (1985). Consciousness-raising and universal grammar. *Applied Linguistics*, *6*, 274-282.

Sachs, R., & Polio, C. (2007). Learners' use of two types of written feedback on a L2 writing revision task. *Studies in Second Language Acquisition*, *29*, 76-100.

Schmidt, R. (1990). The role of consciousness in second language learning. *Applied Linguistics*, *11*, 17-46.

Schmidt, R. (2001). Attention. In P. Robinson (Ed.) *Cognition and Second Language Instruction.* Cambridge: Cambridge University Press.

Schmidt, R., & Frota, S. (1986). Developing basic second language ability in a second language: A case study of an adult learner. In R. Day (Ed.) *Talking to Learn: Conversation in Second Language Acquisition* (pp. 237-236). Rowley, Mass.: New Bury House.

Segalowitz, N., & Lightbrown, P. (1999). Psycholinguistic approaches to SLA. *Annual Review of Applied Linguistics*, *19*, 46-63.

Semke, H. (1984). Effects of the red pen. *Foreign Language Annals*, *17*, 195-202.

Sharwood Smith, M. (1991). Speaking to many minds. *Second Language Research*, *7*, 118-132.

Sheen, Y. (2007). The effect of focused written corrective feedback and language aptitude on ESL learners' acquisition of articles. *TESOL Quarterly*, *41*, 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.* New York: Springer.

Sheen, Y., & Ellis, R. (2010). Corrective feedback in L2 teaching. In E. Hinkel (Ed.). *Handbook of research in second language teaching and learning*. New York: Routledge.

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*, 556-569.

Shintani, N. & Ellis, R. (mimeograph). The comparative effect of metalinguistic explanation and direct written corrective feedback on the acquisition of the indefinite article. *The University of Auckland*.

Swain, M. **(**1984). A review of immersion education in Canada Research and evaluation Studies. *ELT Documents*, 119, 35-51.

Swain, M. (1985). Communicative competence: Some roles of comprehensible input and comprehensible output in its development. In S. Gass & C. Madden (Eds.), *Input and Second Language Acquisition* (pp. 235-253). Rowley, MA: Newbury House.

Swain, M. (1995). Three functions of output in second language learning. In G. Cook & B. Seidlhofer (Eds.), *Principle and Practice in Applied Linguistics* (pp. 125-144). Oxford: Oxford University Press.

Swain, M., & Lapkin, S. (1995). Problems in output and the cognitive process they generate: A step towards second language learning. *Applied Linguistics, 16*, 371-391.

Tomlin, R., & Villa, V. (1994). Attention in cognitive science and second language acquisition. *Studies in Second Language Acquisition, 16,* 183-203.

Truscott, J. (1996). The case against grammar correction in L2 writing classes. *Language Learning*, *46*, 327-69.

Truscott, J. (1999). The case for grammar correction in L2 writing classes: A response to Ferris. *Journal of Second Language Writing*, *8*, 111-22.

Truscott, J. (2004). Evidence and conjecture on the effects of correction: A response to Chandler. *Journal of Second Language Writing*, *13*, 337-343.

Truscott, J. (2007). The effect of error correction on learners' ability to write accurately. *Journal of Second Language Writing*, *16*, 255-272.

Truscott, J. (2009). Arguments and appearances: A response to Chandler. *Journal of Second Language Writing*, *18*, 59-60.

Truscott, J., & Hsu, A. (2008). Error correction, revision, and learning. *Journal of Second Language Writing*, *17*, 292-305.

Uzum, B., & Yazan, B. (2011). Influences of learners' motivation on the processing of corrective feedback. *Applied Linguistics Forum*, *31*. Retrieved from http://umcp.acadamia.edu/ Bernadette Yazan/Papers/1304923

Van Beuningen, C., De Jong, N., & Kuiken, F. (2012). Evidence on the effectiveness of comprehensive error correction in second language writing. *Language Learning*, *62*,1-41.

White, L. (1989). *Universal Grammar and Second Language Acquisition*. Amsterdam: John Benjamins.

Xu, C. (2009). Overgeneralization from a narrow focus: A response to Ellis et al. (2008) and Bitchener (2008). *Journal of Second Language Writing*, *18*, 270-275.

Zamel, V. (1982). Writing: The process of discovering meaning. *TESOL Quarter1y*, *16*, 195-209.

Zamel, V. (1983). The composing processes of advanced ESL students: Six case studies. *TESOL Quarterly*, *17*, 165-187.

Zamel, V. (1985). Responding to student writing. *TESOL Quarterly*, *19*, 79-101.

APPENDIX A: INSTRUMENTS USED IN THE MAIN STUDY

Background Questionnaire

Chinese Name:_____English Name:_____Student ID:_____

1. How old are you?_____
2. Are you male or female?_____
3. What is your major (e.g. marketing, IT)?_____
4. How many years have you been studying English? _____

Task 1

Instructions:

5. Read the newspaper story and underline any new words you do not know.
6. When you have finished, get into groups of four and ask the people in your group what the unknown words mean.
7. If there are some words you do not know, ask your teacher.
8. Look at the pictures and read the story again.

The Lucky Dog

Last Tuesday, a car accident happened on Dunwha South Road. It involved two cars and a lucky dog.

A policeman explained the story. Around 10am, a driver of a black car turned left into the traffic in front of a small red car. The driver then started to drive north when a dog walked onto the road.

The driver stopped suddenly, and he avoided the dog. The red car, however, crashed into the black car, and it damaged the back of the car.

The man in the red car opened his door, jumped out and walked quickly to the other car. He shouted at the man. The man in the black car opened his door and shouted at the man.

Then they both laughed because they knew each other. They were old classmates from university, but they had lost contact twenty years ago.

The two friends next decided to look for the lucky dog. They searched under cars, and they looked around some trees. The dog, finally, walked up to the men and barked. The friends laughed again and returned the lucky dog to its owners.

Chinese Name:_____English Name:_____

Student ID:_____

Instructions:

2. Using the pictures, try and rewrite the story.

3 You may continue writing on the back of the page.

4. The start of the story has been given to help you.

Last Tuesday, a car accident happened on Dunwha South Road...

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____


1
2
3
4
5
6

Task 2

Instructions:

1. Read the newspaper story and underline any new words you do not know.
2. When you have finished, get into groups of four and ask the people in your group what the unknown words mean.
3. If there are some words you do not know, ask your teacher.
4. Look at the pictures and read the story again.

A Landslide in Nantou

Last Friday, a landslide happened in a village in Nantou. It buried a woman alive under tons of mud.

A policeman explained the story. Around 6am on Friday morning, a man returned home from a walk. He heard a loud sound and looked up at the hill behind his house. Terrified, he watched the landslide. The mud and dirt traveled down the hill. It missed his house, but it covered his neighbor's house.

He immediately started to dig because an old woman lived in the house. People from the village joined him, and the police arrived. They all wanted to find the woman, and they searched without rest.

The rescuers worked hard for many hours when suddenly a man shouted out. Everybody

stopped talking. They listened for a few minutes, and then they heard a woman's voice.

Everybody moved towards the sound and started to dig again. They worked and searched even harder. In the early afternoon, they reached the woman and pulled her out of the mud.

An ambulance rushed the woman to the hospital. However, she returned home the same day because she was, amazingly, not hurt.

Chinese Name:_____English Name:_____

Student ID:_____

Instructions:

1. Using the pictures, try and rewrite the story.

4 You may continue writing on the back of the page.

5. The start of the story has been given to help you.

Last Friday, a landslide happened in a village in Nantou...

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____


1
2
3
4
5
6

Task 3

Instructions:

1. Read the newspaper story and underline any new words you do not know.
2. When you have finished, get into groups of four and ask the people in your group what the unknown words mean.
3. If there are some words you do not know, ask your teacher.
4. Look at the pictures and read the story again.

The Lost Bag

Yesterday, a woman misplaced a bag at the Shih Lin Night Market. The bag contained 100,000 NT dollars.

As reported by a policeman, a woman and her son arrived at the market at about 8pm. For 2 hours, they walked around the market and talked with the vendors. They looked at the clothes and enjoyed the food.

At around 10pm, they decided to leave and walked to the train. At the subway, the woman suddenly realized something. She did not have her bag. She asked her son, but he did not have the bag.

So they returned to the market, and they visited each of the places again. They stopped at

every shop, looked for the bag and talked to many people. Finally, they asked a man at a small restaurant. He had the bag.

The woman opened the bag and looked inside. She smiled at the man and thanked him for his honesty.

Then she removed 100,000NT from the bag and handed him the money. The restaurant owner laughed and thanked the woman. The woman just wanted her favorite bag.

Chinese Name:_____English Name:_____

Student ID:_____

Instructions:

2. Using the pictures, try and rewrite the story.

5 You may continue writing on the back of the page.

6. The start of the story has been given to help you.

Yesterday, a woman misplaced a bag at the Shih Lin Night Market...

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____


1
2
3
4
5
6

Task 4

Instructions:

1. Read the newspaper story and underline any new words you do not know.
2. When you have finished, get into groups of four and ask the people in your group what the unknown words mean.
3. If there are some words you do not know, ask your teacher.
4. Look at the pictures and read the story again.

Jade Mountain

Last Sunday, two university students disappeared on Jade Mountain. Two days later, they called their parents in Tainan from their university dormitory in Taipei.

As reported by a local policeman, at around 8am on Sunday, the two students hiked up Jade Mountain for a one day trip. But in the evening, their parents had not heard from them, so they called the police.

On the second day, ten policemen searched the mountain. They looked in rivers and talked to people on the mountain. In the afternoon, more police searched around the mountain. Local people joined the search. They walked for many miles, and they looked everywhere. The army arrived and helped the police, too.

In the evening, however, the students phoned their parents. The parents called the police, and the police stopped the search.

After the hike on Sunday, they had traveled back to Taipei. They needed to study for a test the next day. On Monday, they stayed in the library until late, then returned to their dormitory and prepared dinner.

After they finished their meal, they turned on the TV, and they saw their pictures on TV.

Chinese Name:_____English Name:_____

Student ID:_____

Instructions:

3. Using the pictures, try and rewrite the story.

6 You may continue writing on the back of the page.

7. The start of the story has been given to help you.

Last Sunday, two university students disappeared on Jade Mountain...

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____


1
2
3
4
5
6

Task 5

Instructions:

5. Read the newspaper story and <u>underline</u> any new words you do not know.
6. When you have finished, get into groups of four and ask the people in your group what the unknown words mean.
7. If there are some words you do not know, ask your teacher.
8. Look at the pictures and read the story again.

KTV

A month ago, two students won a trip to Tokyo, but two weeks later they traveled to New York.

A radio DJ explained the story. A radio station in Taipei had a completion to find the best KTV singers. The two best singers would win a trip to Tokyo for two weeks.

For several weeks, many people tried to win the completion, but it was won by two university students from Taipei.

The two students went to Tokyo and stayed at one of the best hotels. They went to temples, visited Disneyland and enjoyed the hot pools.

One day, the noticed there was a KTV completion at the hotel, so they both decided to enter

it, they, of course, won the competition. They won a trip to New York.

They traveled to New York and stayed in an excellent hotel. They enjoyed the many interesting things to see there, ate at different restaurants and bought some interesting clothes.

While they were in New York, there was another KTV completion at their hotel, but they decided not to enter the completion. They just wanted to go back home.

Chinese Name:_____English Name:_____

Student ID:_____

Instructions:

1. Using the pictures, try and rewrite the story.
2 You may continue writing on the back of the page.
3. The start of the story has been given to help you.

A month ago, two students won a trip to Tokyo…

_____

_____

_____

_____

_____

_____

_____

_____

_____

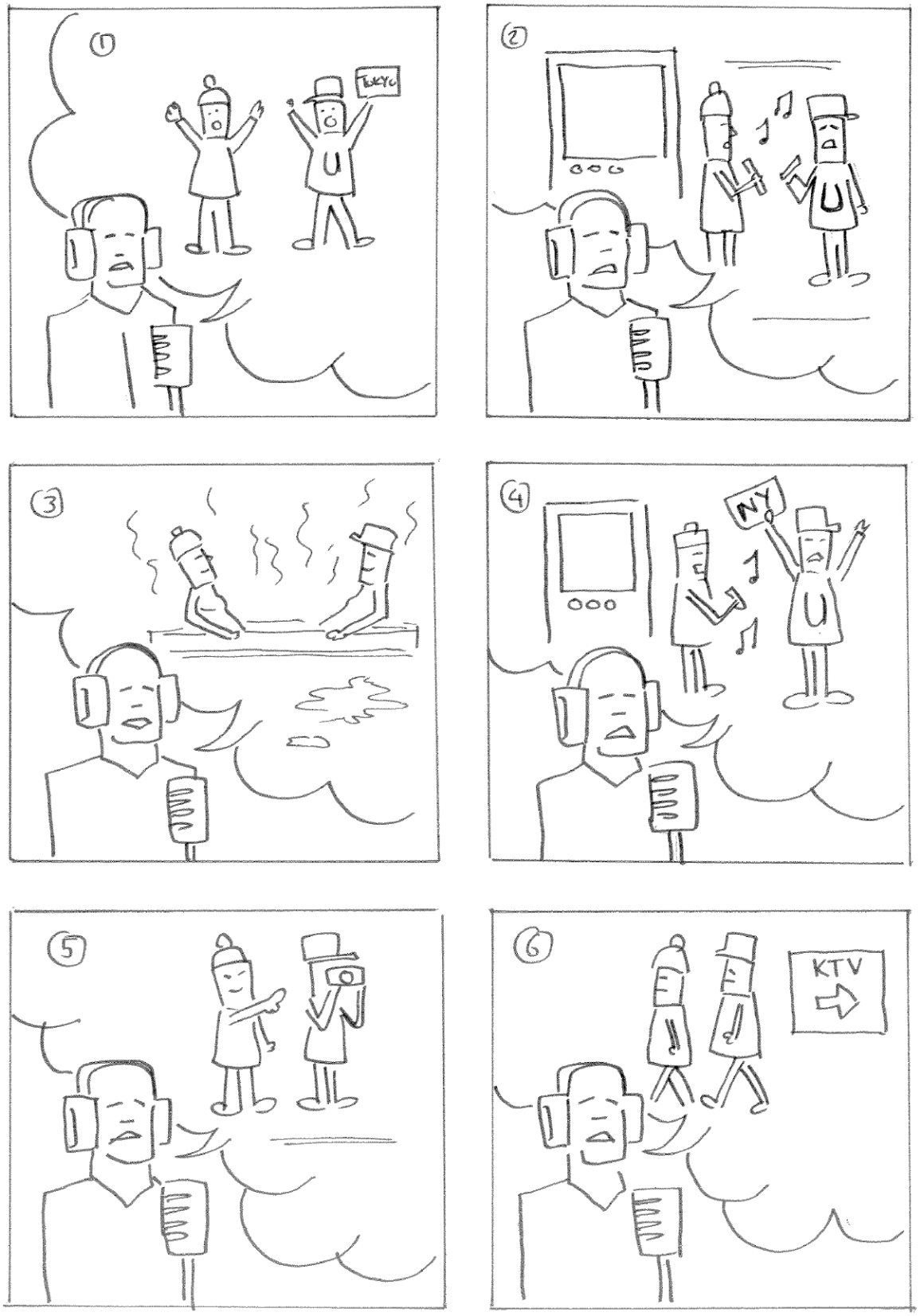
_____

_____

_____

_____

_____


1
TOKYO
2
3
4
NY
5
6
KTV

Revision Task

Chinese Name:_____English Name:_____

Student ID:_____

Instructions:

1. Look at your original version of your writing.
2. Write the story again, so it is better.
3. You may continue writing on the back of the page.

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

Exit Questionnaire

Chinese Name:_____English Name:_____Student ID:_____

1. What do you think the tasks were about?

    a. Practicing writing

    b. Practicing grammar

    c. Practicing reading

    d. Practicing vocabulary

2. What do you think you have learned from doing all the tasks?

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

APPENDIX B: ETHICS DOCUMENTS

THE UNIVERSITY OF AUCKLAND
www.auckland.ac.nz

Department of Applied Language Studies and Linguistics

The University of Auckland
Private Bag 92019
Auckland
New Zealand

Telephone: 64-9-373-7599

PARTICIPANT INFORMATION SHEET FOR INSTITUTIONS

Title: The Effects of Written Corrective Feedback and Revision on Intermediate Chinese Learners' Acquisition of English

To: The Program Director/ Head of Department

My name is David Frear. I am a Doctor of Philosophy student within the Department of Applied Language Studies and Linguistics at the University of Auckland, New Zealand. My major is language teaching. Part of the requirements of my degree is that I undertake research related to language teaching. In particular, I am interested in writing tasks. The research is longitudinal in that it will be conducted on four days over a four week period. Therefore, I need the assistance of some English language students so as to conduct my research. As such, I would like to invite some to participate in this research.

Please could you give the students the participant information sheets which invite the students to participate in the research? I also request your permission to access students on site.

What follows is a brief description of the anticipated implementation of the research covering four episodes of approximately one hour each over four weeks. Week one will involve the students completing a background questionnaire and writing task one. I will take writing task one away and write some feedback on the students' writing. In week two, the students will first work with the feedback from writing task one and then do writing task two. I will take writing task two away and write some feedback on the students' writing. For week three, the students will work on the feedback from writing task two and then revise this writing task. For week four, the students will complete writing task three. A number of weeks later writing task 4 will be completed. The questionnaire will take about ten minutes, the writing tasks about 30 minutes and working on the writing tasks about 30 minutes.
Please note that names of particular language learners and their individual results will NOT be included in the thesis.

Due to the voluntary nature of the participation in the research, I would like to seek your assurance that students' or teachers' participation or non-participation will not affect the student or teachers' relationship with the school or the grades of the students.
Moreover, if the institution decides that it does not want to participate in the research, it may withdraw at any time before and during the research and up to three weeks after the completion of the research. The institution is under no obligation to allow the data to be used.
All of the data in its raw form will be stored in a locked cabinet at Auckland University for six years and will be destroyed before March 31st, 2014.
If you would like further information about this project please feel free to contact me through my Taiwanese or New Zealand contact details below.

David Frear
Doctoral Candidate
Department of Applied Language Studies and Linguistics
University of Auckland
(0064) (9) 373 7599 Ext. 86914 (New Zealand)
d.frear@auckland.ac.nz

David Frear
5F, Number 7, Alley 20,
Lane 391, Section 3, Hoping East Road
Taipei, Taiwan
(00886) (2) 29525760 (Taiwan)
djfrear@hotmail.com

Professor Rod Ellis (Supervisor)
Department of Applied Language Studies and Linguistics
University of Auckland
(0064) (9) 373 7599 Ext. 84876 (New Zealand)
r.ellis@auckland.ac.nz

Associate Professor John Read (Head of Department)
Department of Applied Language Studies and Linguistics
University of Auckland
(0064) (9) 373 7599 Ext. 87673 (New Zealand)
ja.read@auckland.ac.nz

For any queries regarding ethical concerns please contact:
The Chair
Human Participants Ethics Committee
Office of the Vice Chancellor
Private Bag 92019
Auckland 1142
New Zealand
Tel: (0064) (9) 373 7599 ext. 83711

APPROVED BY The University of Auckland HUMAN PARTICIPANTS ETHICS COMMITTEE on 8th October 2008 for 3 years, on 8th October 2008
Reference: 2008 / 333

The University of Auckland
Private Bag
92019
Auckland
New
Zealand

Department of Applied Language Studies and Linguistics

Telephone: 64-9-373-7599

CONSENT FORM FOR INSTITUTIONS
THIS CONSENT FORM WILL BE HELD FOR A PERIOD OF SIX YEARS

Title: The Effects of Written Corrective Feedback and Revision on Intermediate Chinese Learners' Acquisition of English

Researcher: David Frear

From: The Programme Director/ Head of Department

The project has been explained to me and I understand what I have been told. I have had an opportunity to ask questions and have them answered.

- I agree to the researcher using the results of one background questionnaire, three writing tasks, written feedback and one revision of a writing task.
- I understand that the research will be conducted on five days with the research on each day lasting about 30 minutes to one hour. The questionnaire will take about ten minutes, the writing tasks about 30 minutes each and working on your writing tasks about 30 minutes each.
- I understand that the writing tasks will be designed expressly for the sake of research, they have no direct relationship to the routine part of the teaching cycle and that students can receive feedback on their performance after the completion of the research if they request it.
- I understand that the institution can withdraw its participation at any time before and during the research and up to three weeks after its completion.
- I understand that students' names will not be used in the final report on this project, so no one will know who the students are or their performance on the writing tasks.
- I affirm that participation or non-participation will not affect the students or teachers' relationship with the school or the grades of the students. I understand that the institution is under no obligation to participate in this project.

- I understand that any data gathered for this project will be stored in a locked cabinet at Auckland University for six years. The writing tasks, written feedback and revision of writing tasks will be destroyed before the 31st March, 2014.
- I understand that students may withdraw their data from this project at any time without giving a reason, and that they will be informed accordingly.

Signed: _____

Name: ……
Position: ……
Name of Institution……
(please print clearly)
Date: _____

APPROVED BY The University of Auckland HUMAN PARTICIPANTS ETHICS COMMITTEE on 8th October 2008 for 3 years, on 8th October 2008
Reference: 2008 / 333

The University of Auckland
Private Bag
92019
Auckland
New
Zealand

Department of Applied Language Studies and Linguistics

Telephone: 64-9-373-7599

PARTICIPANT INFORMATION SHEET FOR TEACHERS

Title: The Effects of Written Corrective Feedback and Revision on Intermediate Chinese Learners' Acquisition of English

To: Teachers

My name is David Frear. I am a Doctor of Philosophy student within the Department of Applied Language Studies and Linguistics at the University of Auckland, New Zealand. My major is language teaching. Part of the requirements of my degree is that I undertake research related to language teaching. In particular, I am interested in writing tasks. The research is longitudinal in that it will be conducted on four days over a four week period. Therefore, I need the assistance of some English language teachers so as to conduct my research. As such, I would like to invite you to participate in this research.

What do you do?

You will implement a number of procedures within five episodes of data collection – one per week for four weeks and one three weeks later. Each episode will take approximately thirty minutes to one hour of class time. Prior to the data collection I, the researcher, will outline the procedures for collecting the data which involve reading instructions, reading a text twice and collecting the tasks. Week one will involve the students completing a background questionnaire and writing task one. I will take writing task one away and write some feedback on the students' writing. In week two, the students will first work with the feedback from writing task one and then do writing task two. I will take writing task two away and write some feedback on the students' writing. For week three, the students will work on the feedback from writing task two and then revise this writing task. For week four, the students will complete writing task three. Finally, in week eight, the students will complete writing task 4. The questionnaire will take about ten minutes, the writing tasks about 30 minutes and working on the writing tasks about 30 minutes.

What will the researcher do with the information collected during the classes?

Firstly, each participant and their relevant information will be assigned a number so that anonymity can be maintained. The information from the background questionnaire and the writing tasks will be analyzed and form the basis of a doctoral thesis to be presented by the 31st March, 2012. You may see a copy of the thesis upon its completion.

You may request to have some feedback on the instructional procedure upon the completion of the research. Feedback from the researcher prior to this would bias the results of the research.

Due to the voluntary nature of the participation in the research, I would like to seek your assurance that students' participation or non-participation will not affect the students' relationship with the school or their grades.
If you decide that you do not want to participate in the research, you may withdraw it at any time before and during the research and up to three weeks after the completion of the research. You are under no obligation to allow the data to be used. All of the data in its raw form will be stored in a locked cabinet at Auckland University for six years and will be destroyed before March 31st, 2014.

Thank you for considering participating in this research.

If you would like further information about this project please feel free to contact me through my Taiwanese or New Zealand contact details below.

David Frear
Doctoral Candidate
Department of Applied Language Studies and Linguistics
University of Auckland
(0064) (9) 373 7599 Ext. 86914 (New Zealand)
d.frear@auckland.ac.nz

David Frear
5F, Number 7, Alley 20,
Lane 391, Section 3, Hoping East Road
Taipei, Taiwan
(00886) (2) 29525760 (Taiwan)
djfrear@hotmail.com

Professor Rod Ellis (Supervisor)
Department of Applied Language Studies and Linguistics
University of Auckland
(0064) (9) 373 7599 Ext. 84876 (New Zealand)
r.ellis@auckland.ac.nz

Associate Professor John Read (Head of Department)
Department of Applied Language Studies and Linguistics
University of Auckland
(0064) (9) 373 7599 Ext. 87673 (New Zealand)
ja.read@auckland.ac.nz

For any queries regarding ethical concerns please contact:
The Chair
Human Participants Ethics Committee
Office of the Vice Chancellor
Private Bag 92019
Auckland 1142
New Zealand
Tel: (0064) (9) 373 7599 ext. 83711

APPROVED BY The University of Auckland HUMAN PARTICIPANTS ETHICS COMMITTEE on 8th October 2008 for 3 years, on 8th October 2008
Reference: 2008 / 333

The University of Auckland
Private Bag
92019
Auckland
New
Zealand

Department of Applied Language Studies and Linguistics

Telephone: 64-9-373-7599

CONSENT FORM FOR TEACHERS
THIS CONSENT FORM WILL BE HELD FOR A PERIOD OF SIX YEARS

Title: The Effects of Written Corrective Feedback and Revision on Intermediate Chinese Learners' Acquisition of English

Researcher: David Frear

From: The Teacher

The project has been explained to me and I understand what I have been told. I have had an opportunity to ask questions and have them answered.

- I agree to the researcher using the results of one background questionnaire, four writing tasks, written feedback and one revision of a writing task.
- I understand that the research will be conducted on five days with the research on each day lasting about 30 minutes to one hour. The questionnaire will take about ten minutes, the writing tasks about 30 minutes each and working on your writing tasks about 30 minutes each.
- I understand that the writing tasks will be designed expressly for the sake of research, they have no direct relationship to the routine part of the teaching cycle and that students can receive feedback on their performance after the completion of the research if they request it.
- I understand that I can withdraw my participation at any time before and during the research and up to three weeks after its completion.
- I also understand that students' names will not be used in the final report on this project, so no one will know who the students are or what score their performance on the writing tasks.
- I affirm that participation or non-participation of the students will not affect the students' grades.
- I understand that there is under no obligation to participate in this project.

- I understand that any data gathered for this project will be stored in a locked cabinet at Auckland University for six years. The writing tasks, written feedback and revision of a writing task will be destroyed before the 31st March, 2014.
- I understand that students may withdraw their data from this project at any time without giving a reason, and that they will be informed accordingly.

Signed: _____

Name: ……
Position: ……
Name of Institution……
(please print clearly)
Date: _____
APPROVED BY The University of Auckland HUMAN PARTICIPANTS ETHICS COMMITTEE on 8th October 2008 for 3 years, on 8th October 2008
Reference: 2008 / 333

The University of Auckland
Private Bag
92019
Auckland
New
Zealand

Department of Applied Language Studies and Linguistics

Telephone: 64-9-373-7599

PARTICIPANT INFORMATION SHEET FOR STUDENTS

Title: The Effects of Written Corrective Feedback and Revision on Intermediate Chinese Learners' Acquisition of English

To: Students

My name is David Frear. I am a Doctor of Philosophy student within the Department of Applied Language Studies and Linguistics at the University of Auckland, New Zealand. My major is language teaching. Part of the requirements of my degree is that I undertake research related to language teaching. In particular, I am interested in writing tasks. The research is longitudinal in that it will be conducted on five days spread over several weeks. Therefore, I need the assistance of some English language students so as to conduct my research. As such, I would like to invite you to participate in this research.

What do you do?

Firstly, you will complete one background questionnaire and writing task one. I will take writing task one away and write some feedback on your writing. In week two, you will first work with writing task one and then do writing task two. I will take writing task two away and write some feedback on your writing. For week three, you will work on writing task two and then revise this writing task. For week four, you will do writing task three. A number of weeks later you will complete writing task four. We will, thus, meet five times for about thirty minutes to one hour each time. The questionnaire will take about ten minutes, the writing tasks about 30 minutes and working on your writing tasks about 30 minutes.

What will the researcher do with the information collected during the classes?

Firstly, each participant and their relevant information will be assigned a number so that anonymity can be maintained. The information from the background questionnaire and the writing tasks will be analyzed and form the basis of a doctoral thesis to be presented by the 31st March, 2012. You may see a copy of the thesis upon its completion.

You may request to have some feedback on the instructional procedure upon the completion of the research. Feedback from the researcher prior to this would bias the results of the research.

If you decide that you do not want to participate in the research, you may withdraw it at any time before and during the research and up to three weeks after the completion of the research. You are under no obligation to allow the data to be used. All of the data in its raw form will be stored in a locked cabinet at Auckland University for six years and will be destroyed before March 31st, 2014.

Whether you participate in the research or not will not affect you grades or relationship with the school.

Thank you for considering participating in this research.

If you would like further information about this project please feel free to contact me through my Taiwanese or New Zealand contact details below.

David Frear
Doctoral Candidate
Department of Applied Language Studies and Linguistics
University of Auckland
(0064) (9) 373 7599 Ext. 86914 (New Zealand)
d.frear@auckland.ac.nz

David Frear
5F, Number 7, Alley 20,
Lane 391, Section 3, Hoping East Road
Taipei, Taiwan
(00886) (2) 29525760 (Taiwan)
djfrear@hotmail.com

Professor Rod Ellis (Supervisor)
Department of Applied Language Studies and Linguistics
University of Auckland
(0064) (9) 373 7599 Ext. 84876 (New Zealand)
r.ellis@auckland.ac.nz

Associate Professor John Read (Head of Department)
Department of Applied Language Studies and Linguistics
University of Auckland
(0064) (9) 373 7599 Ext. 87673 (New Zealand)
ja.read@auckland.ac.nz

For any queries regarding ethical concerns please contact:
The Chair
Human Participants Ethics Committee
Office of the Vice Chancellor
Private Bag 92019
Auckland 1142
New Zealand
Tel: (0064) (9) 373 7599 ext. 83711

APPROVED BY The University of Auckland HUMAN PARTICIPANTS ETHICS COMMITTEE on 8th October 2008 for 3 years, on 8th October 2008
Reference: 2008 / 333

The University of Auckland
Private Bag
92019
Auckland
New
Zealand

Department of Applied Language Studies and Linguistics

Telephone: 64-9-373-7599

CONSENT FORM FOR STUDENTS
THIS CONSENT FORM WILL BE HELD FOR A PERIOD OF SIX YEARS

Title: The Effects of Written Corrective Feedback and Revision on Intermediate Chinese Learners' Acquisition of English

Researcher: David Frear

From: Students

The project has been explained to me and I understand what I have been told. I have had an opportunity to ask questions and have them answered.

- I agree to the researcher using the results of one background questionnaire, four writing tasks, written feedback and one revision of a writing task.
- I understand that the research will be conducted over five days with the research on each day lasting about one hour. The questionnaire will take about ten minutes, the writing tasks about 30 minutes each and working on your writing tasks about 30 minutes each.
- I understand that my name will not be used in the final report on this research, so no one will know who I am or my performance on the writing tasks.
- I understand that any data gathered for this project will be stored in a locked cabinet at Auckland University for six years. The writing tasks, written feedback and revision of writing tasks will be destroyed before the 31st March, 2014.
- I understand that I can withdraw my participation at any time before and during the research and up to three weeks after its completion.
- I also understand that my participation or non-participation will not adversely affect my relationship with the school or my grades. I understand that I am under no obligation to participate in this project.
- I agree to the use of the background questionnaire, writing tasks, written feedback and revision of a writing task for research about writing.
- I can request feedback upon completion of the research.

- I can request results from the writing tasks.
- I agree to take part in this research.

Signed: _____

Name: _____
(please print clearly)

Date:_____

APPROVED BY The University of Auckland HUMAN PARTICIPANTS ETHICS COMMITTEE on 8th October 2008 for 3 years, on 8th October 2008
Reference: 2008 / 333

APPENDIX C: KOLMORGOROV-SMIRNOV TESTS OF NORMALITY

Research Questions 1 and 2 for the Regular Past Tense

Group	Task 1	Task 2	Task 3
Focused Direct	$D(15) = .13$, $p = .20$	$D(15) = .28$, $p < .01$*	$D(15) = .19$, $p = .14$
Unfocused Direct	$D(9) = .18$, $p = .20$	$D(9) = .23$, $p = .18$	$D(9) = .23$, $p = .19$
Focused Indirect	$D(11) = .32$, $p < .01$*	$D(11) = .15$, $p = .20$	$D(11) = .20$, $p = .20$
Unfocused Indirect	$D(14) = .16$, $p = .20$	$D(14) = .16$, $p = .20$	$D(14) = .18$, $p = .20$
Control	$D(16) = .20$, $p = .10$	$D(16) = .18$, $p = .20$	$D(16) = .18$, $p = .17$

* not normally distributed

Research Questions 3 and 4 for the Regular Past Tense

Group	Task 3	Revision of Task 3
Focused Direct	$D(10) = .28$, $p = .03$*	$D(10) = .39$, $p < .001$*
Unfocused Direct	$D(10) = .22$, $p = .20$	$D(10) = .48$, $p < .001$*
Focused Indirect	$D(16) = .24$, $p = .01$*	$D(16) = .37$, $p < .001$*
Unfocused Indirect	$D(11) = .17$, $p = .20$	$D(11) = .43$, $p < .001$*
Control	$D(10) = .16$, $p = .20$	$D(10) = .20$, $p = .20$

* not normally distributed

Research questions 5 and 6 for the Regular Past Tense

Group	Task 3	Task 4	Task 5
Focused Direct	$D(7) = .18$, $p = .20$	$D(7) = .16$, $p = .20$	$D(7) = .33$, $p = .02$*
Unfocused Direct	$D(9) = .27$, $p = .05$*	$D(9) = .22$, $p = .20$	$D(9) = .17$, $p = .20$
Focused Indirect	$D(15) = .24$, $p < .02$*	$D(15) = .21$, $p = .07$	$D(15) = .19$, $p = .17$
Unfocused Indirect	$D(11) = .17$, $p = .20$	$D(11) = .23$, $p = .12$	$D(11) = .17$, $p = .20$
Control	$D(9) = .19$, $p = .20$	$D(9) = .23$, $p = .19$	$D(9) = .24$, $p = .15$

* not normally distributed

Research Questions 1 and 2 for the Irregular Past Tense

Group	Task 1	Task 2	Task 3
Focused Direct	$D(11) = .12$, $p = .20$	$D(11) = .16$, $p = .20$	$D(11) = .18$, $p = .20$
Unfocused Direct	$D(11) = .24$, $p = .08$	$D(11) = .24$, $p = .07$	$D(11) = .25$, $p = .05$*
Focused Indirect	$D(9) = .22$, $p = .20$	$D(9) = .20$, $p = .02$*	$D(19) = .24$, $p = .15$
Unfocused Indirect	$D(14) = .15$, $p = .20$	$D(14) = .30$, $p < .01$*	$D(14) = .29$, $p < .01$*
Control	$D(11) = .26$, $p = .04$*	$D(11) = .14$, $p = .20$	$D(11) = .11$, $p = .20$

* not normally distributed

Research Questions 3 and 4 for the Irregular Past Tense

Group	Task 3	Revision of Task 3
Focused Direct	$D(10) = .22$, $p = .18$	$D(10) = .52$, $p < .001$*
Unfocused Direct	$D(9) = .18$, $p = .20$	$D(9) = .52$, $p < .001$*
Focused Indirect	$D(9) = .33$, $p < .01$*	$D(9) = .47$, $p < .001$*
Unfocused Indirect	$D(7) = .21$, $p = .20$	$D(7) = .43$, $p < .001$*
Control	$D(10) = .15$, $p = .20$	$D(10) = .15$, $p = .20$

* not normally distributed

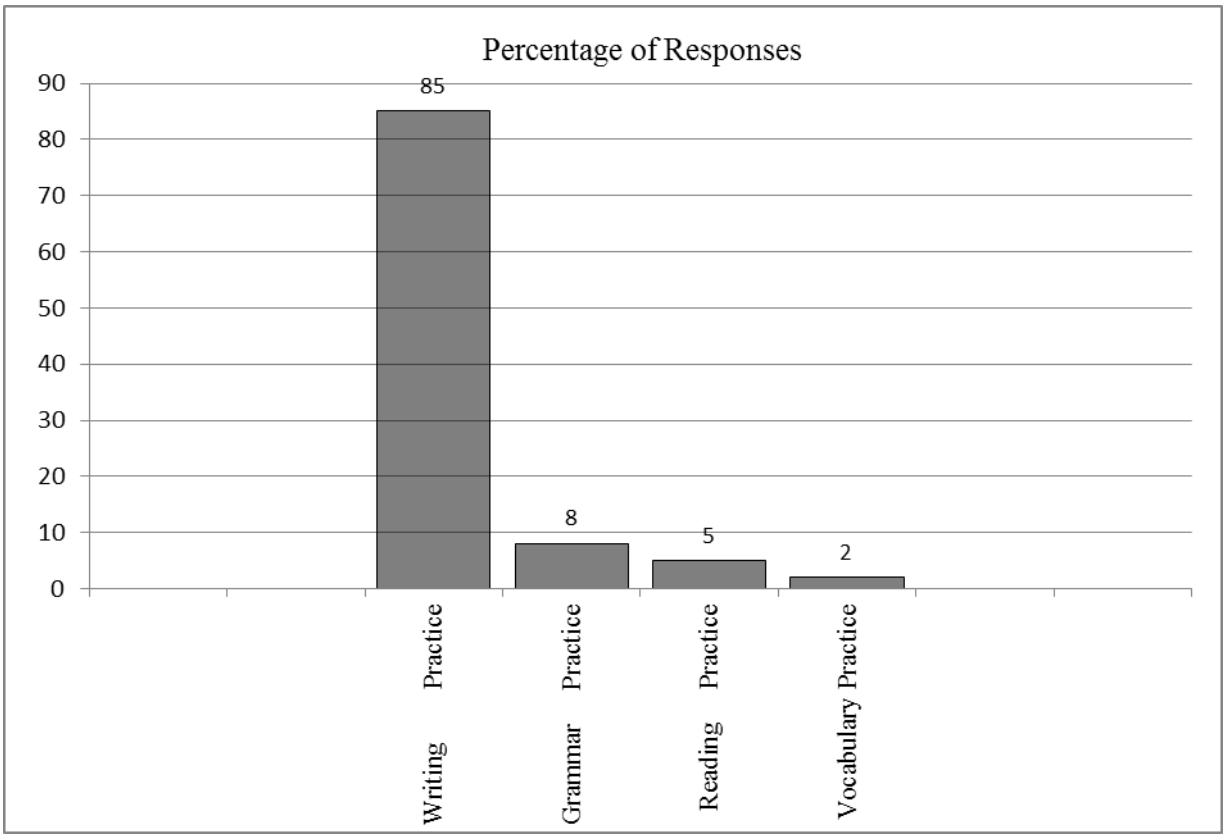
Research questions 5 and 6 for the Irregular Past Tense

Group	Task 3	Task 4	Task 5
Focused Direct	$D(7) = .22$, $p = .20$	$D(7) = .15$, $p = .20$	$D(7) = .15$, $p = .20$
Unfocused Direct	$D(9) = .20$, $p = .20$	$D(9) = .28$, $p = .04$*	$D(9) = .27$, $p = .07$
Focused Indirect	$D(9) = .33$, $p < .01$*	$D(9) = .32$, $p < .01$*	$D(9) = .35$, $p < .01$*
Unfocused Indirect	$D(7) = .21$, $p = .20$	$D(7) = .32$, $p = .03$*	$D(7) = .22$, $p = .20$
Control	$D(10) = .15$, $p = .20$	$D(10) = .17$, $p = .20$	$D(10) = .13$, $p = .20$

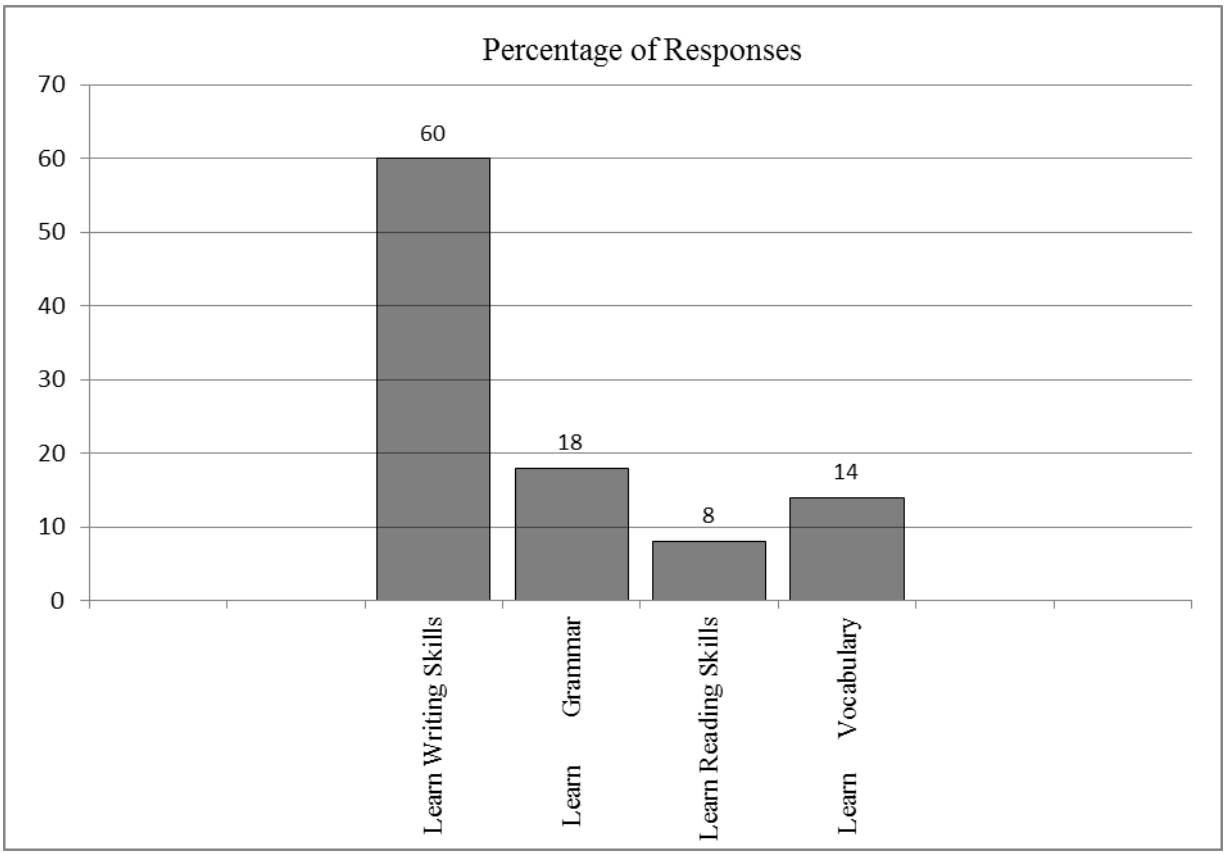
* not normally distributed

APPENDIX D: RESULTS FROM THE EXIT QUESTIONNAIRE

1. What do you think the tasks were about?



2. What do you think you learned from doing the tasks?



APPENDIX E: IRREGULAR PAST TENSE VERB USAGE IN TASK 2 AND TASK 3

Focused Direct CF Group

Learner	Error corrected	Task 2	Task 3
1	Have - had	Used correctly	Used correctly
	Get - got	Not used	Used correctly
	See - saw	Not used	Not used
2	Lose - lost	Not used	Not used
	Has - had	Used correctly	Not used
	Putted - put	Not used	Not used
	Leave - left	Not used	Not used
3	Drived - drove	Not used	Not used
	Finded - found	Used correctly	Used incorrectly
4	Tooked - take	Not used	Not used
5	Know - knew	Used correctly	Not used
	Keep - kept	Not used	Not used
	See - saw	Used correctly	Used correctly
	Get - got	Used correctly	Used correctly
6	Sung - sang	Not used	Not used
7	Hear - heard	Not used	Not used
	Don’t - didn’t	Used correctly	Used correctly
8	Get - got	Used correctly	Not used
	Seen - saw	Not used	Not used
9	Go - went	Not used	Not used
	Feel - felt	Used correctly	Not used
10	Drived - drove	Not used	Not used
	Keep - kept	Not used	Not used
	Shaked - shook	Not used	Not used
11	Run - ran	Not used	Not used
	Find - found	Not used	Used incorrectly

Unfocused Direct CF Group

Learner	Error corrected	Task 2	Task 3
1	Shaked - shook	Not used	Not used
2	Drived - drove	Not used	Not used
	Has - had	Used correctly	Not used
3	Taking - took	Not used	Not used
	Find - found	Used correctly	Used incorrectly
	Buying - bought	Not used	Not used
4	Tells - told	Not used	Not used
5	Have - had	Not used	Not used
6	Get - got	Used correctly	Not used
7	Win - won	Not used	Not used
8	Fonod - found	Used correctly	Not used
9	Tell - told	Not used	Not used
10	Get - got	Not used	Not used
11	Did saw – did see	Not used	Not used
	Give - gave	Not used	Not used

Focused Indirect CF Group

Learner	Error indicated and located	Task 2	Task 3
1	Drived - drove	Not used	Not used
	Leave - left	Not used	Not used
	Think -thought	Not used	Used incorrectly
2	Didn’t gone – didn’t go	Not used	Not used
3	Feel - felt	Used correctly	Not used
	Buy - bought	Not used	Not used
	Have - had	Not used	Not used
	Get - got	Not used	Not used
	Think - thought	Used correctly	Not used
	Lose - lost	Not used	Not used
	Find - found	Used correctly	Not used
	Give -gave	Not used	Not used
	Forget - forgot	Not used	Not used
4	To drove – to drive	Not used	Not used
5	Bend - bent	Not used	Not used
6	Hold - held	Not used	Not used
7	Hold - held	Not used	Not used
8	Choose - chose	Not used	Not used
	Don’t – didn’t	Not used	Used correctly
9	Keeps - kept	Not used	Not used
	Don’t – didn’t	Not used	Used correctly

Unfocused Indirect CF Group

Learner	Error indicated and located	Task 2	Task 3
1	Holded - held	Not used	Not used
2	Might lost might lose	Not used	Not used
	Has - had	Used correctly	Not used
3	Would heard - would hear	Not used	Not used
	Think - thought	Not used	Not used
4	Hang - hung	Not used	Not used
	Have - had	Used correctly	Not used
	Standed - stood	Not used	Not used
5	Spend - spent	Not used	Not used
6	Tell - told	Not used	Not used
7	Beginning - began	Not used	Not used
	Have - had	Not used	Not used
	Drinks – drank	Not used	Not used
	Says - said	Not used	Not used
	Let ... knew – let ... know	Not used	Not used
	Make - made	Not used	Not used
	Tell - told	Used correctly	Not used
	Do - did	Used correctly	Not used
	Take - took	Not used	Not used
8	Come - came	Not used	Not used
9	Taked - took	Not used	Not used
10	Know - knew	Not used	Not used
11	Shaked - shook	Not used	Not used
12	Forgived - forgave	Not used	Not used
13	Shaked - shook	Not used	Not used
14	Taked - took	Not used	Not used

APPENDIX F: IRREGULAR PAST TENSE VERB USAGE IN THE REVISION, TASK 4 AND TASK 5

Focused Direct CF Group

Learner	Error corrected	Revised Text	Task 4	Task 5
1	Lose - lost	Used correctly	Not used	Not used
	Didn't found – didn't find	Used correctly	Not used	Used incorrectly
2	Eaten - ate	Used correctly	Not used	Not used
	Weared - wore	Used correctly	Not used	Not used
3	Tell - told	Used correctly	Not used	Not used
4	Tell - told	Used correctly	Used correctly	Not used
	Sayed - said	Used correctly	Not used	Not used
5	Drived - drove	Used correctly	Not used	Not used
6	Has - had	Used correctly	Not used	Not used
7	Take - took	Used correctly	Not used	Not used
	Find - found	Used correctly	Not used	Used correctly

Unfocused Direct CF Group

Learner	Error corrected	Revised Text	Task 4	Task 5
1	Do - did	Used correctly	Not used	Not used
	Gaved - give	Used correctly	Not used	Not used
2	Get - got	Used correctly	Not used	Not used
	Know - knew	Used correctly	Not used	Not used
	Find – found	Used correctly	Used incorrectly	Not used
3	Find – found	Used correctly	Not used	Not used
4	Drive - drove	Used correctly	Not used	Not used
	Give - gave	Used correctly	Not used	Not used
5	Find – found	Used correctly	Not used	Not used
6	To held – To hold	Used correctly	Not used	Not used
7	Lose - lost	Used correctly	Not used	Not used
	Swared -swore	Used correctly	Not used	Not used
8	Could won – could win	Used correctly	Not used	Not used
	to won – to win	Used correctly	Not used	Not used
	Flue - flew	Used correctly	Not used	Not used
	Has - had	Used correctly	Used correctly	Not used
9	Flied - flew	Used correctly	Not used	Not used
	Buy - bought	Used correctly	Not used	Not used

Focused Indirect CF Group

Learner	Error indicated and located	Revised Text	Task 4	Task 5
1	Think - thought	Used incorrectly	Not used	Not used
	Know - knew	Used correctly	Not used	Not used
	Bring - brought	Used incorrectly	Not used	Not used
2	Didn’t neet - didn’t need	Used correctly	Not used	Not used
3	To won - to win	Used correctly	Not used	Not used
	Having - had	Used correctly	Not used	Not used
	Flied - flew	Used incorrectly	Not used	Not used
4	Take - took	Used correctly	Not used	Used correctly
	Find - found	Used correctly	Not used	Used correctly
5	Keeped - kept	Used correctly	Not used	Not used
6	Flied - flew	Used correctly	Not used	Not used
7	Couldn’t found - couldn’t find	Used correctly	Not used	Not used
8	Finded - found	Used correctly	Used correctly	Not used
9	Drived - drove	Used correctly	Not used	Not used

Unfocused Indirect CF Group

Learner	Error indicated and located	Revised Text	Task 4	Task 5
1	Know - knew	Used correctly	Not used	Not used
2	Find - found	Used correctly	Used correctly	Not used
	Hided - hid	Used incorrectly	Not used	Not used
3	Drived - drove	Used correctly	Not used	Not used
4	Wen - went	Used correctly	Used correctly	Not used
	Sell - sold	Used incorrectly	Not used	Not used
5	Leave - left	Used correctly	Not used	Not used
	Get - got	Used correctly	Not used	Not used
	Take - took	Used correctly	Not used	Not used
6	To won – to win	Used correctly	Not used	Not used
7	Have took – have taken	Used correctly	Not used	Not used