EXPLORING SOME EFFECTS OF DIFFERENT TYPES OF ERROR CORRECTION FEEDBACK ON ESL STUDENT WRITING

by

JACKLINE BONARERI AREGE

Submitted in accordance with the requirements for

The degree of

MASTER OF ARTS WITH SPECIALISATION IN APPLIED LINGUISTICS

at the

UNIVERSITY OF SOUTH AFRICA

SUPERVISOR: PROF E H HUBBARD

December 2010

DECLARATION

I Jackline Bonareri Arege do solemnly declare that this dissertation represents my own work and that it has not been previously submitted for a degree at this or another university.

Signature……..

Date: …….

DEDICATION

I dedicate this dissertation to my loving late husband Dettmer Orori. To my loving children Palmer, Ivy and Sunny, and my loving father for their unfailing support.

ACKNOWLEDGEMENTS

Special thanks are due to my academic supervisor Professor Hilton Hubbard, who patiently guided me throughout the project.

I also extend my thanks to the head teacher, the teachers and the students of YWCA Secondary School (Botswana), where I undertook the research, for their co-operation throughout the study.

My thanks also go to all my friends who encouraged me when the going was tough. Last but not least, my children and my father, who were patient, understanding and gave me a lot of encouragement.

ABSTRACT

This study uses a predominantly quantitative approach to explore the effect of different error correction feedback mechanisms on students' English as a Second Language writing (narrative and descriptive) amongst high school students in Botswana. A longitudinal, quasi-experimental design is used, with a control group that received no correction feedback while the experimental groups received direct, coded and uncoded feedback. Three hypotheses define the study in terms of fluency, correction success and accuracy development over time. No significant increases in fluency were found between the pretests and posttests. Correction success achieved by the three treatment groups when rewriting texts reflected the explicitness of the feedback, with the direct group highest, followed by the coded and uncoded groups. Findings were mixed on the important issue of accuracy development, although they strongly suggest that for spelling, any type of feedback is significantly better than none and that coded feedback is better than direct feedback despite the latter being more explicit. Students from all the treatment groups expressed similarly positive opinions on correction feedback.

Key terms: quasi-experiment; teacher error correction; direct feedback; coded feedback; uncoded feedback; rewriting; revision; fluency; accuracy; ESL writing.

TABLE OF CONTENTS

Page

LIST OF TABLES

Page

CHAPTER 1

INTRODUCTION

1.0 Introduction

This chapter is subdivided into four major sections. In the first (1.1) the background to the study is discussed. Then follow the research problem, aims and questions that are used in the study (1.2), the hypotheses of the study (1.3), and lastly structure of the dissertation (1.4).

1.1 Background to the study

This section gives background information to the study and this is discussed under four subheadings based on what teacher correction entails. The first part discusses the meaning of teacher correction as defined by some scholars. The following part provides a preliminary discussion of accuracy and fluency. The third part gives information on the role that English plays as a second language in Botswana, where the research was done, while the fourth gives a brief history of the school where the study was carried out.

1.1.1 Error correction feedback

Writing is one of the basic skills in the learning of either a first or a second language. It involves mastering the elements of grammar of a language, vocabulary, mechanics, content, organisation and style. In order to develop writing skills, it is essential for learners to receive feedback. Richards, Platt, and Platt (1992, 137) define feedback as:

> Any information which provides reports on the result of behaviour, for example verbal or facial signals which listeners give to speakers. … In the teaching context, these may be comments or information learners receive on the success of a learning task either from the teacher or from other learners.

Error correction feedback in this study should be distinguished from other types of teacher response that can be made to student writing, such as feedback on content or interpersonal feedback. It refers to the correction that students received on the errors that they made in their essay writing from the researcher only, hence ‘teacher error correction feedback’. In the study, ‘teacher feedback’ will often be used to cover the longer label ‘teacher correction feedback’ while ‘teacher correction feedback

mechanisms' in this study refer to the type of correction that the students received on their writing. In this study 'feedback mechanism' will often be used for the longer label 'teacher correction feedback mechanism'.

Teacher correction feedback can take various forms. It can be direct or indirect feedback. In direct correction feedback, a teacher indicates the presence or the location of errors in the text and also provides clues or tips on how the students could correct their error. It includes: (a) underlining a word and providing a written tip, for example *she go* $^{past\ tense}$ *home yesterday*; (b) bracketing a misplaced word or phrase and indicating its proper place in a sentence, for example "*keep quiet [their mother]*" $^{their\ mother}$ *told them*; (c) crossing out a superfluous word, for example *John went to ~~to~~ the party late;* and (d) providing a correct form or structure of an incorrect word or phrase, for example. *They cutted* ^{cut} *the fence yesterday*. In the above examples, (a) is slightly different from the other three in that it is less direct as it requires the students to do some thinking while the others require the students to just transcribe the changes that have been made. In indirect correction, the learners are not given any clue about how they should correct the errors that they have made, however the teacher indicates the presence of an error. The teacher can give indirect correction by: (a) underlining the incorrect forms; (b) circling an inappropriate word; (c) inserting an arrow to indicate a missing word; and (d) placing a question mark alongside a confusing phrase or structure (Hendrickson 1980, 218). However the question arises as to how explicit indirect correction should be. Indirect correction can be coded or uncoded. When it is coded, the learners are given codes for the type of errors made, for example *v* to show that a verb error has occurred. The codes are explained to the students so that they can correct the errors. Uncoded indirect feedback happens when the teacher just circles or underlines the error and the student is left to figure out the type of error that has occurred. There are however no general standards that exist with regard to whether, when, which or how student errors should be corrected and who should correct (Hendrickson 1978).

Arguments have arisen about whether teacher corrective feedback is necessary for students or not: in other words, does error correction benefit students? Feedback is one of the factors in Skinner's operant conditioning model of learning closely tied to behaviourist learning theory. In this learning model, feedback is equated with positive or negative reinforcement. The two aspects of feedback bring about the strengthening or weakening of behaviour (Chaudron 1988). This can be demonstrated by Skinner's

operant conditioning as discussed in Brown (1994). In terms of this theory it is argued that if teachers let errors go uncorrected, students may assume that they are correct in their writing. This means that the teacher has given positive reinforcement and it may lead to the errors becoming internalised in the speech patterns of the students. Once this happens then fossilisation of errors may take place, as the errors may persistently occur in the learners' speech or writing. On the other hand, if the teacher corrects the errors that have occurred it means that negative reinforcement has occurred and this could mean that the student will try to avoid the error occurring again and it may be eradicated eventually. Some non-behaviourist scholars support teacher error correction. Schachter (1984, cited in Pica 1994, 68) argues that if learners do not get any form of feedback, then they may assume that their utterances are accurate. Other scholars who concur with the argument include Lalande (1982), Pica (1994) and Brown (1991). Pica (1984, 69) argues that calling student's attention to differences between the errors and their correct version is very important as it leads to their success in language learning. Brown (1994, 19) adds that, if errors are never corrected, the learners will not get the feedback that they need while Hendrickson (1978) points out that it is difficult for learners to identify their own errors and thus they need someone more knowledgeable to point them out.

However, Truscott (1996) notes that just as much as error correction feedback seems to be necessary, sometimes it has not been found helpful, it is time consuming and such time could be used for other constructive activities. Teachers spend lots of time correcting students' errors, yet the same errors keep occurring again and again in the students' work. Hendrickson (1978, 216) states that:

> One thing is certain, providing all the correct forms in students' imperfect sentences is a time consuming ordeal that can be frustrating to the teacher, especially when they see identical errors recurring again and again in compositions written over a period of time by the same student. It is also disconcerting to receive a corrected composition with many words crossed out, new words added and an array of marginal comments all usually written in blood red ink. Small wonder that some students are embarrassed and lose confidence when they receive their written work corrected this way. This may hinder rather than facilitate the process of second language acquisition in formal situations.

Teacher correction feedback has aroused interest amongst many researchers, with a lot of research being done. Some studies that have been carried out suggest that teacher correction feedback is not useful while others suggest that it is, as will be discussed in the literature review in this study. A

consensus has not yet been arrived at and more research in this area is necessary. This study sets out to investigate whether teacher correction feedback leads to improvement in accuracy and in fluency, which are major components of writing in a second language.

1.1.2 Accuracy and fluency

Richards et al. (1992, 142) define accuracy as 'the ability to use grammatically correct sentences but may not include the ability to speak or write fluently'. Some researchers have taken accuracy as one of the major measurements in their studies. Frantzen (1995) measured accuracy in terms of the number of errors coded as a proportion of the number of correct uses of each structure, as identified with the help of a computer programe that compared the number of coded errors to the number of coded correct uses. Chandler (2003) measured accuracy by calculating the number of errors per 100 words. Improvement in accuracy is measured in terms of learners making fewer errors within a specified time, which could be in the short term (from one draft to the next) or in the longer term (for example, within a semester or an academic year). Accuracy is one of the important elements in language acquisition since mastering the grammar of a language means acquisition of abstract rules and the ability to use them automatically (Havranek 2002).

Accuracy is one of the major components in the learning of a second language and one of the factors that indicate how competent a learner is in a language. As elsewhere, in examinations in Botswana students are assessed in writing and are required to write accurately, losing marks if their grammar is not accurate. At the international level, the scenario is not any different. In some of the international examinations such as the Test of English as a Foreign Language (TOEFL), International English Language Testing System (IELTS) and Scholastic Assessment Test (SAT), which are a requirement for students to join some universities, academic writing accuracy is emphasised. Students are judged on their ability to use grammar correctly.

The issue of fluency in student writing has also been of major concern for some researchers and has been measured in various ways in student writing. Richards et al. (1992, 141) define fluency as 'the ability to produce written and or spoken language with ease'. Frantzen (1995) measured fluency by taking into account: (a) the length of the essay; (b) the expressive ability revealed; and (c) the

complexity of sentence structure. For many researchers fluency is measured in terms of the number of words written, but Chandler (2003) measured fluency differently. Since in her data the length of the composition was stipulated for the students, fluency was measured in terms of the length of time taken to write it. Students were asked to indicate the amount of time they took to write an assignment. Writing was done as homework assignments. The students were simply asked to write approximately five pages about their own life but they were not required to write in a particular order and this was done as a homework assignment. The time each student reported spending on the first and fifth assignment was then calculated per 100 words and change over the semester was one of the outcome measures.

Thus, ability to write more words within a specified time shows improvement in fluency. Fluency is tested in many examinations. In Botswana, in all the primary and secondary school national examinations, students are required to write a given number of words within a limited time. For example, in the Form 5 (Grade 12) English examination, students are required to write an essay of between 350 and 500 words in one hour. At the international level, in examinations such as TOEFL, IELTS and SAT, students are also tested in fluency as they are given tests that are done within a limited time and the minimum number of words is specified. Students lose marks if their scripts are under the minimum word limit (IELTS-International English Language Testing System. www.ielts.com). Teachers of English are therefore expected to produce students who can write fluently. Apart from academic achievement, learners of a second language after completing secondary school should be able to communicate fluently both in oral and written communication. The next section describes the present situation in Botswana with regard to the role of English as a second language as this is where the study was carried out.

1.1.3 English as a second language in Botswana

Botswana is a multilingual and multicultural country with about 26 languages represented in the country. Since independence, the government of Botswana has allowed the use of English as the official language in the country while Setswana, a local language that is spoken by a majority of the people (78.2%) (Languages spoken in Botswana. www.botswanalanguages.com) is used as the national language. The other local languages are not given any public role. English is used as the official language of business, and is widely spoken in urban areas. Most of the written communication

is done in English (Languages spoken in Botswana-www.botswanalanguages.com). English is also the language of administration and the main medium of instruction, playing a major role in national examinations (Nyati–Ramahobo 1991). The educational language policy since independence has been that Setswana is the medium of instruction from Standard 1 to 4 (the first four years of school) and thereafter English is used. This however became controversial in the late 1980s when some teachers disagreed with the policy. A national commission in 1993 (cited in Arthur 1997, 2) recommended that Setswana be used as a medium of instruction in the first year only (Standard 1) and thereafter the medium should be English. This has been followed in the public government schools only, but in the privately owned English medium schools, English is used right from pre-school, the explanation being that the majority of the students who attend these English medium schools are children of expatriates and other foreigners who are non-speakers of Setswana. At the secondary school level English is officially the sole medium of instruction except for other languages offered as subjects, such as French and Setswana, however, a study by UNESCO (2005 in UNESCO Database United nations Educational, Scientific and Cultural Organisation. http://unescodoc.unesco.org/images/0014/001422/14225e.pdf.) revealed that English is usually used alongside other languages such as Setswana and Ikalanga in schools. This could be attributed to the fact that some of the students who join secondary schools have been found not to achieve enough literacy skills in English.

In the academic field, English is used as the subject of assessment not only in English as a subject but also in other content subjects such as Social Studies, Science and Religious Studies (Hule and Mooko 2002). In English language as a subject, students are assessed in writing from primary level to the end of secondary school. At the end of Standard 7 pupils sit for the Primary School Level Examination (PSLE). In this examination, students are assessed in continuous writing as they are expected to write an essay and a letter. At secondary school level, there are two examinations. One is done at the end of Form 3 (Grade 10) while another is done at the end of Form 5 (Grade 12). At the end of Form 3 students sit for the Junior Certificate (JC) examination, where they are again assessed in writing and are expected to write an essay and a letter. At the end of Form 5, students sit for the Botswana General Certificate of Education (BGCSE). In the English examination, the students again are tested in writing as they are expected to write an essay and a report. It is this examination that determines whether a student can join institutions of higher learning or not.

The government of Botswana has set in place very high standards for admission to higher education. In order for any student to enrol at the University of Botswana, English is one of the qualifying subjects, regardless of the faculty. In addition, English must be a minimum of grade C (60%) for non-science degree programmes and a minimum of grade E (40%) for science-based programmes (University of Botswana 2006). Teachers of English are expected to prepare their students to meet the requirements for institutions of higher learning and training, not only locally but also internationally. After high school, the students are also expected to communicate comprehensively, both orally and in writing, in society. The next section gives a profile of the secondary school where the present study was carried out.

1.1.4 A profile of YWCA secondary school

YWCA is the abbreviation for the Young Women's Christian Association. This is an active association in Botswana that has carried out various activities, such as the running of health centres as well as schools. It has opened a secondary school in Gaborone (the capital city of Botswana) where the present study is based. YWCA opened the school with the aim of providing basic education for teenage mothers. Schoolgirls who have fallen pregnant have few opportunities to go back to school, or if they do they fear being stigmatised by the other students. YWCA therefore felt that it would be a good idea if such girls came together in one environment to continue their studies. At the same time, the association would teach the teenage mothers how to take care of and bring up their children. The association even runs a clinic for the students' children within the compound as well as a day care unit and a pre-school. The YWCA initially had the intention of making the school for teenage mothers only, but with time boys were also enrolled. The reason given by the school authorities is that the enrolment of teenage mothers was low and since the school had substantial facilities it was felt that its finances could be boosted by admitting boys.

The school only offers senior secondary classes (Form 4 and Form 5), which is equivalent to Grade 11 and Grade 12 respectively. In a personal communication with the school principal, it was revealed that most of the students enrolled in this school were of a low academic standard, having attained an overall mean grade C and grade D in their Junior Certificate entrance examination at Form 3 level (Grade 10).

The school offers all the academic subjects in the syllabus of Botswana as well as extracurricular activities.

The study also made use of a structured interview (Appendix 4) with the teacher. The interview took place before the rest of the research was carried out. Below I report briefly on demographic information on the teacher and then discuss information elicited on the writing practices in the school where the study took place. The interview gave the study background information on the kind of writing that the students had been exposed to and the kind of feedback that the students were used to in their daily classes.

The co-operating teacher was interviewed by me and I took notes on the responses. The teacher was 32 years old and had a Bachelor's Degree in Education. He had ten years' teaching experience as a high school teacher. He had received some training in giving feedback in composition writing in his teacher training course at the university and in addition he had received training for two weeks in the marking of examinations. He had marked 'O' Level examinations, particularly in the area of composition, for a period of three years in his home country, Zimbabwe.

The main aim of the interview was to elicit information from the teacher on the type of writing that the students had been given during their course of study. This was important as my study was based on the narrative and the descriptive genres and I wanted to establish whether the students were familiar with both these genres or not. On the type of writing that the students did often, he gave four common genres in English writing, namely narrative, descriptive, argumentative and expository writing. The teacher however revealed that the genres were not covered to the same extent as he gave students topics to write on at random for tests but for the purpose of class work topics were written on depending on how they were covered in the syllabus.

On how often the students wrote compositions, the teacher said that composition writing was done at an interval of three weeks, meaning that in a term, the students wrote an average of four compositions. In the normal classroom setting, the learners were not given a word limit. However, they were given a minimum number of pages that they should write, such as, one to one and a half pages initially, then

gradually increasing to two pages and sometimes more, since the learners admitted to this particular school were considered to be of a low academic level.

On the type of error correction that was given to the students, the teacher said that most of the time he underlined the errors and put a code on the error or provided direct feedback, with codes used more often than the direct feedback mechanism. The different types of feedback were used concurrently and comprehensive error correction was done in that all errors were corrected at the same time.

The interview also sought to find out if any form of revision was usually carried out by the students after they had received teacher correction feedback for the essays that they had written. One of the research questions was to find out if revision led to correction success when students rewrote their essays. This was also to help me find out if students had carried out any revisions before and if they did not, seek a way of explaining to the students in the experimental groups why they were rewriting their work. If this had been standard procedure, I needed to seek a way of explaining to the control group why they would not be getting any form of feedback. The interview was also a way of finding out the teacher's opinion on teacher error correction feedback.

In practice the teacher admitted that after giving correction to the students, he did not follow up to find out if the students went through their work or not. There was no follow-up measure to ensure that the students attended to the feedback given. This response was useful to me as I could then explain to the students why all of a sudden they needed to rewrite their work through their teacher.

The teacher felt nevertheless that revision after teacher error correction was necessary in student writing. In particular he felt that underlining and identification of error by use of codes, rather than direct correction, was the best form of error correction, as these helped the students to recognise their errors easily, learn from them, probably not repeat them and improve in their writing. He expressed himself as follows: "How can a teacher of English say that they are teaching if they are not giving students any form of feedback. It is unheard of".

The information from the interview was very helpful as it highlighted aspects of the major issues that were studied in this project, such as the type of feedback given to students, the type of writing that was done by them and whether revision was given after feedback or not.

The school was chosen for this study because the school authorities allowed that the research be carried out in their school and were willing to support the researcher as much as possible.

1.2 The research problem, aims and questions

This section presents the research problem of the study and its aims and then, after providing introductory information on the design and procedures of the study, proceeds to the research questions. Most of the published research on teacher correction as a form of feedback, such as Chandler (2003), Ferris and Roberts (2001), Kepner (1991), Lalande (1982), Lee (1997), Polio, Fleck, and Leder (1998) and Robb, Ross, and Shortreed (1986), has been carried out at university or college level. In the literature review discussed in this study, only one study was carried out by Fazio (2001) at primary school level (Grade 5). The existing studies carried out at high school level made use of different forms of correction feedback such as peer correction but none on teacher correction feedback was found. In addition to this, most of these studies have been carried out in Western first world countries. Not much research on teacher correction feedback in ESL student writing has been done in third world countries such as Botswana. The research problem addressed in the present study, however, is that of investigating the effectiveness of teacher correction feedback on student writing in one secondary school in a developing country. The school chosen for the study was made up of students with low language proficiency and it was assumed that such students would need teacher correction feedback to help them notice their errors.

The overall aim of the study was to find out whether different types of teacher error correction lead to improvement in fluency and accuracy in student writing. This aim can be broken down into the following specific aims.

1. To investigate the effect of different teacher correction feedback mechanisms on students' writing fluency.
2. To investigate the effect of different teacher correction mechanisms on students' success in correcting errors during revision.

3. To investigate the effect of different teacher correction feedback mechanisms on the development of students' writing accuracy in general, and also with respect to different error categories.
4. To investigate students' views with regard to teacher correction feedback.

The subjects in the study are from two Form 4 classes who had sat for their Junior Certificate (JC) examination in Form 3 the previous year. Their ages range between 16 years and 21 years. They are considered to be of a low academic standard as a majority of them obtained grade C or grade D in the JC examination. The study was carried out for a period of nine weeks, which is equivalent to one academic school term in Botswana. The design of the study is quasi-experimental, with one class comprising three experimental groups and another class the control group. A total of 68 students were studied, however the data used for the study was obtained only from 57 students for the narrative essay and 50 students for the descriptive essay. These were the students who completed all eight essay-writing sessions. The teacher correction feedback mechanisms that were used in the experimental groups were: (a) direct feedback; (b) coded feedback; and (c) uncoded feedback. The control group received no correction feedback. For the class that was chosen as the experimental group, learners were assigned to each of the groups randomly. This was done by assigning the pretest scripts of the experimental class randomly to the three different feedback groups. The second class remained as the control group class and the students in each of these groups maintained their groups throughout the term. Students receiving direct feedback had all their errors corrected, those receiving coded feedback received a code for their errors while those receiving uncoded feedback had their errors underlined. The control group did not receive any type of correction but received a summative statement on their work.

Narrative and descriptive essay writing was examined in terms of a total of eight essays, of which four were done in each genre. All essays were written in class in an 80-minute period. After the original scripts were marked the scripts were returned to the students and those in the experimental groups were asked to rewrite their work in class within 40 minutes. The rewriting sessions were in the afternoon, when the students were expected to do their private study. Six error categories were selected for the study, namely: (a) verb errors; (b) article errors; (c) spelling errors; (d) punctuation errors; (e) preposition errors; and (f) word choice errors. These six error types chosen for the study are arguably

the most common ones found in written texts of learners of English as a second language (Ferris 2002, 53; Lee 1997). The limitation to these errors facilitated a more in-depth study which would not have been the case if so many error categories are studied at once

The study attempts to answer the following main research questions:

1. Do students who receive teacher error correction feedback improve in the fluency of their writing more than those who do not and do different feedback mechanisms have different effects on fluency?
2. How successful is teacher correction feedback in ensuring accurate corrections when students rewrite their essays, and do different feedback mechanisms have different effects on correction success?
3. Do students who receive teacher error correction feedback improve the accuracy of their writing over time more than those who do not, and are some feedback mechanisms better than others in this respect?
4. Do students value teacher error correction?

1.3 Hypotheses

For this study three hypotheses were formulated. These are directly linked with the first three aims of the study. There is no hypothesis formulated linked to the fourth aim as this was dealt with in a more exploratory manner. The hypotheses for the study are formulated in a non-directional manner and this is because research that has been done on error correction that I studied, and have discussed in the literature review, provides mixed findings. This made it quite difficult for me at the beginning of the study to assume that some feedback mechanisms are superior to others in any respect. The following were the hypotheses formulated for the study:

H1: There is a relationship between the feedback mechanism used and the development of fluency in students' writing.

H2: There is a relationship between the feedback mechanism applied to student errors and students' success in correcting errors during revision.

H3: There is a relationship between the feedback mechanism used and the development of accuracy in students' writing over time, in both overall terms and with regard to specific error categories.

1.4 Structure of the dissertation

The dissertation is subdivided into five main chapters. This chapter has dealt with the introduction to the study and is followed in Chapter 2 with a review of the research literature that is related to the study. Chapter 3 looks at the methodology of the study, while Chapter 4 presents the findings of the study. Finally, Chapter 5 concludes by discussing some contributions, limitations and implications of the study.

CHAPTER 2

LITERATURE REVIEW

2.0 Introduction

This section reviews the research literature that is relevant to the study. The chapter has been divided into three major sections. Section 2.1 gives a descriptive review of the study and is subdivided into four parts. Section 2.1.1 discusses the issue of teacher correction as an area of controversy. It reviews how studies that have been carried out on teacher corrective feedback have not yet arrived at a consensus as to whether feedback should be given to students or not. Section 2.1.2 reviews studies that have been carried out on the effect of teacher error correction on fluency and accuracy. This is in response to the first and second aims of the study, which are to examine the effect of teacher error correction on students' fluency and accuracy respectively. Amongst the studies reviewed some show that teacher error feedback is not helpful to students' writing, while others reveal that teacher error correction does lead to improvement in students' accuracy and fluency. Studies that have been carried out to find out the effect of revision on student writing are reviewed in this section too. This is to avoid repeated discussion of some studies. Section 2.1.3 reviews only those studies that have been carried out to examine the effect of different feedback mechanisms on different error categories. This is because my study focuses on six error categories and seeks to establish how comparable its findings are to what has already been done. Similarly, for further comparison with my study, Section 2.1.4 reviews studies that have been carried out on students' opinions on feedback. Section 2.2 provides a brief more critical review of the studies that have been discussed in the chapter while Section 2.3 concludes the chapter.

2.1 Descriptive review

This section reviews various studies that have been carried out and particularly their findings that are relevant to my research. The first gives a brief overview of the contradictory arguments that exist about error correction. This is followed by a review of studies in terms of the effect of different correction feedback mechanisms on fluency, overall accuracy, accuracy in terms of individual error categories and lastly on students' opinions on teacher feedback.

2.1.1 Correction feedback: an area of controversy

Feedback is an essential part of language acquisition and correction seems to have been accepted as one form of feedback (Horner 1988). There has been debate on whether or not to give grammar correction and the empirical evidence that exists is divided on this issue. Some research studies (such as Chandler 2003; Greenslade and Felix-Brasdefer 2006; Lalande 1982; and Lee 1997) suggest that teacher correction feedback is helpful to students. However, others (such as Bitchener, Young, and Cameron 2005; Fazio 2001; Kepner 1995; Robb et al. 1986; Semke 1984; and Ferris et al. 2001) suggest that it is not helpful to students. Students' preferences for teacher correction have been sought and some studies (Chandler 2003; Ferris and Roberts 2001; Greenslade and Felix-Brasdefer 2006; Hedgecock and Lefkowitz 1994; Lalande 1982; Lee 2004) found that learners showed strong preferences for teacher grammar correction. All the studies cited above are discussed in detail in the next section. If some studies show that error correction is helpful while others show that it is not and students want their errors corrected, what should the teacher do? Teachers are put in a situation where they are not sure whether to correct or not to correct students' errors and if they are to correct, what errors should be corrected and how these errors should be corrected. Hendrickson (1978) raised this issue over 30 years ago and to date no answer has been arrived at despite the numerous studies that have been done.

Truscott (1996) argues that grammar correction is harmful and according to him it should be abolished. He takes this strong point of view by reviewing studies such as Kepner 1991, Semke 1984 and Sheppard 1992, which revealed that there was no significant difference across the various types of feedback that were given in each of the studies. Truscott argues that if grammar correction is helpful, then students who receive it should perform much better than those who do not. He goes ahead and says that if no significant difference is found between students who receive grammar correction and those who don't, then grammar correction is not helpful. Truscott adds that if students who do not receive grammar correction perform better than those who receive it, then in this case he considers correction as harmful (Truscott 1996, 329). According to Truscott on the studies that he reviews, teacher correction feedback was not considered to be helpful and he questions why it should be given, while such time could be used in more constructive work since the teachers spend too much time correcting student errors and likewise the students spend a lot of time going through their work.

Ferris (2004), in response to Truscott (1996), argues that the evidence given by Truscott in his article is not enough to arrive at such a conclusion that grammar correction is harmful. She argues that Truscott compares studies that differ in a lot of parameters such as: the type of writing that was used, the length of study, the design, type of feedback given and who gave the feedback, and says that it is actually like comparing apples and oranges. Chandler (2004) also challenges Truscott's (1996) argument and says that some studies such as Ashwell (2000), Ferris and Roberts (2001) and Lee (1997) found that all groups that received error feedback outperformed those that did not in measures of accuracy. Truscott is accused of just mentioning studies that support his view. Chandler (2003) found that students who received error correction and carried out revision performed much better in measures of accuracy and fluency compared to those who did no revision. Truscott also argues that Sheppard (1992) showed that teacher correction was not helpful yet in this study the content group improved in sentence boundary markers (periods, semi-colons and question marks). Chandler (2004) also argues that sometimes Truscott gives the wrong interpretation to some of the studies that he reviewed; for example, Truscott (1996) says that Robb et al. (1986) revealed that error correction was not helpful, yet all the four groups that were used in the study showed improvement in syntactic complexity. The same study also did not have a control group yet Truscott wants to make us believe that the group that only had the number of errors in each line recorded was equivalent to a no-feedback control group. But this group received some form of feedback. Since Robb et al.'s study did not have a true control group; it is too easy for Truscott to argue that it is not the error feedback that produced the positive change.

Ferris (2004) advises teachers to make use of whatever research is available and deal with their own classroom situations with flexibility. Truscott (1999) concludes that the decision on whether to correct or not should be left to the teacher's decision. Both Truscott (1996) and Ferris (2004) however agree that more research needs to be done so as to arrive at a consensus on whether to give feedback or not and the present study addresses this issue of how effective teacher error correction is for student writing. The next section reviews studies that have been carried out using various treatments and the effects that they had on student writing.

2.1.2 The effect of teacher correction feedback

This section discusses the effects of different teacher feedback types on accuracy and fluency. It first deals with studies that found no significant difference amongst the various treatments that were given

and then follow those that found significant differences across treatments. The studies that are discussed consists of both short term and long term studies.

Accuracy refers to 'the ability of an individual to produce grammatically correct sentences' (Richards et al. 1992, 142) and is usually measured by how error free a text is. Fluency on the other hand in the writing context is 'the ability to write a specified number of words with ease' (Richards et al. 1992, 141) usually within a word and time limit. Fluency is usually measured by the number of words written although some researchers such as Chandler (2003) measure fluency in terms of the time taken to write a specified number of words.

Increasingly, studies have compared whether different feedback types lead to more improvement in accuracy and fluency than others. The findings are quite varied. Some studies, such as earlier ones discussed by Truscott (1996) including Kepner (1991), Semke (1984) and Sheppard (1992)) and others such as, Bitchener, Young, and Cameron 2005; Fazio (2001); Ferris and Roberts (2001); and Polio et al. (1998), found no significant difference amongst the various treatments that were given to the subjects. On the other hand, some studies found significant differences across treatments (Chandler 2003; Fathman and Whalley 1990; Ferris 1997; Ferris et al. 2000 as discussed in Ferris 2002, 20; Frantzen 1995; Greenslade and Felix-Brasdefer 2006; Lalande 1982; Lee 1997). In this section, the studies that are discussed had findings which showed that there was no relationship between feedback given and accuracy development.

Sheppard (1992) carried out a short term study for a period of 10 weeks amongst 26 immigrant students from a USA college for whom English had been a foreign language. He experimented with two different types of feedback in a writing class based on narrative writing. The students were divided into two groups of which group A received coded error correction whereby the type of error and the location were indicated. This was followed by a conference with the teacher and then the students were asked to make a second corrected copy. Group B received feedback that dealt only with the content of student writing. This group received requests for clarifications that were written in the margin of the students' papers. Conferences with the teacher were based on these comments in this group. When the revised texts for the two groups were compared, there was no significant difference. Regarding accuracy, both groups improved significantly in the verb form, however there was no significant difference between them. The content group improved significantly in punctuation while

the correction group did not and the difference between them was significant. Truscott (1996) uses Sheppard's study to argue that if error correction was helpful, then the content group should have not shown any significant improvement at all.

I saw a few weaknesses in Sheppard's study. Sheppard does not give a report on the effects that the conferences may have had on the student writing, making it difficult to conclude whether improvement was based on the conferences, the feedback or both. Sheppard also made use of only 26 students who if divided into two groups show that the sample may actually not be a true representation of a larger group. The study also lacked a control group that did not receive feedback to make comparisons easier. The study having looked at the effect of different feedback mechanisms on accuracy makes it relevant to my accuracy development hypothesis.

A short term study similar to Sheppard's study (1992) was Semke (1984), which also studied narrative writing which was done amongst 141 German students from a German university for a period of 10 weeks. The study looked at free writing and a cloze test was also given. In comparison with Sheppard (1992), Semke made use of a pretest, more subjects and more groups that were assigned to four different feedback categories. Group one received comments on content only, group two direct corrections on errors, group three comments on content and direct corrections on errors, while group four had their errors coded and had to self-correct their errors and submit their rewrites one week later. Unfortunately Semke does not clearly show what the other three groups did after receiving feedback.

In the study by Semke (1984) there was no significant difference in accuracy among the groups similar to Sheppard's (1996) finding. The group that received comments on content only outperformed the other groups in fluency and in a cloze test, while the group which had direct corrections was actually inferior to all the other groups. A major weakness in Semke's study and that of Sheppard is that both lacked a true control group that did not receive any feedback. Truscott (1996) uses Semke's study to argue that if error correction gives positive results, then it should have been superior to all the other feedback mechanisms given but it is not. However, Truscott forgets that Semke's study did not have a true control group. Semke's use of various teacher correction mechanisms makes her study relevant to the current study, which also made use of different teacher feedback mechanisms.

Similar in findings to Semke's study discussed above, though done in the long term, was Robb, Ross, and Shortreed's (1986) study that was carried out amongst 134 Japanese college students for a period of one academic year. The study focused on more genres of writing compared to that of Semke (1984) and Sheppard (1992) and these were: expository, narrative and descriptive essays. The subjects wrote a total of five compositions each and were divided into four treatment groups: (a) direct correction; (b) indirect coded feedback; (c) indirect highlighted feedback (no codes); and (d) indirect marginal feedback. Content feedback was not given, as in the case of Semke and Sheppard discussed above. The group that received direct correction had their papers marked completely by the instructor. The correction covered lexical, syntactic and stylistic errors. The students had to rewrite their work and all that they had to do was copy out their composition incorporating the instructor's corrections. The second group received coded feedback whereby the type of error made by the student was indicated on the paper. The students received a guide on the codes on their papers in order for them to understand the instructors' marking. The indirect highlighted feedback group had their errors highlighted with a yellow test-marking pen. This differed from the coded feedback in that it indicated the places that needed editing but did not indicate the reason why it was marked. For the fourth group, which received indirect marginal feedback, the total number of errors per line was written in the margins of the student's paper. The subjects were all expected to rewrite their work based on the type of feedback that was given by the instructor. The data was analysed on a set of measures designed to assess accuracy, fluency and complexity.

The study found that there was no significant difference across the four treatment groups on any of the three measures investigated. For fluency, there was no improvement. The results suggest that methods of correction that are less time consuming may suffice since as per the finding of this study, there was no significant difference in students' writing, despite the fact that some very time consuming feedback mechanisms were used. This study however had one shortcoming: it, too, lacked a true control group that received no feedback at all, which would have made it easier to draw comparisons. The study relates to mine in that similar types of writing and feedback mechanisms were used and subjects had to rewrite their work. The present study, however, had a control group that received no feedback and also went ahead to find out the effect of feedback on individual error categories that Robb et al (1986) did not look at.

Related to Sheppard's (1992) and Semke's (1984) studies in terms of methodology and findings was that of Kepner (1991), who carried out a study amongst 60 intermediate college level Spanish students for 15 weeks. Kepner looked at journal writing (eight entries), unlike Sheppard and Semke who investigated narrative writing. The subjects were divided into two treatment groups of which one received comments on content while the other received comprehensive correction on surface level errors. The group that received comments on content experienced a whole language approach. The researcher responded to the students' message as a whole following four criteria: (a) personalisation (the student's name was used); (b) summarisation statement of the main point of the piece; (c) readers' reaction and evaluation of how the message was integrated; and (d) a question or suggestion for extending or improving upon the topic in subsequent entries. The second group that received correction had all their surface-level errors corrected. Reasons for correcting the errors were provided in a brief rule or note. Mechanical, grammar, vocabulary and syntax errors were corrected.

The study found no significant difference in accuracy between the two groups. It may therefore be concluded that error correction feedback here was not helpful in improving student accuracy. I saw a few shortcomings in the study such as that Kepner (1991) did not have any pretest measures and did not even check the first journal entry so as to know the level of proficiency of the students. Kepner's study did not have a true control group that did not receive any feedback at all, which could have made comparisons easier. The present study makes use of a pretest-posttest design, students in the experimental groups revised their work after revision and also a control group was used.

Cumming and So (1996) carried out a short term study for a six week-period amongst 20 volunteer adult ESL learners who were enrolled in academic courses at a Canadian University. The learners were considered to be of intermediate to advanced levels of proficiency. The subjects were required to write between 300 and 400 words on one of the four topics that were given at home and bring them to class for revision with the tutors. This differs from the present study in that the students in my study were to do their writing in class. Text revisions of the compositions were done for each learner in four individual tutoring sessions with the tutors and these were spaced one week apart. Each of the tutors was required to apply conventional error correction for one session in each language (English as a second language or mother tongue), while for the other two tutoring sessions the tutors were to make use of five thinking prompts with techniques of procedural facilitation. In the conventional error

correction the learners had their errors corrected in areas that were problematic at the end of their writing while in procedural facilitation the learners were helped both in the revision and writing procedure by use of the provided prompts. In total each learner received four sessions of which two were in conventional error correction and two were in procedural facilitation.

In the findings similar patterns were observed across both approaches and across the languages that were used. When procedural facilitation and error correction were compared statistically, it was found that there were no significant differences when either of them was used in the revision of student work. The study revealed that the students on their own noticed very few errors compared to the number of errors that their teachers noticed in their work. This is a sign that students need to be helped to notice their errors by someone more knowledgeable.

I felt that the study had some flaws, Firstly, the subjects did not receive any pretest to determine their level of proficiency and again bearing in mind that the subjects were volunteers from a larger group, they may not necessarily have been a representative of the whole group. Secondly, the tutors were not given a standard approach that was to be followed in any of the tutoring sessions and relied on their own experiences. The authors observed that there was great variation amongst the tutors' approaches showing inconsistency when feedback was given. Cumming and So (1996) admit that this is an important factor that should be considered in the future. Thirdly, the study also took place in the short term (six weeks) and so it becomes difficult to conclude whether the results would be similar in the long term. The study differs from the current study in that students had to write their drafts at home and they received help on their drafts from their tutors but in my study, students' writing was done in class and within a limited time. It is however similar to my study since revision of students' work was considered.

Another study that found no significant difference between the feedback mechanisms used was that of Fazio (2001). Fazio carried out an experiment amongst 112 Grade 5 students aged between 10 and 13 years in the context of two French language schools in Montreal, Canada for five months. Majority of the students spoke French as their first language while a minority spoke French as their L2. The subjects were in four classes and these were randomly assigned to three feedback groups giving each of the students an equal chance of belonging to any of the feedback groups. One group received

feedback on form, the second received feedback on content, while the third received feedback on form and content. The group that had form-focused feedback had all their errors in spelling underlined and then the correct form was given above the error that had been located. Students in the content-based group received a comment regarding the content of their work which was given at the end of their writing. Students that received both form focused and content feedback had their feedback delivered in the same way as to the other two groups Fazio used a similar methodology to that of Ashwell (2000) but differed in that in all groups, students received the same feedback mechanisms, unlike Ashwell where the feedback was rotational.

Fazio (2001) examined journal writing which was new to all but one student, which differs from the current study where students were familiar with the type of writing that they carried out. Accuracy of the journal writing was determined by calculating an error ratio: the number of errors in spelling as a proportion of the total number of occurrences of the structure. The findings of the study were that although there was no significant change in accuracy in grammatical spelling in any of the groups, the group receiving corrections maintained its error ratio throughout the study while the group receiving comments and the other that received a combination of both feedback types got worse. An explanation for these results by Fazio is that since a different teacher who was not a classroom teacher gave feedback; this could have had an effect on how the students reacted to the feedback.

I noticed three flaws in Fazio's (2001) study. The first was that Fazio did not conduct a pretest for her students to find out their proficiency levels which could have been a contributing factor to the findings. The second problem was that Fazio's study lacked a control group that did not receive any feedback at all, which could have made comparisons easier. The third problem was that the period for Fazio's study was also quite short for any significant improvement to have taken place. The study is relevant to my study as spelling errors were studied and underlining as a form of feedback was used in mine too.

Polio et al. (1998) carried out a short term study amongst 65 ESL students in a US university for a period of seven weeks. The study investigated whether revision led to better-written texts by considering journal writing. An experimental group and a control group were made use of in the study. The control group wrote four journal entries weekly and did not receive any feedback at all for a period of seven weeks. On the other hand, the experimental group did regular journal entry and was asked to

revise one of the two entries, had grammar review, and edited exercises. The experimental group received error correction on both the editing exercises and the journal entries and was expected to do revision. It is however not clear who gave the feedback. Only grammatical errors, inappropriate use of lexical items and punctuation errors were counted. In terms of accuracy, the study found that the students improved significantly when they revised their texts (short term) and from pretest to posttest (long term), but there was no significant difference between the control and the experimental groups, so the improvement cannot be attributed to the treatment. The study does not report on individual error categories, unlike the present study, which does so. The study is relevant to mine in that the subjects were expected to revise and rewrite their work after receiving feedback.

Ferris and Roberts (2001) carried out a study where three treatment conditions were investigated amongst 72 university ESL students. Five error categories (verb, noun endings, articles, word choice and errors of sentence structure) were studied. The subjects were assigned to three groups, whereby one group received no feedback; another had errors underlined while a third had errors underlined with codes. The subjects carried out editing after feedback had been given. The study found that the subjects that had their errors underlined corrected 60% of their errors while the group that had their errors both underlined and coded corrected 64% of their errors. The control group however only managed to correct 18% of their errors. Both groups that received feedback outperformed the control group that received none. This was attributed to the revision that was done by the students which was similar to Fathman and Whalley's (1990) finding. There was no significant difference between the two groups that received feedback even though the 'codes' group corrected a higher percentage of errors. This contradicts Greenslade and Felix-Brasdefer's (2005) findings, where coded feedback was associated with significantly better error correction compared to underlining of errors. The study is similar to the present study in terms of the methodology used and error categories studied, but differ in the type of population used. It is relevant to this study in that it supports the correction success hypothesis.

Schlig (2005) investigated how student writing could be improved amongst 35 third year intermediate Spanish students. Two groups were compared, whereby group A wrote 15 journal entries and the students received comments on their writing but grammar mistakes were not marked. Group B submitted 15 reading comprehension exercises and received feedback on content, organisation and accuracy. Group A had their errors identified using a 16-item code while group B that had their errors

identified using a five item code. Both groups carried out revision after receiving corrective feedback using the correction codes that had been provided. The findings of the study were that both groups corrected most of the errors; however group B was more successful in correcting errors compared to group A. Unfortunately the results were not statistically tested to find out if there was a significant difference between the two groups. This is a major shortcoming of the study.

I found the study to have other shortcomings too. To begin with the two groups did not write essays on the same topic, which may not make it easy to compare the two groups, as some topics may be considered easier than others. Secondly, the errors were identified differently as group A had a 16-item correction code which was very detailed and could have made it cumbersome for the students to follow as well as the researcher to give consistent marking, which Schlig (2005) admits. Group B had a shorter correction code of the items only, which could have been easier to understand by the participants. These differences could have affected the findings of the study. Another flaw of the study is that the students who had to re-type their essays after revision could probably make new errors that did not exist in the revised version, compared to Group B, which had their errors electronically indicated. Furthermore, Group B was allowed to use the editing tools in Microsoft Word in order to make corrections, which actually indicates that this group received aid from both the teacher and the computer programme. The study also did not have a pretest to find out the proficiency level of the participants either. The study does not therefore show whether the students were of the same or different proficiency levels to start with and this could have affected the results. The study is relevant to my study in that students received feedback and carried out revision but it does not say much since the data was not statistically tested.

Bitchener et al. (2005) carried out a study in New Zealand amongst 53 adult immigrant students who had just entered a post-intermediate ESOL programme for a period of 12 weeks. Three treatment groups were used that were based on whether participants studied full time for 20 hours per week, part time for 10 hours a week or part time for four hours per week. The full-time class received direct written correction feedback with explicit corrections above the underlined words followed by conferences with the researcher with the major focus on the three targeted errors: prepositions, past tense and the definite article. The 10-hour part-time class received direct written correction feedback

only, while the four hours per week part-time class received no correction feedback at all, but received feedback on the quality and organisation of their work for purposes of ethical satisfaction.

During the conference sessions, the participants were allowed to ask questions on the feedback given. The researcher took this opportunity to give any additional explanations and examples of errors that they made, while paying attention to prepositions, past simple tense and the definite article errors.

Each of the participants was expected to complete four writing tasks that were 250 words long based on informal letter writing within a time limit of 45 minutes which differs from the present study that used narrative and descriptive writing. The study found that the type of feedback provided did not have a significant effect on accuracy when the three targeted categories were considered as a group. However, the group that received direct feedback with individual conferencing feedback achieved greater accuracy when past simple tense and definite articles were used than was the case in the use of prepositions. In the study of Bitchener et al. (2005) study had various strengths, such as that it was well conducted with a control group and the period was long enough to determine the effects of different feedback on students' accuracy. The data was also statistically tested and hence can be more easily relied on. The study, however, neither tells us what the students did with the feedback that they received nor reports on the effect of the teacher conferences. A similar weakness was seen earlier in Sheppard's (1992) study where there was no report on the teacher conferences. This makes it difficult to determine what effects the teacher conferences had in the study. The study is relevant to my accuracy hypothesis in the overall and in individual error categories in that it dealt with the three targeted errors as a group and at the same time as individual error categories.

The studies described and commented on above (Bitchener et al. 2005; Cumming and So 1996; Fazio 2001; Ferris and Roberts 2001; Kepner 1991; Polio et al. 1998; Schlig 2005; Sheppard 1992; and Semke 1984) all suggested to a greater or lesser extent, that error correction feedback did not significantly improve writing development over time. There were no significant differences amongst the various feedback mechanisms that were used in the discussed studies. It was also noted that only Bitchener et al. 2005; Polio et al. (1998) and Ferris and Roberts (2001) made use of control groups which made comparisons easier. Due to differences in feedback mechanisms given in the various studies it is very difficult to make generalisations from the findings.

The studies that follow, however, reveal that feedback can show more positive results. These studies show that there is a positive relationship between error feedback and accuracy development. Both short term and long term studies are discussed, starting with the short term ones.

Fathman and Whalley (1990) carried out a study amongst 72 ESL students who were enrolled in intermediate ESL composition classes. The subjects were asked to write a composition based on a picture sequence in 30 minutes. The students were divided into four groups and the four treatments given were: (a) no feedback; (b) feedback on content; (c) feedback on grammar and content; and (d) feedback on grammar. Grammar feedback consisted of underlining of all errors while content feedback consisted of general comments that were not text specific. The group that received both content and grammar feedback received both underlining of errors and comments on the text. All the subjects were asked to revise and rewrite their work after feedback had been given.

The findings of the study were that all of the students improved in grammatical accuracy for both groups that either received grammar feedback only or grammar and content feedback and this was significant for the former group. It was also noted that all groups showed an increase in the number of words in the revised and rewritten scripts. The no feedback group actually had the highest increase in number of words written in the rewrites, showing that they wrote more fluently in their rewrites. Fathman and Whalley (1990) had their data tested statistically, which is one of its strengths. The errors were not specified as the errors were marked comprehensively so it becomes difficult to tell whether there was improvement on any specific errors. The study was carried out in the short term since the writing that was assessed was revisions from draft to draft and so these results support my error correction success hypothesis but say nothing about longer term accuracy development.

Another short term study was that carried out by Lee (1997) amongst EFL College students in Hong Kong which sought to investigate ESL learners' performance in error correction in writing. The subjects were given an error correction task that contained 20 errors that were implanted in a text. The subjects were divided into three groups of which one group had all its errors underlined, the second had errors indirectly indicated as all the lines that were error free were ticked. Students in this group

therefore had to try and find errors in the lines that were not ticked. The third group did not receive any form of feed back at all.

It was found that the students whose errors were underlined had the highest mean correction scores compared to the other two groups and this was significant. This result is similar to that of Chandler (2003), discussed in this section where the most explicit correction led to the highest correction compared to the other feedback mechanisms used. On the other hand, this finding differs from that of Lalande (1982), where indirect feedback was superior to the direct feedback. However, the studies differ in both the period and the type of writing that was investigated. The study differs from the present study in that in Lee (1997) the errors studied were not those made by students in a naturalistic situation as the errors were implanted in an existing text, while the present study dealt with errors made by students in their own writing which gives a clearer picture of the errors that students make in their own writing. Lee admits that in real life error correction may be more complex than it is shown here. Lee's study, however, supports the accuracy development hypothesis of this study.

Hong (2004) carried out a study amongst 119 students enrolled in an ESL course at Brigham Young University. The study examined five different error categories which were the same as those studied in Ferris and Roberts (2001), discussed above. Three treatments were used: coded feedback, non-coded feedback and no feedback. The subjects were given an in-class essay on the same topic and the scripts were self corrected in 20 minutes two weeks later after they had received their various feedback mechanisms. The findings were that the control group corrected the lowest mean number of errors followed by the non-coded and finally the coded group. There was a significant difference between the two groups that received feedback and the control group but there was no difference between the two experimental groups. The same finding is in Ferris and Roberts (2001), where the experimental groups outperformed the control group. There is a clear indication that students did benefit from the teacher correction feedback given. The study also has similar findings to that of Fathman and Whalley (1990), where subjects who received feedback and revised their work improved more compared to those who received none, but the two studies differ in the type of writing used in that Fathman and Whalley used picture composition, while Hong used a diagnostic essay.

Greenslade and Felix-Brasdefer (2006) compared coded and uncoded feedback which was similar to Hong's (2004) design although in the former there was no control group. The subjects were 21 students of Spanish as an FL of Intermediate/advanced proficiency at a university in the USA. Two compositions were given in the study. In the first composition, errors were underlined while in the second errors were underlined and coded. Syntactic, lexical and mechanical errors were focused on and these totalled up to 19 error types. For both compositions, the subjects received feedback on their scripts and after feedback had been given they were asked to do corrections in 20 minutes. In order to correct the second composition, participants were given a list of codes used in the feedback. The findings were that accuracy improved in the revised drafts under both conditions, but the coded feedback enabled the students to improve in accuracy more than the uncoded feedback. As with the other studies just discussed this particular study was done in the short term, in that students' compositions were studied from one draft to the next draft and only two compositions were considered, which therefore means that we should be cautious about drawing conclusions from it about the value of coded as opposed to uncoded feedback, but because it focuses on this distinction, this research is directly relevant to my study and supports the correction success hypothesis.

The four studies discussed above (Fathman and Whalley 1990, Hong 2004, Greenslade and Felix-Brasdefer 2006, Lee 1997) all reveal that students' writing improved when students rewrote their work after revisions. Findings for studies that made use of coded and uncoded feedback (Hong 2004, Greenslade and Felix Brasdefer 2006) showed that coded feedback led to more improvement in students' writing compared to uncoded feedback. Even though the type of writing and methodology differed in the above mentioned studies, the findings are similar. These studies are however short-term ones, and so they say nothing about whether the same results would apply in the long term. The studies that follow were done in the long term and are those that reveal that feedback showed positive results.

Lalande (1982) carried out a study amongst 60 intermediate students for one semester. The study made use of a pretest essay and two draft essays. Two treatments were used and these were direct and indirect feedback. The direct group had their errors corrected directly and rewrote their work while the indirect group had correction codes, rewrote their work and had an error awareness sheet. The study found that the students who received indirect correction made significantly fewer errors compared to

the direct correction. Guenette (2007) criticizes Lalande's pretest results that showed significant differences in the students' writing abilities in that it is not clear how these results were arrived at. Lalande's study also lacked a true control group that did not receive any feedback at all. This could have made comparison easier as to whether some feedback was superior to no feedback. In fact, Chandler (2004, 347) says Lalande's study 'proved nothing about the efficacy or inefficacy of corrective feedback' if a control group is lacking. This is one of the reasons why the present study made use of a control group that did not receive any feedback. The feedback mechanisms used are similar to those in my study.

Another study that made use of direct and indirect feedback mechanisms was that of Frantzen (1995). Frantzen carried out a study amongst 67 students that were enrolled in four sections of an intermediate Spanish content course for 15 weeks. Two groups were used in the study of which one was a grammar supplement group while the other was not. The grammar supplement group had grammar exposure that involved subjects in this group getting a brief daily grammar review together with all errors corrected on their written work (direct feedback). For the in-class essays, the non-grammar class had their errors indicated by circling or underlining, but the correct word was not given (indirect feedback).

One finding of the study was that neither group improved in fluency over the semester. This finding contradicts that of Chandler (2003), where subjects improved in fluency. A study by Frantzen (1995) revealed that both groups showed significant improvement in grammatical accuracy, though the grammar supplement group showed significantly greater improvement compared to the non-grammar supplement group. This finding contradicts that of Ferris, Chaney, Komura, Roberts and Mackee (2000) and Lalande (1982), who found that the indirect feedback group performed better than the direct feedback group. These three studies again share one weakness and that is they all lacked control groups. Frantzen's study seemed to be problematic because it had several treatments given at the same time, for example, direct versus indirect feedback, grammar review versus no grammar review, revision versus no revision and grading on accuracy versus no grading on accuracy. It becomes difficult to conclude which of these treatments had an effect on student writing. The study is relevant to this study in that it supports the accuracy development hypothesis.

The study by Ferris et al. (2000) can be compared to that of Lalande (1982) as it also made use of direct and indirect feedback. The study looked at the effects of indirect and direct feedback both in the short term (from one draft to the next) and in the long term (from the beginning to the end of the semester). All students revised their work after teacher error feedback had been given. The findings of the study were that in the short term, the direct feedback group outperformed the indirect feedback group. However, by the end of the semester (long term), the indirect feedback group's error ratio had reduced much more than that of the direct feedback group. This finding is similar to that of Lalande (1982), where the indirect feedback outperformed the direct group. The two studies however have one shortcoming in that none of them had a control group to allow comparisons. The study by Ferris et al. (2000) is relevant to mine in that the methodology used was similar and it supports the correction success hypothesis of this study.

Another long term study that supports feedback was that of Ashwell (2000), which was carried out for a period of one year. Ashwell carried out a study amongst 50 Foreign Language students in a Japanese University who were enrolled in two writing classes. The study investigated whether content feedback followed by form feedback was the best method of bringing about improvement in writing. All the subjects were expected to draft their essay twice (D1 and D2) before writing a final version (D3). A total of four assignments written out of class were written in the year. The study made use of four types of feedback: (a) content feedback followed by form focused feedback; (b) form feedback followed by content feedback; (c) a mixture of both form and content feedback; and (d) no feedback. Participants were assigned to one of the four groups. The form-focused feedback was indirect and was given by underlining or circling grammatical, lexical and mechanical errors indicating omissions. The errors focused on were: (a) lexical choices; (b) articles/determiners/plurals; (c) spelling; (d) prepositions; (e) punctuation or sentences and clause boundaries; (f) clause boundaries; (g) agreement; and (h) verb/tense. Content feedback addressed organisation, paragraphing, cohesion and relevance.

The findings of the study concerning accuracy revealed that the three groups that received feedback improved significantly in accuracy between D1 and D3, and the group which received a mixture of form and content feedback made the largest gain. It was however noted that there was no significant difference between the groups that received feedback. The no feedback group got slightly worse. This is an indication that even though the groups that received feedback do not differ from each other,

feedback seems to have a positive effect on student writing. The study is supportive of error correction feedback. I noticed one shortcoming in the study: the study had the teacher being the researcher at the same time, which could affect its validity. Ashwell studied similar error categories to those in my study and also looked at overall accuracy and so is relevant to my accuracy development hypothesis.

In a more recent study that compared direct and indirect feedback mechanisms Liu (2008) studied 12 first-year students in the USA who were involved in a course whose aim was to improve students' writing. The participants met for 50 minutes weekly for a period of 16 weeks and were taught by the same teacher researcher. Three genres were studied: a) a rhetorical analysis; b) an argumentative essay; and c) a reflective essay. All the genres were new to the students and the teacher had to demonstrate to the students what was expected of them before they wrote the essays. This was quite different from the present study, which looked at narrative and descriptive genres which were familiar to the students and so no demonstrations were done before the writing.

In the direct feedback group, students' errors were underlined and corrected while in the indirect feedback, errors were only underlined. Even though Liu's study indicates that the two groups received direct and indirect feedback, it is indicated that this was only done for the first essay while for the other essays, the participant received underlining and or description on latter essays. This was an inconsistency in the study as the feedback mechanisms were not used as they were supposed to be. The present study tried to eliminate such inconsistency by ensuring that each feedback group received one feedback mechanism throughout the experiment, making it easier to draw conclusions on the effect of each of the feedback mechanisms used.

Liu's (2008) study found that both groups corrected most of the errors when they wrote a different draft, which is similar to Ferris and Roberts's (2001) finding, where the two groups that received feedback outperformed those that did not receive any. Liu found that those who received direct feedback outperformed those who received indirect feedback concurring with Chandler's study (2003), where direct feedback was the best for correcting errors. Liu's results however were not tested for significance, making it difficult to conclude whether the improvement in the students' writing was significant or not. The study also lacked a control group.

Chandler (2003) carried out two studies, both reported on in the same article. The first addressed the question as to whether to give feedback or not while the second investigated the effects of different teacher responses to error on the correctness both of revision and subsequent writing. In the first study, the subjects were drawn from an American conservatory and it took place over 10 weeks (a semester). The study had an experimental and a control group of which both groups received indirect feedback (errors underlined). Even though Chandler calls one a control group, it was not a true control group but rather a different treatment group since this group received feedback too. The experimental group revised each of the assignments, correcting all the underlined errors before handing in the next assignment. On the other hand, the control group did revision of the underlined errors at the end of the semester. In total all the subjects wrote five assignments in autobiographical writing. Comprehensive error correction was done on all the students' texts. For accuracy, the total number of errors made per 100 words was calculated, while fluency was measured by looking at the amount of time it took to write each assignment.

Chandler (2003) found that the experimental group showed a significant increase in error correction while the control group did not and so accuracy of student writing improved significantly where subjects were asked to carry out revision. The study thus also suggests that if students are given feedback and do not revise their texts then that is as good as giving no feedback at all. This result can be compared to that of Fazio (2001), where students had their errors corrected but they were not asked to revise and despite the feedback that was given, none of the groups improved in their spelling accuracy. Fazio found out through interviews conducted that many of the students did not look at the feedback given. On the measure of fluency, Chandler (2003) found that both the control and experimental group showed a significant increase in fluency, concurring with Fathman and Whalley's (1990) study, but contradicting that of Frantzen (1995), where none of the groups showed improvement. Chandler's study differs from my study in the measurement of fluency in that the length of writing was not specified while in this study it is. This study looked at six error categories while Chandler (2003) studied about 23 errors, but did not give a report on whether the students improved in these individual categories or not. One of the strengths of Chandler's (2003) study is that it had a control group even though this can be challenged in that the control group received some kind of feedback. The study is relevant to my study also because it supports the correction success, fluency and overall accuracy development hypotheses.

The second study by Chandler (2003) was carried out on the effects of various kinds of error correction amongst students in the same ESL writing course as the one above but in a different year with different students. The subjects wrote five assignments within the semester and were expected to do revision after the teacher had given feedback before they wrote the next assignment. Chandler's previous study marked errors by underlining comprehensively but this second study involved four different types of feedback: (a) correction; (b) underlining with description; (c) description of type only; and (d) underlining. Those who received correction got the correct versions of the errors made. In underlining with description, errors were underlined and an instruction of the type of error made was written in the margin. For description only, the type of error made was shown in the margin without pointing out where the error was, while with underlining, errors were simply underlined. In this study, each student received the four kinds of feedback in different orders (rotational). Accuracy was measured in terms of the number of errors made per 100 words on the revised and subsequent scripts, while fluency was measured in terms of the time spent writing each assignment.

The findings of the study were that student writing improved significantly over the semester in both fluency and accuracy. Correction was found to lead to the greatest increase in accuracy both for the revised texts and subsequent writing, while underlining and correction together led to more accurate writing on the next assignment compared to the other two feedback mechanisms.

The procedure was quite different from my study in that in my study each student received only the one feedback type that was assigned to them, while in Chandler's (2003) study, the students received each of the feedback mechanisms that were given in a rotational way. Over time it is difficult in Chandler's study to conclude what effect each feedback mechanism had on the student writing. Both studies discussed by Chandler above looked at autobiographical writing while my study looks at descriptive and narrative writing, though as with Chandler (and unlike Liu (2008), for example), all the subjects wrote on the same topics. This makes it easier to make comparison as the level of difficulty should be the same for all the students. Chandler's study supports the present study's hypotheses: correction success, fluency and accuracy development in students' writing.

Similar to Chandler's (2003) second study in terms of methodology discussed above is Delgado (2007). The participants were second semester students in one university in the USA. The students remained in their intact classes but were randomly assigned to three treatment groups namely: coded, uncoded and no feedback. Delgado's study differs from that of Chandler in that it made use of a control group that received no feedback at all. Students wrote on specified topics and received the different feedback mechanisms in rotational order for a period of 14 weeks. Accuracy was measured in terms of number of errors per 100 words in the same way as in Chandler (2003). The study found that both coded and uncoded feedback led to significant improvement in student writing. A comparison made between the two feedback mechanisms showed no significant difference. The study lacked a pretest and also a posttest that would tell whether the students improved because of the treatment given or something else. Just as in Chandler's study, the study did not look at specific error categories so it becomes difficult to tell which error categories were improved on and which were not.

Ferris (1997) used a different approach compared to the other studies discussed above in that the study made use of comments only. She carried out a study amongst 47 ESL students enrolled in three sections of an advanced University ESL composition course that were taught for two consecutive semesters. This study made use of personal, narrative, expository and persuasive writing. It therefore looked at a wider range of genres than any of the other studies reviewed here. The type of feedback given was that subjects received both endnotes and marginal comments in their first drafts and revised their texts. The findings were that both types of comments led to successful revision. For the comments that were given on grammar and mechanics, endnotes led to more corrections than marginal notes. It was also found that some comments led to less successful revision probably because they were either ignored or deleted. Though the study made use of endnotes and marginal comments, it is relevant to my study in that the students were expected to do revision after receiving feedback and it supports my correction success hypothesis.

The studies reviewed in this section reveal mixed findings. Some show that regardless of the type of feedback given, students did not improve in the accuracy of their writing (Kepner 1991; Polio et al. 1998; Robb et al. 1986; Semke 1984; and Sheppard 1992). Others reveal that revision led to correction success (Chandler 2003; Fathman and Whalley 1990; Ferris 1997; Greenslade and Felix-Brasdefer 2006; Hong 2004; and Liu 2008). Different feedback mechanisms had different effects on accuracy as

some showed that indirect feedback led to more improvement in writing compared to direct feedback (Ferris et al. 2000, Lalande 1982,) while in some direct feedback was superior (Chandler 2003, Frantzen 1995, and Liu 2008). Unfortunately some studies, such as those of Schlig (2005) and Liu (2008), did not test results statistically, making it difficult to assess whether there was improvement or not. The discussion in this section has also shown how different the studies are from one another, which makes it very difficult to interpret the different findings and come to a generalisation about whether correction feedback mechanisms lead to improvement or not. Such problems will be considered in more detail when I move from descriptive review to critical review in the next major section (2.2), but the following section concentrates on the narrower issue of research on the effect of feedback across specific error categories.

2.1.3 Effect of feedback across error categories

This section focuses on the effect of different feedback mechanisms on specific error categories as explored in various studies, most of which have been more generally discussed above. Ferris (2002) groups errors into 'treatable' and 'untreatable errors'. Treatable errors are those that can be corrected by the use of a rule. Examples of treatable errors include those of tense, spelling and capitalisation. Untreatable errors are those that are not rule governed, such as errors of word choice and sentence structure. Some studies reveal that certain errors are easily noticed by students after error feedback has been given, while others show that even without teacher error feedback, some errors are still noticed by the students. Below is a discussion of some studies that looked at the effect of different feedback mechanisms on certain error categories.

Lee (1997) gave students a correction task which contained surface and meaning errors. A newspaper article was made use of in the study and it had 20 errors implanted in it. These errors were further categorised by four independent judges into surface and meaning errors but some did not fall into these two categories and were referred to as 'non-classified items'. Lee does not tell us which of the identified errors falls into each of the three categories just mentioned. The study found that the student's ability to categorise the errors was quite limited and that while some of the errors, such as spelling, prepositions, articles and punctuation were easily noticed by the students, others such as verbs, adjectives, possessives and part of speech were difficult to notice. The results revealed that all

feedback conditions did significantly better at correcting surface errors. The finding is similar to Liu (2008), who found that morphological errors were easier to correct compared to semantic errors (word choice). Lee's study dealt with errors that were implanted in a given text and so is rather different in this respect to my own study, which deals with errors made by students themselves while writing in a more naturalistic situation. Thus full comparison between the two studies is not possible but the subjects used by Lee were similar to those used in my study in that they were low academic performers. Lee's study is also relevant in that five of the error categories used were also studied in my research, which had additionally the wrong word category. The study is thus relevant to my investigation of my accuracy development hypothesis with respect to individual error categories.

Frantzen (1995) examined 12 grammatical categories. In this study, two groups were used of which one received direct feedback and grammar exposure while the other received indirect feedback with no grammar exposure. As regards the effect of this treatment on error categories, it was found that both groups improved as regards tense, aspect and overall grammar usage. It should be noted that even the group that received only content feedback improved in grammar.

Fazio (2001) gave corrections, commentaries and both corrections and commentaries to her participants. The spelling error category was studied and it was found that the group that received corrections only maintained the same error ratios throughout the study. On the other hand, the groups that received content feedback and both content and corrections made more spelling errors by the end of the study. Fazio explains this strange finding as a result of the students' reaction to the feedback as it was given by a different teacher rather than their own. She also says that it could be because spellings in the French language are a difficult category to improve in easily within a short time and is quite challenging even for native speakers. Fazio's findings contradict that of Lee (1997), where amongst the error categories studied, the spelling error category was the easiest to correct although this discrepancy should be treated cautiously as the settings for the two studies were quite different. In Lee's study the errors were implanted in a text while those in Fazio's study were made in the students' own writing. Fazio's study is relevant to my study in that the spelling error category was studied in mine.

Ferris and Roberts (2001) studied five error categories: (a) verbs; (b) nouns; (c) articles; (d) word choice; and (e) sentence structure. The number of error categories studied was more than the three (preposition, past simple and definite article) studied by Bitchener, et.al. (2005) but a lot fewer than Chandler (2003) and Lee (1997) who had 23 and 21 types of errors studied respectively. The study compared coded feedback to uncoded feedback and had a control group that received no feedback at all. The study revealed that overall, the highest mean number of errors was observed in verbs followed by sentence structure, while the article was the least problematic. When the feedback and non-feedback groups were compared, there was a wide gap in the success ratios of error correction. The feedback groups corrected 64% of the verb errors while the control group only managed to correct 14%, for the noun error category, the feedback groups corrected 79%, while the control group corrected only 8%, for the sentence structure errors, the experimental groups corrected 52% while the control group corrected 23% and lastly, regarding the word choice error category, the experimental groups corrected 63% of the errors while the control group corrected 31%. It can be noted that the control group subjects performed relatively better in correcting untreatable errors compared to their treatable errors. Although there was a statistically significant difference in feedback groups' ability to correct treatable errors and 'untreatable' sentence structure errors compared to the control group, they were not significantly better at correcting the 'untreatable' word choice errors. Ferris and Robert's (2001) study is relevant to the current study in that it looked at the effect of different feedback mechanisms on selected error categories, with the current study including similarly the verb, article and wrong word choice categories.

In a similar study to Ferris and Roberts (2001), Hong (2004) studied the same type of error categories. It was found that for all the subjects, the sentence structure, the wrong word and verb error categories had the highest mean corrections while the article and noun ending categories had the least. The article errors remained constant for all the feedback groups. Hong explains that this article finding could be attributed to the Asian background of many of the subjects who have problems with article usage due to a difference in grammatical structure of their native language. The study also found that the two experimental groups significantly outperformed the control group in all error categories, similar to what Ferris and Roberts (2001) found. Hong's study is relevant to my study as a similar methodology was used and the same error categories studied.

Schlig (2005) studied five error categories: nouns, verbs, articles, lexical and sentence structure errors, similar to those studied by Ferris and Roberts (2001) and Hong (2004). Two groups were used in the study of which group A received a 16-item correction code while group B received a 5-item correction code. The results showed that for group A, the noun ending followed by article error category had the highest correction rate while the verb error category had the lowest. Group B also corrected noun endings best followed by lexical errors and the least corrections were in the article category. Unfortunately Schlig did not test the data collected for significance so not much can be concluded. However, for the three studies overall (Ferris and Roberts 2001, Hong 2004, and Schlig 2005), it can be concluded that some errors were easier to correct than others.

Bitchener et al. (2005) studied the effect of different feedback mechanisms on three error categories: prepositions, past tense and the definite article. The study made use of direct, explicit feedback and a five minute student conference, direct, explicit written feedback and no correct feedback. It found that the direct, explicit written feedback combined with conference feedback had a significant effect on both past tense and the definite article. The errors were some of those that were studied in the present study and so it is relevant to the accuracy development hypothesis on individual error categories.

Liu (2008) investigated three error categories: morphological, semantic and syntactic. Semantic errors were more difficult to correct than morphological errors and the direct feedback group corrected more semantic errors compared to the indirect feedback group.

The three studies that looked at the same error categories (Ferris and Roberts 2001; Hong 2004; Schlig 2005) can be compared in terms of findings. It was found that the sentence structure, wrong word and verb errors were easily corrected in Ferris and Roberts (2001) and Hong (2004), while, for both studies, the noun and article errors were the least corrected. In fact, in Hong's study, the article error category remained constant for all the feedback groups throughout the study. On the contrary, the noun ending error category was the one most corrected in Schlig (2005). For other researchers, such as Lee (1997), spelling errors were reported to have been noticed easily by the students but in Fazio (2001) the subjects made more spelling errors by the end of the study. In two of the studies reviewed (Lee 1997 and Schlig 2005), the verb error category was found to be the most problematic. Again, regarding students' success in correcting different categories, it is not easy to make comparisons as different

studies used different methodologies as well as different error categories. In conclusion it can at least be stated that students find some errors easier to correct than others, which is an indicator that different error categories should be treated differently.

As my study made use of different correction mechanisms, it was important to find out what the students' opinions were about these mechanisms and findings on this matter are discussed in the following section.

2.1.4 Students' evaluation of teacher error feedback

In this section, studies that have investigated students' opinions on teacher error correction have been reviewed. Research that has been done on students' preference for error correction seems to deliver similar findings so that at least in this area of study not many discrepancies exist as students have shown a definite overall preferences for error correction, as exemplified by some of the studies described earlier, as well as others, which are discussed from this point of view. These issues are relevant to my study because it, too, gathered information on students' opinions of different types of feedback.

At the end of the end of the study by Lalande (1982), data was collected from the students by the use of a questionnaire to ascertain whether they considered the respective treatment they had received to be desirable in the development of their writing skills. The finding was that 72% of the respondents in the group who received direct feedback and 86% of those in the group who received coded feedback felt that they improved significantly by the end of the course. This is in agreement with the findings of the study where the coded feedback group made significantly fewer errors compared to the direct feedback group by the end of the study. There is an indication that the students were not discouraged upon receiving feedback as they felt that it helped them improve in their accuracy.

A survey was carried out by Hedgcock and Lefkowitz (1994) amongst 110 ESL and 137 EFL university students who were all considered being of elementary level in English. A 45-item questionnaire that was based on a 6-point Likert scale elicited information from the students on issues such as: the type of feedback that the students found least and most helpful, level of enjoyment, continued motivation to improve L2 writing skills, self-assessment of L2 writing proficiency,

preferences for various feedback types on intermediate and final drafts and beliefs about the benefits of specific teacher intervention behaviour. On the preferred type of feedback, majority of the students (60%) preferred written feedback combined with writing conferences, while 30% preferred written feedback only and 10% of the subjects preferred verbal feedback only for both ESL and EFL combined. When the two groups were compared separately on the same question, the ESL subjects showed a greater preference for the combined feedback type (65%) compared to the EFL subjects (56%).

Both EFL and ESL writers' responses showed a strong preference for teacher correction if formal text features like lexical and grammatical errors were corrected. There was a higher preference among ESL writers for comments on idea development and sequencing compared to the EFL writers. The ESL writers also expressed a strong preference for teachers' comments on writing styles. Of moderate concern were comments on mechanical features (punctuation, capitalisation and spelling). Both groups expressed preference for the use of correction symbols on the part of their teachers, this finding thus linking with others showing preferences for coded feedback.

Ferris and Roberts (2001) also sought students' preferences through the completion of a grammar knowledge questionnaire at the end of the study. Amongst the questions asked, students were to give information on their prior grammar instruction, the types of grammar problems they had in writing and their error feedback preferences. On the question of error preferences, all the respondents wanted error correction, while 2% of the respondents wanted only the most serious errors corrected. Those who wanted their errors circled formed 19% of the responses, 31% of the respondents wanted all their errors corrected by the teacher while 48% of the respondents wanted their errors circled and labelled by error type. This is a similar finding to that of Hedgcock and Lefkowitz (1994), discussed above where students preferred coded feedback to the other types of feedback.

After carrying out a study on the efficacy of various kinds of error correction feedback for improvement in accuracy and fluency on student writing (as discussed above), Chandler (2003) sought the views of 21 students on the various feedback mechanisms that were used in the study. The survey sought information on which approach they found easiest to correct their mistakes, which approach was the easiest for them to see their mistakes, from which approach they learnt most, which approach

was likely to help them to write correctly in the future, which they liked most, whether they understood the feedback mechanisms that were used and whether they felt discouraged or not from the feedback that was used. The same questionnaire was adapted by Liu (2008) for the first five items and a comparison of the two studies' findings is done here. In both studies students found underlining and describing to be the easiest to correct their errors, though for Liu's study the number of respondents who preferred full correction was the same as those preferring underlining and describing. Thus again, for both studies, underlining with description (coded feedback) was seen as the easiest way to see the kind of mistakes made, the way that the respondents said they learnt from most, and the way that most felt would help them write correctly in the future. Students in the two studies seem to share same views even though they were from different settings. Chandler's (2003) study tried to find out whether students understood the type of feedback used and full correction received the largest response as having been understood. Lastly the respondents were asked whether they felt discouraged when they received feedback and for all the feedback types, all the students except one did not feel discouraged when they received feedback except for one student from the direct correction. These findings concur with those of Lee (2004), which revealed that students showed highest preference for direct correction as they felt that it was the easiest way to see their errors.

Lee (2004) carried out a study on teachers' perspectives, practices and problems regarding error feedback in Hong Kong. A questionnaire was administered to the students to find out about their perceptions, beliefs and attitudes regarding error correction. The questionnaire included statements on teachers' correction practices and students' own belief and attitudes about error correction. The questionnaire was both in Chinese and English. In the survey, students' attitudes on comprehensive versus selective error feedback, direct versus indirect error feedback and the use of error codes was sought.

The survey found that 83% of the correspondents said that they wanted their teachers to mark all errors. In the follow-up interview, the main reason given was that it was easy for the students to locate their errors. With regard to direct versus indirect feedback, 76% of the students wanted teachers to give direct feedback on all errors while 22% of the students wanted teachers to give direct feedback on some of the errors but it does not specify what some of these errors were. The study does not indicate what the other 2% preferred. In response to the question on whose responsibility it was to correct

errors, 45% of the students felt it was the teachers' job while 55% of them thought it was the students' job. This study thus shows in various ways that on balance students expected some form of feedback.

In her dissertation, Hong (2004) investigated students' concern for grammar correction and preferences and found that the majority of the students (64.4%) were concerned about their grammatical accuracy. Most of them (53%) preferred coded feedback which is similar to that of Ferris and Roberts (2001), where nearly a majority of the students (48%) preferred coded feedback. For the other types of feedback, Hong found that 21% of the students preferred direct correction, 19% preferred non-coded feedback, and 4% said they did not want any form of grammar correction while 3% wanted global correction. Nearly all the students (96%) wanted their grammatical errors corrected.

Greenslade and Felix-Brasdefer (2006) sought to investigate students' perceptions of the effectiveness of coded vs. uncoded feedback. Data was collected by the use of a student questionnaire. For the uncoded group, 11% of the students were 'very satisfied' with the feedback, 84% were 'somewhat satisfied' while 5% (one student) was 'not satisfied' with the correction at all. Of those students who received coded feedback, 42% reported being 'very satisfied' with the correction, 58% were 'somewhat satisfied' and none of them said they were 'not satisfied at all'. The results also revealed that the students preferred if symbols were used when their errors were corrected. A majority also expressed a preference for feedback if it was coded rather than uncoded as 15 out of the 19 students said that they understood better what was expected of them. They felt the corrections could be made easily for the mistakes had been clearly outlined.

In a recent survey amongst intermediate and upper intermediate students who were studying English for Specific Purposes (ESP), Anusiene Kaminskienne and Kavalianskienne (2009) made several findings on students' opinions on teacher correction feedback. Participants were asked if they preferred immediate teacher correction of spoken errors and for the first year students, 62% agreed, while for the second year students, 90% agreed. Participants were also asked if they found it hard to notice their own mistakes and for the first year students, 47% agreed that they found it difficult while for the second year students, 70% agreed. The survey also required that the students indicate whether they wanted every mistake in their writing to be corrected. For the first year students, 81% agreed while for the second year students, 95% agreed. On whether the teachers' individual correction of students' written

mistakes is useful for learning ESP, for the first year students, 88% agreed and all the second year students agreed. The results show that a very large majority of the students felt that teachers' individual correction was useful.

Even though learners seem to want to have errors corrected, Truscott (1996) is of the opinion that learners must not be given what they want. However, in response to Truscott, Ferris (2004) argues that students' desire for error correction cannot just be dismissed. Hedgcock and Lefkowitz (1996) had earlier argued that students' perceptions about error feedback and the type of feedback that they get from the teacher may actually influence the quality of work written by students, which in turn is likely to influence the students' writing proficiency.

The studies discussed in this section reveal that some students preferred direct correction as they felt that this helped them see the errors they made and so correct them (Chandler 2003; Lee 1997; and Liu 2008), while in others coded feedback was preferred by most of the subjects (Hedgcock and Lefkowitz 1994; Ferris and Roberts 2001; and Greenslade and Felix-Brasdefer 2006). In conclusion, it can be said that the vast majority of the subjects wanted some form of feedback and were not discouraged when they received feedback. The findings on students' preferences are relevant to my study as it too sought to investigate students' opinions on teacher correction feedback that they received in their writing.

The following section gives some general critical comments on the studies that have been reviewed in this chapter and considers how difficult it is to make comparisons across these studies.

2.2 Some general critical comments

In this section, I focus more on critique of the studies that have been reviewed above, and also on the field as a whole. This section discusses various parameters that were used in the studies reviewed in this chapter and tries to show how much variation exists concerning them. The parameters discussed are: the proficiency level of subjects, classroom settings, age of subjects, number of subjects, type of subjects, period within which the studies were carried, the type and length of writing, error categories, type of feedback given and data collection.

Many of the studies have been carried in intact classes where students have already been enrolled in terms of criteria set by the institutions. Guenette (2007) points out that in such settings, proficiency levels of students can vary widely. It is therefore wrong to assume that such students would be at the same proficiency level and there is a need to determine the proficiency level of students before carrying out any study. Many of the studies that have been reviewed lacked a measure of the initial proficiency level of their students. Some studies made use of pretests and posttests (such as Polio et al. 1998; Robb et al. 1986; and Semke 1984) but some do not report on the results of the pretest, such as Semke (1984), who reports only on the posttest scores. Guenette (2007) criticizes Lalande (1982), who used a pretest in his study and reported that his participants were on the same proficiency level, but it is not clear how this result was arrived at. In some studies, such as that of Ferris and Roberts (2001), that had a pretest, it was found that there were considerable differences in the fluency and accuracy levels between the three groups used in the study, but this could be expected since the subjects were from two different groups (immigrant and international students) and the researchers did indicate that these differences could have affected the findings. Other studies made use of holistic ratings (e.g. Fathman and Whalley 1990) to say that the participants were of the same level of proficiency, but Polio et al. (1998, 52) point out that "…holistic scales are not fine-grained enough to use on a homogeneous population, that is, on students who are placed in the same class." On the other hand many studies did not make use of any pretests at all (Bitchener et al., Young and Cameron 2005; Chandler 2003; Cumming and So 1996; Fazio 2001; Ferris 1997; Greenslade and Felix-Brasdefer 2006; and Kepner, 1991). For example, Bitchener et al. (2005) divided the participants into groups based simply on whether they were full time or part-time students in the course. Cumming and So (1996) used 20 volunteers who were adult ESL learners involved in an intensive ESL writing course, but being volunteers, it cannot be assumed that these participants would be representative of a particular population. Fazio (2001) on the other hand divided the participants in the study into two groups depending on whether they were speakers of French as a first language or of other minority languages. These studies do not find out the subjects' proficiency levels at the beginning of the studies and Guenette (2007) is correct in claiming that if the level of proficiency of participants is not measured well at the beginning of a study, it becomes difficult to tell if improvement or no improvement is attributed to the proficiency level of the student or to the feedback given.

The ages of the students in the studies vary greatly, although in all the studies reviewed here, only a few studies give the ages of their subjects such as, Fazio (2001), the subjects' ages ranged from 10 to 13 years and in Lee (1997), the ages ranged from 19 to 23 years. Most studies reviewed here do not give the ages of the subjects but report that the students were enrolled either in college or university courses (Chandler 2003; Cumming and So 1996; Fathman and Whalley 1990; Ferris and Roberts 2001; Greenslade and Felix-Brasdefer 2006; Liu 2008; Semke 1984). Age plays a major role in the learning of a second language (Brown 1994) and if studies are to be compared, ideally, those that have been carried out on subjects in the same age bracket should be compared.

Another aspect related to subjects is that of the number of subjects that were used in the reviewed studies, and here too there is wide variation. Some studies had as few participants as 20, 21 and 26 subjects, as in the case of Cumming and So (1996), Greenslade and Felix-Brasdefer (2006) and Sheppard (1992) respectively, while Liu (2008) involved just 12 students and these were divided into two feedback groups. The majority of the studies had between 50 and 100 participants, while only three of the studies discussed had over 100, namely Fazio (2001), with 112, Robb et al. (1986), with 134 subjects and Semke (1984), with 141 subjects. In instances where the subjects are only about 20 for example and these were divided into different treatment groups, the findings may not be representative of the whole population. Kombo and Tromp (2006, 84) recommend a large sample as they say 'in general, the larger the sample, the more representative it is likely to be, and the more generalisable the results of the study are likely to be.' In circumstances where the sample is too small, it is quite difficult to make generalisations.

As implied above, the type of students also varied a lot in the studies reported in this literature review. In a majority of the studies reviewed in the literature review, the subjects were ESL writers that is English was used in the context within which they were writing. In a few (Lalande 1982; Robb et al. 1986; and Semke 1984), the writers were EFL writers, that is the learners were learning English as a foreign language but is not spoken in the context. In some the subjects were studying another language other than English as a foreign language such as Greenslade and Felix-Brasdefer (2006) where Spanish was the target language. In some of the studies, subjects were elementary students while in others they were college or university students. With different learners having different interests, for example EFL learners in academic contexts having an interest in passing examinations and so paying more attention

to grammatical features of the language, while immigrant workers may just want to learn the language for basics such as daily communication. Immigrant workers usually do not have any formal instruction in the learning of a language but some how they have learnt it subconsciously. Comparing these two for example on reduction of grammatical errors may not be very meaningful. Overall, this too means that it becomes difficult to generalise from the findings of EFL, FL, ESL and immigrant worker learners.

The period over which the studies were done also varies considerably. Some studies were relatively long term, ranging from 10 weeks to 15 weeks (Ferris 1997; Kepner 1991; Lalande 1982; Robb et al. 1986; Semke 1984), while others took place in the short term (from draft to draft). Those that mainly looked at the effect of feedback from one draft to the next include Fathman and Whalley 1990; Ferris and Roberts 2001; Ferris et al. 2000 as cited in Ferris 2002, 20; Greenslade and Felix-Brasdefer 2006. In the short term studies, improvement in accuracy was shown by the subjects, however this should be treated cautiously as one cannot simply conclude that the correction success will have a positive effect over time. Some of the long term studies reviewed, such as Kepner (1991), Sheppard (1992), and Semke (1984), found that feedback did not have a significant effect on the students' accuracy development. There is a need for studies to be carried out for even longer periods to enable more valid conclusions to be drawn.

The type of writing done in the studies reviewed is not homogenous. Writing that was used includes: journal writing (Fazio 2001; Kepner 1991; and Polio et al. 1998), autobiographical writing (Chandler 2003), composition writing (Fathman and Whalley 1990; Ferris 1997; Polio et al. 1998; Robb et al. 1986), errors implanted in a text (Lee 1997), diagnostic essays (Ferris and Roberts 2001) and free-form journal entry writing, where students could just write on anything they wanted (Semke 1984). With all these variations in the type of writing used, making comparisons in this area at the moment is, as Ferris (2004, 52) puts it, like comparing "apples and oranges (and pears and grapes and nectarines…)".

In connection with the type of writing used, is the issue of the length of writing and whether the writing was done in class or out of class. In some studies, the length of essays written was stipulated. In Ashwell (2000), participants wrote 500-word essays, Bitchener et al. (2005) asked the participants to write 250 words, in Chandler (2003), students wrote 25 pages of autobiographical writing, for

Frantzen (1995), all classes wrote between 200 to 250 words while Cumming and So (1996) had participants write between 300 and 400 words. On the other hand, some studies did not specify the amount of writing to be done. Fazio (2001) for example, asked subjects to write as much as they could and it is reported that the amount of writing varied between just a short paragraph and several pages in the weekly entries and this was done out of class. Greenslade and Felix-Brasdefer (2006) did not give a word limit but the students were limited to write within a limited amount of time similar to that of Fathman and Whalley (1990) and Polio et al. (1998). Comparisons become tricky to make between cases where there is no time limit, no word limit or both and cases where there are such limits.

Amongst the studies that have been reviewed in this study, on the issue of individual error categories a lot can be discussed. Some studies looked at errors comprehensively such as Fathman and Whalley (1990), but the results of that study do not report on individual error categories. In Ferris et al. (2000 cited in Ferris 2002, 50), it was noticed that even though the teachers involved in the research tried to mark errors comprehensively, they did not manage to notice all of the errors. This means that there were inconsistencies in the marking of errors in that study. Other studies, such as Liu (2008) just mention that morphological, semantic and sentence structure errors were studied but these are broad terms as they each consist of several error types (Ferris 2004, 53). The result then does not specify which particular error types were studied and their findings. Other studies looked at certain individual error categories but these were varied in terms of number and type. For example, Fazio (2001) only studied spelling errors, Bitchener et al. (2005) studied three error types (prepositions, past simple tense and the definite article) while others such as Chandler (2003) and Lee (1997) studied 23 and 21 error categories respectively. Some studies reviewed looked at similar error types such as those of Ferris and Roberts (2001), Hong (2004) and Schlig (2005) but again the results are varied. For example, both Ferris and Roberts (2001) and Schlig (2005) found that article errors were easy to correct but for Hong (2004) this error category was the most problematic for the students in the study. This means of course that factors such as differences in student background could affect results of studies even if the same error categories are studied, and so full clarity should be provided concerning these factors.

Regarding the type of feedback given there is a wide range in the studies reviewed. Some of the feedback mechanisms used included: feedback on form, content, form and content, error correction with additional grammar instruction, revision on feedback, and teacher conferences after feedback had

been given, marginal comments and no feedback. Some studies that used different forms of feedback assigned students to the same treatment group for the whole study. In others, such as Chandler (2003) and Ashwell (2001), feedback mechanisms were used in rotational form. It would be easier to draw general conclusions if comparisons are made across studies that make use of the same feedback mechanisms assigned to students throughout the study.

As regards data collection, some studies were not consistent. For example, Schlig (2005) uses two groups in her study. For one she collected data using a 16-item system adapted from Lalande (1982) while for another a 5-item system was used, adapted from Ferris and Roberts (2001). The group receiving the 16-item coding could have found it more challenging compared to the one receiving the 5-item code and so it is difficult to compare the two groups' results. In Schlig's study (2005), one group of students received electronic feedback and was allowed to use the editing tools of Microsoft Word. Under such circumstances, it is difficult to tell to what extent these students noticed the errors and corrected them themselves or just used assistance from the computer programme.

It is indeed difficult to come up with definite general conclusions on the value of feedback because of the differences that exist in research designs. None of the studies is a replica of the other. For example if the error categories studied were similar in two studies, there could have been differences in the feedback mechanisms used or the type of subjects or the type of writing or even the period of study. Ferris (2004, 52) points out that amongst the existing studies, 'none of the studies constitute 'replication' by any stretch of imagination' and that there is a need for at least more partial replication. My study is in some ways such an attempt as it uses feedback mechanisms and studies similar errors to those that have been used in some existing studies.

2.3 Conclusion

Having reviewed several studies, certain conclusions can be drawn. Firstly, on the effect of revision on student writing, all the studies reviewed had similar findings in that students wrote better texts when they revised their work compared to the original texts. Most of these studies were in the short term in that this was from draft to draft. It is difficult to tell if these results would remain the same in the long term. My own study looks at both the effects of some correction feedback mechanisms in the short

term (from draft to draft) and in the longer term, namely over a period of nine weeks. More valid comparisons can be made on the effect of error correction in the long term.

Little research has been done in the area of fluency. In the literature reviewed above, only four studies (Chandler 2003; Frantzen 1995; Robb et al. 1986; Semke 1984) actually measured fluency. In Frantzen (1995) and Robb et al. (1986) none of the groups used showed significant improvement in fluency. Interestingly, however, in Chandler (2003), both the experimental and the control groups showed significant improvement in fluency and Semke reported similar findings. More research is needed in this area since fluency is fundamental in the learning of a second language. My study sought also to investigate whether different feedback mechanisms affected students' fluency in their writing.

For accuracy, a substantial amount of research has been done in comparison with fluency. The major problem in this area, as mentioned above, is that the findings appear to be contradictory as some results reveal that accuracy improved while in others there was no improvement. More detailed research is needed in this area to determine the effect of feedback but as Ferris (2004) cautions, this should be done with proper methodology.

Several studies sought to investigate the effect of feedback on different error categories. Some specified the error categories studied, (such as Ferris and Roberts 2001, Frantzen 1995, Lee 1997 and Schlig 2005) while others did not. More research is needed that deals with all error categories that exist so as to allow better comparisons to take place.

The sometimes similar and sometimes contradictory findings of the studies reviewed in this chapter could be attributed to different designs and different methodologies that have been used. To alleviate the controversy as to whether to give feedback or not, more research is needed - especially, as Ferris (2004) puts it, more controlled longitudinal studies that are comparable in design and methodology. Given this background, the present study attempts to explore and shed more light on this controversial area even though it is by no means free of the shortcomings that are typical of research in this area.

CHAPTER 3

RESEARCH METHODOLOGY

3.0 Introduction

This chapter describes the research methodology that was used in the study. It is divided into five sections. Section 3.1 covers the design of the study, followed by Section 3.2, which looks at the subjects involved in the study and Section 3.3, which describes the materials and procedures. Section 3.4 then deals with data collection and procedures and finally Section 3.5 deals with processing of results.

3.1 Design of the study

The study made use of a quasi-experimental design. 'Experiments are usually carried out in order to explore cause-effect relationships between variables' (Nunan 1992, 23). In experiments, relationships between variables can be tested. A quasi-experiment differs from a true experiment in that in a true experiment, subjects are randomly assigned to control and experimental groups, while in a quasi-experiment this is not done. In educational studies it is often difficult to carry out true experiments, especially in situations where groups such as classes are already intact, and then quasi-experiments become appropriate. Although the internal validity of quasi-experiments is not as high as that of experiments, the experimental and control groups are at least studied under naturalistic conditions. As in true experiments, both pretest and posttests are used in the study.

This study made use of a pretest-posttest design as outlined in Nunan (1992, 34). A pretest is a measure that '… is taken for all cases prior to the introduction of the independent variable in the experimental group while a posttest is a second measure that is taken for all cases after the experimental group has been exposed to the independent variable' (Frankfort-Nachmias and Nachmias 2004, 101). Both the experimental and the control classes that participated in this study wrote essays from each of the genres and the first essay in each was used as the pretest while the last written in each genre written constituted the posttest. The narrative pretest topic was 'The most interesting holiday that I have ever had', while the posttest topic was 'Write a short story on the following topic: A midnight

visitor.' The descriptive pretest topic was 'Describe your room after a party is over and the guests have left', while the posttest topic was 'Describe a natural disaster you have witnessed.'

The study also made use of a students' questionnaire to elicit information from students on their feedback preferences and a structured interview to elicit information from the teacher who participated in the study on teacher practices in the classroom regarding teacher correction feedback in writing.

3.2 Subjects

At the beginning of the year 2006, I identified the school to be used in the study. At the time when this study was carried out, I was not teaching in a secondary school. I therefore had to seek authority from a school where the study could be carried out. Then a letter was provided from the University of South Africa requesting that the researcher be given access to the school. I gave letters of introduction to the school authorities, who responded positively and guaranteed permission for the study to be carried out.

As indicated in Chapter 1, the school that was selected for the study was a private co-educational school located within Gaborone City in Botswana. The school was chosen based on the feasibility and the willingness of the school administration to allow the study to be carried out. It had students with a lower academic profile than their counterparts who had been admitted to senior government schools. The government of Botswana has a policy of admitting only students who score a mean grade of A or B overall in the JC examination to join Form 4 senior government schools. Those who fail to achieve these mean grades have to join elsewhere, such as the privately run senior secondary schools. The weak mean grades achieved at the JC exams by these students are an indication that these students are of a relatively low academic level, with most of the students in the school having attained an average grade C or D in their junior certificates (Form 3) and as a result the government schools could not absorb them. The students were expected to sit for their final secondary school examination the following year (2007) at the end of Form 5. This is an examination that would determine whether the students could join institutions of higher learning or not. The students had learnt English as a subject right from their first year of learning (Standard 1) while the other subjects were taught in Setswana up to Standard 4 with a switch to English as from Standard 5. English is the main medium of instruction at both junior and senior secondary school and so the students had had six years of English as the Language of Learning and Teaching (LoLT).

The study involved 68 students and one teacher of English. The students were drawn from two Form 4 classes from one private co-educational secondary school. The Form 4 class was chosen because it was a pre-candidate class and the students were expected to sit for their BGCSE examination the following year (2007) at the end of Form 5. There were 28 males and 40 females and their ages ranged between 16 years and 21 years. All the students spoke Setswana as their first language and were learning English as their second language.

At the beginning of the study, a total of 68 subjects were involved. However not all these subjects completed writing all eight essays. To allow consistency, only those who completed writing all the eight sessions had their results processed for this study. By the end of the study, only 50% of the control class students had succeeded in writing all the eight essays, 82% from the coded group, 91% from the uncoded group and 82% from the direct group. Data was collected from a total of 57 subjects who successfully participated in the study, and a total of 214 essays were analysed. The analysed essays were only from the pretest and the posttest. The experimental group resulted in 30 subjects for both the narrative genre and the descriptive essay for the pretests and the same number of scripts for the posttests respectively. The control group yielded 27 subjects for the narrative essay and 20 subjects for the descriptive essay in the pretests and a similar number of scripts in the posttests respectively. The number of subjects varies between the narrative and the descriptive essays because for each category scripts were analysed only for those who completed writing all the eight essays.

The school chosen for the study was a three-streamed school. Students were assigned to the classes according to when they reported to the school without any other consideration. Two of the streams had an equal number of students while one had fewer students. The two with an equal number of students were selected for the study. Once the two classes used for the study had been selected, one was labelled heads and the other tails. A coin was then tossed to select which class would be used as the control class and which would be experimental. The part of the coin that came up determined the class that became the experimental class while the other became the control group. After the pretests had been written, the experimental class had its subjects assigned randomly to the three different feedback groups. I did this by putting three labels of the feedback mechanisms used in the study on my desk. Then I distributed the scripts of the experimental class into these three groups randomly. Since the

scripts had not been collected in any particular order, then each script had the chance of landing in any of the three groups. Random sampling allowed each student in the class to have an equal chance of being selected to any of the groups (Richards et al. 1992; Kombo and Tromp 2006). The subjects maintained their various groups throughout the term. This was to enable the researcher to investigate the long-term effects of different teacher feedback mechanisms on student writing.

A teacher of English who taught both classes was involved in this study. He was a graduate teacher with nine years of teaching experience in high school. As the study required intensive analysis and students writing and re-writing compositions, it was not really feasible to carry it out in more than two classes. Making use of only one teacher was quite limiting but more positively, this did mean that it was easy for me to control for the teacher variable across the two classes. The teacher co-operated with me in administering the writing sessions, collecting the scripts from the students at the end of each session, giving the scripts to me to provide the different types of feedback and then returning them to students.

3.3 Materials and procedures

A quasi-experiment, a student questionnaire and a structured teacher interview were used in the data collection. The quasi-experimental design was made use of because it would allow the researcher to '… draw causal inferences and observe with relatively little difficulty whether or not the independent variable caused changes on the dependent variable' (Frankfort-Nachmias and Nachmias 2004, 103).

The quasi-experiment made use of compositions that were drawn from one text prescribed in the Botswana schools syllabus (Ngoh and Fereira 1980, 240, 248). The compositions that were chosen for the study were among those given in the textbook. Of the eight compositions, four were from the narrative genre while four were from the descriptive genre (see Appendix 1). The two genres were chosen for the study because it was felt that they had already been introduced to the subjects both at the primary school level and the junior school level and thus learners would not have much problem in their writing. The narrative genre is usually considered one that is easily understood by students and also has different patterns of tense use to the descriptive genre, which was also used to allow comparisons.

For the quasi-experiment, prior to the main study, a pilot study was conducted by the researcher in the co-educational private school, as already mentioned in Gaborone city within Botswana. This was carried out in the first term of 2006. The purpose of the pilot study was to enable the researcher to gain access to the school used for the study, interact with both the teachers and the students of the school and to test the instruments that would be used in the study and make any possible adjustments that would be necessary. The main study took place in the second term of 2006. The second term of the year was chosen because it was the longest (14 weeks) in the year and so the best for investigating long-term effects of the correction feedback types as independent variables.

At the beginning of the second term of the year, I met with the co-operating teacher. This was the teacher who was currently teaching both classes that were used in the study. The teacher and I decided upon the topics that would be used in the study as well as the schedule for the study, which included: the dates when the writing and rewriting sessions would take place, when the researcher would collect students' scripts for feedback, when the scripts would be returned to the teacher to give back to the students and when the rewritten scripts would be collected for the researcher to mark and have them returned to the students.

The co-operating teacher then informed the subjects that they were going to have several writing sessions in the course of the term, more than what they normally had, and this was intended to help them improve in their writing. They were also informed that the average mark of the essays would contribute towards the continuous assessment mark at the end of the term. The subjects were not informed that they were under observation until the end of the study. This was to prevent effects such as the Hawthorne effect, as if the students knew that they were under observation, they might change their way of writing, the results were likely to become contaminated and the situation would no longer remain naturalistic. It worked quite well as the students did not discover that they were under observation. At the end of the study, I requested the students for consent to use their results in completing a dissertation project, which the students gave.

The study made use of an experimental class and a control class. The experimental class had subjects assigned to three different groups which were: direct feedback group, coded feedback group and uncoded feedback group. The direct feedback group (n=11) had their essays marked by having their

errors given the correct forms. The correct forms of the errors were written above the error that had occurred. The coded feedback group (n=11) had their errors coded, using the codes shown in Table 3.1 below. The errors were underlined and a code was written above the error. The codes were explained to the students at the end of their essays so that they could understand them. The codes only gave students an idea of the kind of errors that they had made and nothing else. Some of the codes were adapted from Ferris (2002, 155) and are shown in table 3.1 below. The uncoded feedback group (n=12) received feedback whereby the errors were underlined without an indication of what type of errors they were. The students were to try and work out the errors for themselves and correct them. The control group had 34 students and this group did not receive any form of feedback except for a summative statement. This was for ethical reasons such that the students had a feel that their essays were looked into. Students maintained their groups throughout the term to allow consistent comparisons to take place.

Table 3.1: Codes used in the coded group

Code	Explanation
Art	Article use errors
V	Verb errors
Prep	Preposition errors
Sp	Spelling errors
P	Punctuation errors
Ww	Word choice errors (‘wrong word’)

Both the experimental class (n=34) and the control class (n=34) were taught by the same teacher. They received the same number of teaching hours weekly and did the same coursework throughout the term. The eight essays (Appendix 1) were written in the course of the study, of which four were narrative while the other four were descriptive, with both groups responding to the same topics. The students were asked to write each essay (first draft) in about 300 words within a time limit of 80 minutes. All the essays were done in class.

The experimental class had to rewrite their essays after receiving their previous drafts, complete with feedback. The students in the experimental class were assigned to the three groups mentioned and they

maintained their feedback groups throughout the term. The control group did not receive any kind of feedback except for a summative statement so that they could feel that their essays were indeed marked. The summative comments were about matters such as the length of the essays, organisation of the work, content and any other area that the researcher felt that the students had to work on. The control group was not asked to rewrite their work either as it was felt that if they were asked to, they would not have seen this as meaningful since only general statements were given for feedback. In this study, only the researcher gave the feedback in all the essays, so as to allow consistency in marking as well as more valid comparisons between the pretest and posttest results.

All the essay writing sessions were in-class and were supervised by the co-operating teacher in the study. The first scripts were written during normal class teaching hours and they were collected by the teacher and passed on to the researcher to give feedback. I gave feedback and the scripts were returned to the students via their teacher. The students were asked to rewrite their essays in 40 minutes. The rewriting sessions were done weekly in the afternoon sessions when students had free time to do their own studies. Though the term was 14 weeks long, only nine of the weeks were made use of. This is because the first two weeks were used in making arrangements on how the study was to take place; one week was a mid-term break while the last two weeks of the term were examination weeks. The questionnaire was administered on the 12th week after the last feedback had been given. The rewrites were done immediately after feedback was given and the scripts collected for marking, as only the rewrites received a mark. Any errors that the learners were not able to correct were corrected for them and their scripts given back to them via their teacher. Correcting the errors that the students were not able to correct after the first feedback was done essentially for ethical reasons and to meet students' expectations, but it must be noted that doing so might have had the effect of lessening differences between the treatments given. This could therefore mean that the study's results would be an underestimation of effects, though on the other hand the students were not required to do anything with those final corrections and so they might not have had much impact on their learning. All the writing of the first scripts and the rewritten scripts was done in class within a given time so as to obtain maximum control over the students' writing and also help in making accurate comparisons of the effects of various feedback treatments that were given.

A students' opinion questionnaire (Appendix 3) was administered to the students at the end of the study. The questionnaire consisted of five closed questions. The students were to indicate whether it was easy to see the kind of mistakes that they made, whether teacher correction helped the students to correct their work, whether they understood the correction given by the teacher, whether they felt discouraged upon getting feedback and how often they would like the teacher to correct their errors. The questionnaire used a 5-point Likert scale. The main purpose of the questionnaire was to help draw comparisons between the students' views on feedback and the findings of the study.

After completing the study, I wrote a letter of appreciation to the school administration thanking them for allowing the study to take place. I also thanked the subjects verbally for their co-operation.

3.4 Data collection procedures

Brown (1994, 171) defines error analysis as 'The study of errors made by learners regardless of the source.' In any error analysis study, it is important for a researcher to identify the errors to be studied before carrying out research. This study focused on a selection of the most common types of errors made by students. Comprehensive correction of errors would have been not only tedious for the researcher but also overwhelming for the students (Ferris 2002), while focusing on just a few of the errors can help students for example to improve their understanding of the grammatical rules associated with some errors.

In order to identify the errors to be analysed in this study, a pilot study was carried out which provided information on the errors that students made in their writing. From the results of the pilot study, it was found that the six common errors that were most problematic in the students' writing were: article, verb, punctuation, preposition, spelling and wrong word choice errors. The study then chose to cover only these six error categories. Similarly, Bitchener et al. (2005) studied the most problematic errors that were found in the first writing task that was given to his subjects. These errors fall into two broad groups, namely treatable and untreatable errors. According to Ferris (2002, 23), treatable errors are rule governed. This means that when treatable errors occur, a rule can be pointed to, which can be used to correct the error. Untreatable errors on the other hand are not rule governed. Such errors require the student to use prior knowledge that has been acquired to solve them. The wrong word choice fell into the untreatable group while the rest were classified as treatable errors.

After identifying the error types to be studied, the next step was to describe them. Details of these error descriptions are found below in Table 3.2.

Table 3.2: Error categories used for feedback analysis

Verb errors	Errors in verb tense, subject-verb agreement and omission of verbs
Articles errors	Wrong choice of article, incorrect addition or omission of articles
Spellings	All errors in spelling
Prepositions	Wrong choice of preposition, incorrect addition or omission of prepositions
Wrong word	Wrong choice of lexical item (i.e. noun, verb, adjective or adverb), including wrong category choice (e.g. *dramatically* for *dramatic*).
Punctuation	Omissions or wrong choice of punctuation marks and capitalisation.

The pretest consisted of the first narrative and descriptive essays (Appendix 1) that were written by the students at the beginning of the term, which were copied to be used later in data analysis. These essays were also used for the first treatment. The posttest consisted of the last two essays of which one was narrative and the other descriptive (Appendix 1). These were also copied to be used in the data analysis. It should be noted that the posttest essays were not rewritten.

For the fluency hypothesis (H1), both the experimental and the control group were asked to write about 300 words for each essay. The number of words written by each of the students for each of the eight essays were counted and recorded. For analysis only the pretest and posttest number of words were compared.

To collect data for the correction success hypothesis (H2), students were asked to write a first draft of the given essays and upon receiving feedback they were to rewrite their essays. The number of errors made was tallied according to the error categories. These were recorded and for the rewrites the number of errors not corrected was also recorded. The percentage of errors corrected was calculated by dividing the number of errors corrected by the number of errors in the original essays and multiplying the result by 100. I marked the scripts according to the different feedback mechanisms that had been

assigned to the eight essays. The errors were tallied according to the different error categories in the study for both the original and the rewritten essays.

For the accuracy development hypothesis (H3), the total number of errors in the individual error categories studied as well as the totals for each individual category was calculated for all the essays marked. The pretest results and the posttest results were compared and statistically tested for significance using a t-test. Improvement was confirmed if errors made in the posttest were significantly fewer than in the pretest.

I requested a colleague who is also a teacher of English to mark 10% of the scripts that were written by the students from each of the experimental groups. The teacher was to identify and categorise the errors in terms of the six error categories that were examined in the study. There was some disagreement between the rater and myself on what categories some errors fell into, and a limitation here is that I did not calculate interrater reliability, partly because of the difficulties around the interpretation of results (Polio 1997). However, our disagreements were discussed and consensus was arrived at. There will always be some degree of subjectivity in error analysis, but allowing another rater to examine some of the essays assisted me to identify and categorise the errors more consistently.

Error analysis in this project was a difficult task in various ways. Firstly, it was sometimes difficult to distinguish whether an error arose as a result of spelling or wrong choice of words as in the example … *I found shoes on <u>their</u> way out* (Appendix 5; UN6 pretest descriptive essay). Presumably the student intended to say …*I found shoes on the way out.* In this case it could be analysed as a wrong word error but it could also be a spelling error. In my analysis I took it as a spelling error since in the same essay the same student had problems differentiating the use of *their* with other words like *there.* Another example is … *I excepted their kindnesss* (Appendix 5, C9 pretest descriptive essay). This could be treated as a wrong word choice, with the student confusing *excepted* for *accepted* or equally it could be analysed as a spelling error. In such cases, I made use of the context and also of what I know about the students' background to help me classify the errors. In this case, it was treated as a spelling error as it could have arisen as a result of wrong pronunciation. Amongst the Setswana speakers, there is a tendency to raise the initial vowel in *accept*, leading to an *e* spelling, and hence confusion with another word's spelling.

Secondly, some errors fell into two error categories such as verb/wrong word as shown by the example … *I even picture about my bedroom* (Appendix 5; C9, pretest descriptive essay). Here the student could have wanted to say *I even pictured my bedroom*, hence a verb error, or they could have wanted to say *I even thought about my bedroom*, hence a wrong word error. From the context, I categorized this as a wrong word error because in the sentence that follows, the student says *I thought that they would respect it* ….

Thirdly, sometimes it was difficult to tell if an error was a verb or a spelling error, for example …*I found them shivering and they came to me and apologies* (Appendix 5; C9, pretest descriptive essay). The word *apologies* could be treated as a spelling error as well as a verb error. In such instances, I again looked at the context in which the word was used and made a decision on where the error belonged. This was classified as a verb error because from the context, an apology was being made so the student may have intended to say … *they came to me and apologised* .

Fourthly, where one word was spelled wrongly on several occasions in the same essay, it was difficult to decide whether the word should be counted as one error or as an error each time it occurred. In this study, each one was tallied such that, for example, where *oclock* (Appendix 5, UN6 posttest narrative and pretest descriptive essays) appeared several times in one of the student's essay this was counted as six errors and each of them was underlined in the students' script whenever it occurred.

Fifthly, it was sometimes difficult to comprehend what the student intended to say, for example, in one of the student's essays, the following sentence was used: *I started to become shuddered wondering what is happening* (Appendix 5, C9 posttest narrative essay). I did not have any idea what the student meant by the word *shuddered*, therefore making it difficult for me to give the right correction. This was difficult to categorise and again I tried to guess what the student meant from the context. This was taken as a wrong word error. Similar situations occurred in Frantzen (1995) and Lee (1997) where some errors could not be categorised and Lee labels them as unclassified.

In some cases it was difficult to decide whether to classify an error as a preposition error or a wrong word error. For example in the following statement, *When I got at the party it was good* (Appendix 5,

UN6 posttest narrative essay) the student could have meant *arrived at,* thus a wrong word error, or it could have been *got to,* making it a preposition error. Classifying this kind of error was quite problematic and I classified it as the former but the student could have meant the latter.

Sometimes some words were left out and these were not marked as punctuation errors since a missing word could have been more appropriate. The study does not analyse such missing words. For instance, (Appendix 5, UN6, pretest descriptive essay) … *On Saturday the party started at two in the afternoon it was good and I was very happy.* This could read better if the word *and* is inserted … *afternoon and it was good*…. Another example, (Appendix 5, UN6, posttest descriptive essay) is *I was the only one who was not burnt I was admitted for being afraid.* The sentence would read better if it was *I was the only one who was not burnt and I was admitted for being afraid.* The error is the omission of a word, here and, rather than a punctuation error, but word omissions are not one of the categories analysed in the study, so neither punctuation nor omission errors are analysed in such cases.

Lastly, in this study the verb error category is very general, covering a lot of different verbs. One student may have had problems with verb forms, another with subject-verb agreement while another with the tense. Comparisons of improvement may therefore not give a fully clear picture as improvement may have varied depending on the type of verb error. It was decided, however, that analysing the six error categories at further levels of detail would be beyond the scope of the present study, and that in this respect it would remain more directly comparable to other similar studies.

3.5 Processing the results

The data collected for the study related in the first place to the three main hypotheses that were derived from the first three aims of the study. There was no hypothesis derived from the fourth aim as this was dealt with in an exploratory manner.

For Hypothesis 1, the fluency hypothesis, the students were expected to write each essay in about 300 words. To check if the students' writing increased in the number of words written, the number of words written in each first draft essay was counted and recorded. For statistical analysis, only the pretest and posttest were compared in terms of the mean number of words written. An increase in the mean number of words in the posttest was seen as an improvement in fluency while a decline in the

mean number of words was seen as lack of improvement in fluency. To test for significance, t-tests were run, using the SPSS program, version 10. The t-test is a statistical procedure for testing the difference between two means. The t-test allowed easy comparisons to be made in terms of the effect of each feedback mechanism as well as no feedback on fluency.

For Hypothesis 2, the correction success hypothesis, the number of errors made in the first draft was tallied according to the respective error categories studied. The students then revised and rewrote their essays after teacher corrective feedback had been given. I then checked out the errors that had been corrected successfully and those that had not. This was applicable only to the experimental groups as the control group did not revise and rewrite their work. The percentage number of errors corrected in the student rewrites was calculated and to find out if there were statistically significant differences in correction success between the groups, Chi-square tests were run. This test is designed to evaluate whether the differences between the observed frequencies and the expected frequencies under a set of theoretical assumptions is statistically significant' (Frankfort-Nachmias and Nachmias 2004, 496). To find out whether the three feedback mechanisms could be ranked in terms of correction success, subsidiary Chi-square tests were undertaken for each pair of feedback types relative to one another (using the VassarStats website: http://faculty.vassar.edu/lowry/vassarStats.html).

For Hypothesis 3, the accuracy development hypothesis, errors were tallied and totaled according to the error categories and their grand totals calculated as well. The mean density of errors per 100 words was then calculated for all the essays but only the pretest and posttest error densities were compared. Improvement in accuracy was calculated in terms of the mean density of errors per 100 words in the pretest and posttest. The hypothesis was first tested for all the errors combined for each of the experimental groups and the control group. The t-test was used to find out if there was any significant improvement in accuracy for each of the feedback types. The second part of the accuracy development hypothesis investigates whether there is a relationship between feedback mechanism used and the development of accuracy with regard to specific error categories in student writing. A t-test was carried out to find out if there was any significant improvement in each of the error categories.

The last aim, which was dealt with in a more exploratory manner investigated the students' attitude towards teacher correction feedback by using a questionnaire, though a statistical test was used here

too. Students' responses were analysed by the use of a five-point Likert scale. The Likert score means for each group and each question were calculated and these were used for analysis. Overall totals were calculated for each of the feedback groups and in order to see whether there were significant differences in the different groups' attitude to error correction the Chi-square test was used (using the VassarStats website: http://faculty.vassar.edu/lowry/vassarStats.html).

My study is, however, essentially a quantitative one that uses a pretest-posttest design and tests its hypotheses statistically. The use of statistical testing made possible more generalisable comparisons and conclusions with regard to the effect of different error feedback mechanisms on fluency, correction success and accuracy development.

CHAPTER 4

FINDINGS

4.0 Introduction

This chapter presents and discusses the results of the study. The first three main sections are concerned with the hypotheses of the study (see Section 3.6). Section 4.1 deals with findings on fluency (Hypothesis 1). Section 4.2 looks at students' success in correcting errors during revision (Hypothesis 2) and Section 4.3 looks at the findings on students' success in accuracy development over time (Hypothesis 3). This latter section is divided into two parts. The first deals with the overall accuracy results, where all the error categories are analysed together to find out whether the different groups of students made any improvement in accuracy in general, while in the second, individual error categories are analysed to find out whether the groups of students made any improvement regarding these categories. This is followed by Section 4.4, which provides a complementary perspective on the overall hypothesis findings by presenting a more qualitative account of samples of individual students' errors, specifically of those students who improved most in spelling error categories between the pretest and the posttest. Section 4.5 then analyses findings from the questionnaire administered to elicit students' views on teacher error correction feedback while Section 4.6 discusses the structured interview held with the teacher. Section 4.7 concludes the chapter.

4.1 Effects of corrective feedback type on students' fluency (Hypothesis 1)

This section deals with findings of the first Hypothesis (H1) namely:

> There is a relationship between the feedback mechanism used and the development of fluency in students' writing.

Fluency has been described as 'the rapid production of language' by Wolfe-Quintero, Inagaki, and Kim (1998) and has been measured in different ways. Chandler (2003) measured fluency by how much students wrote (see Section 2.1.2). In the present study fluency is examined in terms of the change in the number of words written from pretest to posttest. The number of words was counted both in the

pretest and posttest essay and any changes were an indication of either improvement in fluency if the number of words increased or none if the number of words declined (see Section 3.3).

The analysis for fluency for all the experimental groups and the control group was carried out. The mean number of words in each feedback group is shown in Tables 4.1 (narrative essays) and Table 4.2 (descriptive essays), followed by the mean number of words written by the students. The standard deviations were also calculated. To find out if students wrote longer or shorter texts in the posttest, the pretest means were subtracted from the posttest means. To find out if there was any significant improvement in the number of words written in each feedback group a t-test was carried out whereby the pretest and posttest results were compared. The data was computed using the SPSS programme for Windows 10. The following headings have been used and their meanings include: SD stands for standard deviation, T stands for the t-value, df stands for degrees of freedom while Sig (2-tailed) stands for the p-value or level of significance (given a 2-tailed test). The results were tested in terms of the 0.05 level of significance. A negative sign in the analysis for fluency in the pretest-posttest change column shows that the students wrote fewer words in the posttest. The findings are arranged according to each feedback group. Fluency results for the narrative essay are shown on Table 4.1 below.

Table 4.1: Narrative essay (fluency)

	Pretest			Posttest			pretest-posttest change	t-test outcome		
Group	Totals	Mean	SD	Totals	Mean	SD		T	df	Sig (2tailed)
Direct (n=10)	3340	324.00	68.86	3073	307.30	66.69	-16.7	.781	9	.455
Coded (n=10)	3622	362.20	73.59	3563	356.30	91.22	-5.90	.321	9	.755
Uncoded (n=10)	3692	369.20	118.87	3745	365.50	99.44	-3.70	.122	9	.905
Control (n=27)	8559	317.00	78.59	8849	327.74	114.55	10.74	.555	26	.584

Analysis for students' fluency was done and in the narrative essay it was revealed that in the pretest essay, the highest mean number of words was written by the uncoded feedback group (369.20) while the control group wrote the lowest mean number of words (317). In the posttest the uncoded feedback

group still wrote the highest mean number of words (365.50), while the direct feedback group wrote the lowest (307.30). All the experimental groups wrote a lower mean number of words in the posttest with the direct feedback group having the largest decline in mean number of words (-16.7). However, the t-test results revealed that none of these differences were significant. The control group showed that the mean number of words written in the posttest increased by 10.74 words, but this result was also not significant, and so, neither those who received feedback nor those who did not showed any real change in fluency.

The number of words for the pretest and the posttest descriptive essays for the experimental groups and the control group were analysed. The results are shown in Table 4.2.

Table 4.2: Descriptive essay (fluency)

	Pretest			Posttest			pretest-posttest change	t-test outcome		
Group	Totals	Mean	SD	Totals	Mean	SD		T	df	Sig (2tailed)
Direct (n=9)	2843	315.89	69.76	2579	286.56	62.54	-29.33	1.8	8	.103
Coded (n=11))	3506	318.73	58.50	3389	308.10	72.34	-10.63	.667	10	.520
Uncoded (n=10)	3412	341.20	73.99	3118	311.80	89.86	-29.40	1.8	9	.110
Control (n=20)	5823	291.15	74.65	5141	256.85	86.56	-34.30	1.747	19	.097

For the descriptive essay in the pretest the uncoded feedback group wrote the highest mean number of words (341.20) while the control group had the lowest (291.15). In the posttest essay the uncoded feedback group had the highest mean number of words (311.80) while the control feedback group still had the lowest (256.85). For all the experimental groups and the control group, the mean number of words written by the students decreased in the posttest. The highest decline was in the control group, which had a mean decline of -34.30 words, and although this decline was not significant, it could be argued to show a weak tendency towards significance (being less than p=.010 at p= .097). When the results for the experimental groups were compared they were not found to be significantly different either and so none of the groups showed any improvement in fluency.

The results of the study do not therefore support the hypothesis that teacher error correction of any sort leads to improvement in students' fluency.

4.2 Effects of corrective feedback on students' success in correcting their essays (hypothesis 2)

This section focuses on Hypothesis 2 namely:

> There is a relationship between the feedback mechanism applied to student errors and students' success in correcting errors during revision.

In this study a total of eight essays were written of which four were descriptive and four were narrative. The first essay in each of the genres was the pretest. Feedback was given to students in the experimental groups on this pretest essay and so it also takes the role of the first essay in the study. The percentage of errors corrected was calculated by subtracting the number of errors not corrected in the rewrites from the number of errors earlier made in the original scripts. The results were then divided by the number of errors made in the original written scripts and then multiplied by 100 to arrive at a percentage. The results in Table 3, Table 4 and Table 5 below refer to the three essays which students had to correct after feedback. It was important to find out how the students performed in each of the essays that they corrected and whether they got better at correcting over time. The fourth essay has not been included as this was the posttest and the students never received any feedback on it. Table 4.3 shows the percentage of errors successfully corrected by students in their rewrites for the first narrative and descriptive essays.

Table 4.3: Percentage of errors corrected (Essay 1)

Type of Feedback	V	AR	Prep	P	SP	WW	TOTAL
Direct (N=10)							
Narrative	96	94	96	97	98	97	97
Descriptive	92	86	68	100	84	60	84
Coded (N=11)							
Narrative	81	67	92	75	85	80	81
Descriptive	79	100	88	48	100	84	85
Uncoded (N=11)							
Narrative	80	67	78	75	81	91	80
Descriptive	25	83	73	84	86	41	58

In both the narrative and descriptive essays, students corrected most of their errors in the rewrites. On the individual error categories, the spelling errors were corrected best in the direct feedback (narrative essay) and for both coded feedback and uncoded feedback in the descriptive essay. In both the narrative and the descriptive essays, the least explicit feedback (uncoded) led to the lowest corrections while direct feedback led to the most corrections in the narrative essay and coded feedback to the most in the descriptive essay.

Table 4.4 shows the percentage of errors corrected by students in their rewrites in the second narrative and descriptive essays.

Table 4.4: Percentage of errors corrected (Essay 2)

Type of Feedback	V	AR	Prep	P	SP	WW	TOTAL
Direct (N=10)							
Narrative	96	100	100	50	89	100	95
Descriptive	89	100	89	94	83	90	89
Coded (N=11)							
Narrative	87	100	93	0	92	73	89
Descriptive	73	45	100	55	68	84	72
Uncoded (N=11)							
Narrative	88	0	40	90	45	100	83
Descriptive	85	80	75	92	46	62	79

In the second essay for both narrative and descriptive genres, direct feedback showed the highest number of errors corrected in both the narrative essay (95%) and in the descriptive one (89%). The percentage of errors corrected in all groups was high, ranging between 72% (coded feedback, descriptive) and 95% (direct feedback, narrative). Overall, the percentage of errors corrected in the narrative essay was higher (ranging from 83% to 95%) compared to the descriptive essay (from 72% to 89%). With regard to the error types, in some categories, students noticed and corrected all the errors, for both the narrative and descriptive essay as indicated by 100% in the above table. There are two instances where none of the errors made were corrected as indicated by 0% in the table. This was the case for the uncoded feedback group where in the narrative essay the students were not able to correct the article errors made. The coded feedback group was also not able to correct punctuation errors made

in the narrative essay. There does not seem to be a consistent trend in the correction of different errors in the set of essays.

Table 4.5 shows the percentage of errors corrected by students in their rewrites in the third narrative and descriptive essays.

Table 4.5: Percentage of errors corrected (Essay 3)

Type of Feedback	V	AR	Prep	P	SP	WW	TOTAL
Direct (N=10)							
Narrative	78	75	83	79	75	82	78
Descriptive	84	67	84	30	100	76	79
Coded (N=11)							
Narrative	61	67	44	81	63	77	63
Descriptive	75	82	42	53	54	65	65
Uncoded (N=11)							
Narrative	62	89	67	53	42	56	65
Descriptive	60	50	69	48	50	46	53

The percentage of the total errors corrected seemed quite low in this set of essays compared to the first and second sets. For the narrative essay, this ranged between 63% (coded feedback) and 78% (direct feedback), while for the descriptive essay, the total percentage of errors corrected ranged between 53% (uncoded feedback) and 79% (direct feedback). When all groups were compared, for descriptive and narrative essays, the highest percentage of errors corrected was by the direct feedback group. For the individual error categories, in the narrative essay, the highest error correction was in the punctuation category while in the descriptive essay, it was in the article category.

For the direct feedback group, the highest error correction in the narrative essay was in the preposition category while for the descriptive essay, the highest correction was in the spelling category. For the coded group, the highest correction in the narrative essay was in punctuation category while in the descriptive essay the highest correction was in the article category. For the uncoded group, the highest error correction in the narrative essay was in the article category while in the descriptive essay, the highest correction was in the preposition category.

When the first, second and third essays were compared in terms of the percentage of total error corrections the direct feedback group had the highest percentage of errors corrected in both the narrative and descriptive essays in all instances except for the first descriptive essay. There was however no consistent indication of the ease with which different error categories was corrected. Overall, the result is a little strange in that the second essay was corrected best while the third essay was worst corrected for all three groups. This could be attributed perhaps to the students getting tired of writing and rewriting, which they were not used to before. This result could however also just be due to random variation. This indicates that correction did not get more successful with time.

In order to arrive at a broader perspective, the results of all three sets of essays were amalgamated. Table 4.6 shows overall percentages for errors corrected for the experimental groups.

Table 4.6: Overall percentage of corrected errors

Type of Feedback	V	AR	Prep	P	SP	WW	TOTAL
Direct (N=10)							
Narrative	90	90	93	76	88	93	90
Descriptive	88	84	80	75	89	76	84
Coded (N=11)							
Narrative	76	78	76	53	80	77	78
Descriptive	75	76	77	52	74	78	74
Uncoded (N=11)							
Narrative	77	52	62	73	56	83	76
Descriptive	57	71	72	75	61	50	63

The average percentage totals were calculated for each of the feedback types for the three essays as a whole in each genre. For the narrative essay the direct feedback group had the highest percentage of total errors corrected, followed at a distance by the coded feedback group with the uncoded feedback group slightly lower still. All three feedback groups' results revealed that the students were able to correct more than three-quarters of their errors when rewriting their essays. The overall trend remained the same in the descriptive essay, with the direct group highest, followed by the coded and then the uncoded. The results clearly indicate that the most explicit form of error correction (direct feedback) led to the most successful correction, even though this may not be surprising as in this kind of

feedback the students were given the correct forms of the errors that they made and therefore they did not have to struggle to see where they went wrong. What is perhaps rather surprising is that as much as 10 to 16 percent of the errors in the two genres were still not corrected. It is also worth noting that students made higher corrections in the narrative essay compared to the descriptive essay for all the different types of feedback. This could be attributed to the fact that the narrative essay is thought to be a much easier genre for students, compared to the other genres. Also, it could be assumed that the students had had more practice in narrative essay writing compared to descriptive essay writing.

For the effect of type of feedback on different error categories, for the narrative essay, the results revealed that direct feedback led to the most successful correction in all the error categories. However, it was not automatic that it was followed by the coded feedback, as was the case in the average totals for all the errors. For example, the uncoded feedback group had more corrections in the verb, punctuation and wrong word categories compared to the coded feedback group despite uncoded feedback being the least explicit form of feedback.

For the descriptive genre the results did not duplicate what took place in the narrative essay. The direct feedback group had the highest error corrections in four out of six of the error categories, the punctuation and wrong word categories being the two exceptions. Uncoded feedback - the least explicit category - actually led to the most corrections for punctuation errors, though followed very closely by direct feedback. For the wrong word category, the coded group had the highest error correction level although this was again very close to the direct feedback group.

The results reveal that overall the rate of error correction varied between the narrative and the descriptive genres, with a clear indication that the students made more successful corrections in the narrative essay compared to the descriptive essay. The results also suggest that the more explicit a feedback type was, the higher the correction level, but not for all cases, and there were even instances where the least explicit form of error correction (uncoded feedback) led to higher correction levels than even the direct feedback.

To test the correction success hypothesis (Hypothesis 2) statistically, the total number of errors corrected and those not corrected for each of the feedback types for each genre was analysed using

Chi-square. The data presented in Table 4.7 and Table 4.8 was processed accordingly (using the VassarStats website: http://faculty.vassar.edu/lowry/vassarStats.html). Total errors corrected and not corrected in the narrative essay are shown in Table 4.7

Table 4.7: Narrative essay (a comparison of corrective feedback mechanisms)

Feedback group	Errors corrected	Errors not corrected	Total
Direct	557	64	621
Coded	447	126	573
Uncoded	426	140	563
Total	1430	330	1760

For the narrative essay, the value of Chi-square was 46.32, well above the 0.05 threshold of 7.815 (two degrees of freedom and two-tailed, given the non-directional hypothesis) and the probability value was computed as $p<0.0001$. For the narrative essay, then, there is a very significant relationship between the type of feedback provided on errors and student success in correcting them during revision. To establish if the three feedback mechanisms could be ranked in terms of correction success, subsidiary Chi-square tests were undertaken for each pair of feedback types relative to one another. These showed that direct feedback was very significantly ($p< 0.0001$) better than the other types, but the latter were not significantly different to one another. Total errors corrected and not corrected in the descriptive essay are shown in Table 4.8.

Table 4.8: Descriptive essay (a comparison of corrective feedback mechanisms)

Feedback group	Errors corrected	Errors not corrected	Total
Direct	423	81	504
Coded	364	120	484
Uncoded	304	182	486
Total	1091	383	1474

For the descriptive essay the value for the Chi-square was 59.32 which is above the 0.05 threshold of 7.815 (two degrees of freedom and two tailed) and the probability value was computed as $p< 0.0001$.

For the descriptive essay, there was thus also a very significant relationship between the type of feedback provided on errors and students' success in correcting them during revision. To establish whether the three feedback mechanisms could be ranked in terms of correction success, subsidiary Chi-square tests were again undertaken for each pair of feedback types relative to one another. It was revealed that the direct feedback was very significantly ($p<0.0009$) better than the coded feedback and the coded feedback in turn very significantly ($p<0.0001$) better than uncoded feedback and thus for the descriptive essay there is a very clear ranking.

These statistical results provide strong support for Hypothesis 2, indicating that different types of feedback can make a difference to correction success. The results showed that students achieved very significantly more success in correcting their errors during revision of both the narrative and descriptive essays if they received direct feedback. For the narrative essays there was no significant difference between the coded and uncoded feedback groups but the results for the descriptive essay indicated a definite ranking in terms of correction success from direct to coded to uncoded types. This ranking goes from most explicit to least explicit feedback and so is not unexpected (see Section 4.2), although this does not apply so clearly to the narrative essays (see Section 4.2), and various other exceptions that are discussed below with regard to specific error categories. The result is similar to that of Chandler (2003), where correction led to the highest correction success compared to the other feedback mechanisms that were used. Delgado (2007) also found that students wrote better texts after receiving explicit (coded feedback) compared to when they received less explicit feedback (uncoded feedback).

4.3 Effects of different types of feedback on students' accuracy development (Hypothesis 3)

This section deals with Hypothesis 3 namely:

> There is a relationship between the feedback mechanism used and the development of accuracy in students' writing in both overall terms and with regard to specific error categories.

In the two sections that follow, the hypothesis is first of all dealt with in general terms whereby all the errors are dealt with together, followed by the second section that looks more specifically at individual error categories.

4.3.1 Overall accuracy development

All the six types of errors were counted for each student and the totals for all the students were added for each group, as well as the means and standard deviations (SD). To control for variation in the length of the student essays, this part of the study worked with error densities rather than frequencies. The number of errors per 100 words was calculated for each student and this was used to calculate the mean density of errors per 100 words in the different groups. A t-test was then used to find out if there was any significant difference between the pretest and posttest means. The findings are shown below. A negative sign in the pretest-posttest change column in the analysis for accuracy indicates a drop in the mean density of errors.

All the errors in the pretest and posttest narrative essays were analysed for all the experimental groups and the control group and the results are shown in Table 4.9.

Table 4.9: Narrative essay (all errors combined)

	Pretest			Posttest			pretest-posttest change	t-test outcome		
Group	Totals	Mean error density	SD	Totals	Mean error density	SD		T	df	Sig (2tailed)
Direct (n=10)	226	7.11	2.12	139	4.67	1.85	-2.44	2.647	9	.027
Coded (n=10)	199	5.46	1.37	158	4.57	1.13	-0.89	2.218	9	.054
Uncoded (n=10)	207	5.45	2.04	198	5.38	2.64	-0.07	.091	9	.929
Control (n=27)	737	8.37	4.37	708	8.93	5.37	0.56	-.597	26	.555

In the pretest essay, the control group had the highest mean density of errors, with the direct group not far behind, while the coded and the uncoded groups were very similar, with considerably lower

densities of errors. In the posttest the control group error density increased, though only marginally. For all the experimental groups, however, the mean density of errors declined in the posttest. The direct feedback group had the highest mean density error decline and this was found to be statistically significant (p=.027). For the coded group the means just missed being statistically significant (p=.054). Thus the results for the direct feedback group provide support for the hypothesis with arguably some support also from the coded group results, indicating that the two most explicit feedback types appeared to be helpful to the students in improving their accuracy over time in their narrative essays.

All the errors in the pretest and posttest descriptive essays were analysed for all the experimental groups and the control group and the results are shown in Table 4.10.

Table 4.10: Descriptive essay (all errors combined)

	Pretest			Posttest			pretest-posttest change	t-test outcome		
Group	Totals	Mean error density	SD	Totals	Mean error density	SD		T	df	Sig (2tailed)
Direct (n=9)	152	5.79	3.02	137	5.52	2.90	-0.27	.447	8	.667
Coded (n=11)	196	5.75	1.78	168	5.15	1.88	-0.60	1.096	10	.229
Uncoded (n=10)	178	5.45	2.04	165	5.38	2.64	-0.07	.091	9	.929
Control (n=20)	369	6.67	3.82	402	8.00	3.85	1.33	-1.875	19	.076

The results revealed that in the pretest essay the control group here too had the highest mean density of errors, while all the experimental groups had very similar mean densities in the posttest. In the posttest the control group again showed an increase in mean density of errors which, although not significant, indicating a tendency in the direction of significance (p=.076). For all the experimental groups, the mean density of errors decreased in the posttest but none of these decreases were statistically significant.

Thus, the results for both genres showed in absolute terms that the control group students, who did not receive any feedback, worsened in accuracy over the time of this study while all the experimental groups improved, and the uncoded group, which received the least explicit feedback, showed the least improvement in respect of both genres. However, the only real support for the accuracy development hypothesis came from the statistically significant reduction in error density for the narrative essays of the direct group (and the nearly significant result for the coded group). To supplement and possibly deepen understanding of these findings, this study also examined the changes in respect of each error category, as detailed in the following section.

4.3.2 Accuracy development effects for specific error categories

This section looks at the findings for the second part of Hypothesis 3, which suggested a relationship between the feedback mechanism used and the development of accuracy with regard to specific error categories in students' writing.

Analyses similar to those in the previous section were carried out for each error category. To find out if there was any significant improvement in the individual error categories, t-tests were again carried out, whereby the pretest and posttest means were compared. The findings are arranged according to individual error categories for each feedback group, beginning with the narrative essay and then followed by the descriptive essay.

The verb errors for the pretest and in the posttest narrative essays for all the experimental groups and the control group were analysed. The results are shown in Table 4.11.

Table 4.11: Narrative essay (verbs)

	Pretest			Posttest			pretest-posttest change	t-test outcome		
Group	Totals	Mean error density	SD	Totals	Mean error density	SD		T	df	Sig (2tailed)
Direct (n=10)	76	2.44	1.75	83	2.79	1.64	0.35	-610	9	.557
Coded (n=10)	54	1.45	0.69	108	3.14	1.38	1.69	-5.093	9	.001
Uncoded (n=10)	103	2.92	1.83	119	3.29	2.60	0.37	-5.48	9	.597
Control (n=27)	325	3.73	2.51	354	4.01	2.92	0.29	-542	26	.592

The results revealed that the control group had the highest mean density of verb errors in the pretest while the coded feedback group had the lowest mean density. In the posttest the control group still had the highest mean, followed by the uncoded feedback. Rather surprisingly, for all the groups the number of errors increased and error density in the coded feedback group more than doubled, resulting in a statistically very significant difference between pretest and posttest (p=.001). This finding is indeed strange and difficult to explain but this could be because the verb form is quite wide and problematic and the students had not mastered the basics for the different forms of verb usage. For the other feedback groups, increases in the number of errors were not found to be statistically significant.

The verb errors for the pretest and posttest descriptive essays were analysed for all the experimental groups and the control group. The results are shown in Table 4.12.

Table 4.12: Descriptive essay (verbs)

	Pretest			Posttest			pretest-posttest change	t-test outcome		
Group	Totals	Mean error density	SD	Totals	Mean error density	SD		T	Df	Sig (2- tailed)
Direct (n=9)	57	2.24	1.94	75	3.03	1.89	0.79	-1.383	8	.204
Coded (n=11)	70	2.05	1.22	70	2.23	1.25	0.18	-594	10	.556
Uncoded (n=10)	79	2.36	1.23	88	2.89	1.46	0.53	-1.110	9	.296
Control (n=20)	157	3.12	2.60	171	3.31	2.02	0.19	-340	19	.737

For the descriptive essay, the results revealed that the control group had the highest mean density of errors in the pretest followed by the uncoded feedback group. In the posttest the control group again had the highest mean density of errors while the coded feedback group maintained the lowest mean density of errors. All the groups showed the same trend as in the narrative essay, namely an increase in the mean density of verb errors in the posttests, though none of the differences were found to be statistically significant for any of the groups.

All verb errors increased for all groups in both the descriptive and narrative genres even though only the coded group (narrative genre) had a significant increase. A possible explanation for this could be that the students had gained in confidence by the time of the posttest writing and had tried to focus on

writing more interesting content, but lost focus on form in the process. Another explanation for this result could have been due to the broadly defined nature of the verb category, as the study did not focus on specific types of verb error (thus a possible limitation to the study). Lee (1997) too found that the subjects found it difficult to correct verb errors in comparison to spelling, punctuation and article errors.

The results for article errors for the pretest and posttest narrative essays for all the groups are shown in Table 4.13 below.

Table 4.13: Narrative essay (articles)

	Pretest			Posttest			Pretest-posttest change	t-test outcome		
Group	Totals	Mean error density	SD	Totals	Mean error density	SD		T	df	Sig (2tailed)
Direct (n=10)	16	0.54	0.43	10	0.31	0.34	-0.23	1.272	9	.235
Coded (n=10)	11	0.26	0.26	6	0.17	0.23	-0.09	.735	9	.480
Uncoded (n=10)	4	0.12	0.21	8	0.20	0.26	0.08	-.662	9	.525
Control (n=27)	43	0.56	0.57	26	0.29	0.35	-0.27	2.178	26	.039

The uncoded group was the only group that showed an increase in mean error density, and although this nearly doubled (from 0.12 to 0.20), the numbers involved are so small that this was not at all statistically significant. Although the mean densities of the direct and the control groups showed similar changes, the change in the control group was statistically significant (p=.039). The significant drop in article error densities by the control group is difficult to explain, particularly since this did not happen in the descriptive essay discussed next.

The results for article errors for the pretest and posttest narrative essays for all the groups are shown in Table 4.14 below.

Table 4.14: Descriptive essay (articles)

	Pretest			Posttest			Pretest Posttest Change	t -test outcome		
Group	Totals	Mean error density	SD	Totals	Mean error density	SD		T	df	Sig (2tailed)
Direct (n=10)	6	0.23	0.33	13	0.58	0.76	0.35	-1.289	8	.233
Coded (n=11)	7	0.23	0.32	9	0.31	0.35	0.08	-.711	10	.493
Uncoded (n=10)	11	0.33	0.32	8	0.27	0.43	-0.06	0.605	9	.560
Control (n=20)	19	0.34	0.38	15	0.31	0.43	-0.03	.221	19	.828

As with the narrative essays, the article errors in the descriptive essays show low numbers and mean densities, making it more difficult to meet statistical requirements for significance. In both the direct and coded feedback group, the mean density of errors increased in the posttest with the direct group having the highest mean density increase of errors. However, when the means for each of the groups were compared, the increases were not found to be statistically significant. The uncoded group and the control group showed a decline in the mean density of errors but this too was not found to be statistically significant.

The results for the preposition errors are shown in Table 4.15.

Table 4.15: Narrative essay (prepositions)

	Pretest			Posttest			pretest-posttest change	t-test outcome		
Group	Totals	Mean error density	SD	Totals	Mean error density	SD		T	df	Sig (2tailed)
Direct (n=10)	25	0.77	0.55	12	0.37	0.34	-0.40	1.922	9	.087
Coded (n=10)	11	0.32	0.28	7	0.17	0.19	-0.15	1.649	9	.134
Uncoded (n=10)	28	0.82	0.86	16	0.48	0.48	-0.34	1.527	9	.161
Control (n=27)	60	0.70	0.53	65	0.84	0.60	-0.10	-0.854	26	.401

All the three experimental groups as well as the control group showed a decline in the mean density of errors in the posttest, with the direct feedback group having the highest drop. When the mean error densities of each of the groups were tested for statistical significance there were no such differences, although the direct feedback group's improvement tended toward significance (p=.087). Despite the density of errors declining by more than half, a significant result was not reached, partly because the standard deviations were high compared to the means, which is an indication that the students in this group had their results varying a lot.

The results for preposition errors in the descriptive essays are shown in Table 4.16 below.

Table 4.16: Descriptive essay (prepositions)

	Pretest			Posttest			pretest-posttest change	t-test outcome		
Group	Totals	Mean error density	SD	Totals	Mean error density	SD		T	df	Sig (2tailed)
Direct (n=9)	25	0.88	0.97	9	0.38	0.33	-0.50	1.334	8	.219
Coded (n=11)	16	0.51	0.50	17	0.50	0.46	-0.01	.048	10	.963
Uncoded (n=10)	10	0.64	1.10	8	0.23	0.27	-0.37	1.060	9	.317
Control (n=20)	20	0.34	0.37	21	0.53	0.44	0.19	-1.502	19	.490

When the pretest means were compared with the posttest, it was found that in the control group, the mean density of errors increased, but this was not statistically significant. All the experimental groups showed a decline in the mean density of errors in the posttest, but none were statistically significant. The direct group had a very high drop in the mean density of errors but no statistical significance resulted, largely because here too there were high standard deviations, an indication that the individual students' results varied widely.

Punctuation error findings are shown in Table 4.17.

Table 4.17: Narrative essay (punctuation)

	Pretest			Posttest			pretest-posttest change	t-test outcome		
Group	Totals	Mean error density	SD	Totals	Mean error density	SD		T	df	Sig (2tailed)
Direct (n=10)	32	1.06	1.47	9	0.33	0.48	-0.73	1.438	9	.148
Coded (n=10)	17	0.46	0.46	17	0.51	0.48	0.05	-.246	9	.811
Uncoded (n=10)	12	0.40	0.42	28	0.72	0.54	0.32	-1.790	9	.107
Control (n=27)	39	0.46	0.49	57	0.67	0.79	0.33	-1.109	26	.278

Interestingly in the posttest, the direct feedback group had the lowest mean density of errors yet it had the highest density of errors in the pretest. This does not however lead to a statistically significant result, again largely because of the very high standard deviations in both the pretest and posttest, where they are higher than the means. All the other groups showed increases in error densities but none were significant.

The punctuation error results for the descriptive essay are shown in Table 4.18 below.

Table 4.18: Descriptive essay (punctuation)

	Pretest			Posttest			Pretest-posttest change	t-test outcome		
Group	Totals	Mean error density	SD	Totals	Mean error density	SD		T	df	Sig (2tailed)
Direct (n=9)	13	0.50	0.54	6	0.30	0.51	-0.20	1.007	8	.343
Coded (n=11)	21	0.68	0.44	25	0.74	0.63	0.06	-.258	10	.802
Uncoded (n=10)	25	0.79	0.74	18	0.56	0.46	-0.23	.784	9	.453
Control (n=20)	31	0.51	0.56	27	0.54	0.66	0.03	-.203	19	.842

For the punctuation error category in the descriptive essay, the direct and uncoded feedback groups had their mean density of errors decline in the posttest, but there was no statistically significant difference. On the other hand, the coded and the control groups had their errors increase in the posttests, though again there was no statistically significant difference.

The spelling error results for the narrative essays are shown in Table 4.19 below.

Table 4.19: Narrative essay (spellings)

	Pretest			Posttest			Pretest-posttest change	t-test outcome		
Group	Totals	Mean error density	SD	Totals	Mean error density	SD		T	df	Sig (2tailed)
Direct (n=10)	48	1.40	1.10	16	0.56	0.62	-0.84	2.159	9	.059
Coded (n=10)	52	1.40	0.83	6	0.19	0.36	-1.21	3.800	9	.004
Uncoded (n=10)	40	0.94	0.79	12	0.32	0.48	-0.62	1.870	9	.094
Control (n=27)	127	1.52	1.08	114	1.58	1.92	0.06	-197	26	.845

For the spelling error category in both the pretest and the posttest, the control group showed the highest mean error density with a slight but totally insignificant increase in the posttest. On the other hand, all the experimental groups showed a decline in the mean density of errors in the posttest, and this decline was statistically very significant (p=.004) in the coded feedback group, and tended towards significance in the direct feedback group (p=.059) and the uncoded groups (p=.094). With the very marked decline in errors in the latter two groups, a statistically significant difference might have been expected, but this was not so, again because of the standard deviations being very high relative to the means. This shows that the individual results in these groups varied widely and so affected the results.

The spelling error results for the descriptive essays are shown in Table 4.20.

Table 4.20: Descriptive essay (spellings)

	Pretest			Posttest			pretest-posttest change	t-test outcome		
Group	Totals	Mean error density	SD	Totals	Mean error density	SD		T	df	Sig (2tailed)
Direct (n=9)	37	1.40	0.84	16	0.56	0.56	-0.84	2.80	8	.023
Coded (n=11)	63	1.80	1.04	10	0.36	0.61	-1.44	3.961	10	.003
Uncoded (n=10)	22	0.71	0.55	5	0.13	0.24	-0.58	3.086	9	.013
Control (n=20)	86	1.59	1.32	68	1.35	1.17	-0.24	1.167	19	.258

In the spelling error category (descriptive essay), all the experimental groups and the control group showed declines in the mean density of errors made in the posttest. The highest decline was found in the coded feedback group, while the control group had the lowest decline. When tested for statistical significance, the coded feedback group revealed a very significant (p=.003) change and the uncoded (p=.013) and direct groups (p=.023) revealed significant changes. Thus all the experimental groups showed significant improvement in spelling, while the control group's change was not significant

The results for the spelling error category show this as the only error category where, for both the descriptive and narrative genres, all the experimental groups showed significant or near significant results, while the control group showed no significant improvement. In both genres the coded group showed very significant results, and this could be attributed to the nature of the type of feedback that they received. As the error feedback was coded and not fully explicit as in the direct feedback group, the students should have had to think more deeply about their corrections, and so they were more likely to learn from the experience over time. Unlike the uncoded feedback, the coded group benefited from at least being informed that the relevant errors were spelling errors and nothing else. The direct feedback group might have been expected to perform better than the other groups but this was not so. This contrasts with the finding of Hypothesis 2, where the direct feedback group had the highest error corrections compared to the other groups. This contrast could be attributed to the direct feedback students not paying as much attention to the errors that they corrected, as they did not have to struggle

to correct their spellings and so did not learn as much from the experience. Some spelling accuracy development does however appear to have taken place amongst all the feedback groups and this could suggest that teacher corrective feedback led to learning because the control group did not show any significant improvement in either of the two genres.

The wrong word error results for the narrative essays are shown in Table 4.21.

Table 4.21: Narrative essay (wrong word)

	Pretest			Posttest			Pretest-posttest change	t-test outcome		
Group	Totals	Mean error density	SD	Totals	Mean error density	SD		T	df	Sig (2tailed)
Direct (n=10)	29	0.89	0.51	9	0.28	0.27	-0.61	3.417	9	.008
Coded (n=10)	54	1.51	1.12	14	0.40	0.37	-1.11	2.956	9	.016
Uncoded (n=10)	19	0.61	0.76	20	0.53	0.40	-0.08	.340	9	.741
Control (n=27)	143	1.76	1.46	95	1.25	1.17	-0.51	2.001	26	.056

In the narrative essay (wrong word), for the pretest, the control group revealed a near-significant decline in mean density of errors. The direct feedback group revealed a statistically very significant decrease (p=.008) and the coded feedback group a significant decrease (p=.016). The uncoded group showed a slight decline, but of no statistical significance. Even without feedback, students in the control group managed to make fewer wrong word errors in the posttest, which is an indication that to an extent student can deal with untreatable errors without help from a teacher.

The wrong word error results in the descriptive essays are shown in Table 4.22.

Table 4.22: Descriptive essay (wrong word)

	Pretest			Posttest			Pretest-posttest change	t-test outcome		
Group	Totals	Mean error density	SD	Totals	Mean error density	SD		T	df	Sig (2tailed)
Direct (n=9)	14	0.52	0.36	18	0.67	0.40	0.15	-.871	8	.409
Coded (n=11)	19	0.51	0.38	37	1.05	0.97	0.54	-1.971	10	.077
Uncoded(n=10)	32	0.99	0.62	37	1.24	0.80	0.25	-.986	9	.350
Control (n=20)	45	0.78	0.53	93	1.93	1.18	1.15	-4.896	19	.000

When the wrong words for the descriptive essay were analysed, they revealed totally different patterns to those of the narrative essays. For all the groups the mean density of errors increased. The highest mean error density increase was in the control group which showed a statistically very significant increase ($p=.000$ according to SPSS output, i.e. $p<.0001$). The increase for the coded feedback group was found to be nearly significant. The uncoded and direct feedback groups' increases were not significant. The increase in the wrong word errors could be attributed to the topic that the students wrote on in the posttest, which could be considered to have been unfamiliar to the one that they wrote on in the narrative posttest (Appendix 1). The narrative pretest topic was 'The most interesting holiday that I have ever had', while the posttest topic was 'Write a short story on the following topic: A midnight visitor.' The descriptive pretest topic was 'Describe your room after a party is over and the guests have left', while the posttest topic was 'Describe a natural disaster you have witnessed.' For instance, some of the posttest essays written by some of the students were considered to be off topic as some students may never have experienced a disaster or could not imagine on a midnight visitor, so it was likely that some students were not familiar with it completely and so they were likely to make some wrong word choices.

In summary, with respect to specific error categories, the verb error category showed lack of improvement for both genres as the verb errors increased for all feedback groups. In one instance, in the narrative essay (coded feedback), the increase in verb errors was significant. For all groups in both genres, the article errors dropped but none was significant. In the preposition category, for both genres,

it was only the direct feedback group (narrative essay) that tended towards being significant (p=.087). For the punctuation error category no group showed any significant improvement in either genre, while for the wrong word error category, the number of errors showed no consistency. The spelling error category, however, showed consistent improvement for all the experimental groups in both genres, and so it is only with regard to spelling errors that Hypothesis 3 is supported. This result contradicts that of Fazio (2001), where spelling errors increased regardless of feedback, but it is similar to that of Lee (1997), where the spelling error category was one where students improved most. A second relatively consistent finding, applicable particularly to the spelling errors, is that the coded feedback group often revealed more improvement in accuracy than any of the others, including the direct feedback group. This suggests that correction should not be too explicit and that students should have to do a certain amount of cognitive work when correcting as they are then more likely to learn from the experience.

To shed a more qualitative light on the nature of accuracy development on certain individual error categories, and to illustrate the kind of writing dealt with in this study, the next section traces the progress of a small sample of students more specifically, namely those who showed the most improvement in spelling. This category was focused on because it was the only one to show significant improvements across all the feedback groups.

4.4 Sample students' accuracy development: spelling errors

To analyse individual student performance, one student who was found to be the most improved in terms of error reduction in the spelling category was selected from each of the experimental feedback groups for each genre. This was done on the basis of the individual improvement for the two genres taken together. This was to help make comparisons between the two genres and see whether any improvement took place. Their performance was followed in terms of the spelling errors that were made in the pretest and posttest in these students' essays. The narrative genre is dealt with first followed by the descriptive genre for each student (transcripts of the essays including the spelling error feedback are provided in Appendix 5). Each student's script was indexed. The direct feedback scripts received were indexed as D along with a number for each individual that is D1, D2 and so forth. The coded scripts received were indexed as C, the uncoded feedback scripts received the index UN while the control group scripts received were indexed as Cont. Scripts for the following students were selected: D8 (direct feedback group), C9 (coded feedback group) and UN6 (uncoded feedback group).

For the direct feedback group (narrative essay), D8 had the highest improvement in mean density of spelling errors.

In the pretest, D8 made the following spelling errors:

… *themselfs* = … themselves

…*practicing* = … practising

… *relux* = … relax

…*well come* = … welcome

…*quete* = … quiet

… *massage* (×2) = message

In the posttest D8 made the following spelling errors:

…*quete* = …quiet

…*scrim* = …scream

We have visitors you now = …We have visitors you know.

…*this drinks* =…these drinks

…*my self* = …myself

The length of the two essays is actually almost the same and the student made almost the same number of errors in the posttest as in the pretest for this genre. This shows that not much improvement had been made. The errors varied but the word *quiet* is misspelt in the same way in both the pretest and posttest. This shows that the student had not learnt from the feedback that had been given in connection with this error. This student also had problems distinguishing the different vowels in *this* and *these* most probably because they are pronounced in the same way – a common problem amongst many African language learners of English.

For the descriptive pretest essay the following spelling errors were made by D8.

…*tryed to explain* = …tried to explain

…*to live my room* = … to leave my room

… *did not lessen* = … did not listen

… *I feel asleep* = … I fell asleep

I desided to go… = I decided to go…

I will never plane … = I will never plan….

In the posttest the D8 made three errors only and these were:

…*stricking* = striking

…*prefered* =…preferred

…*lestened* =… listened

Student D8 made fewer errors in the posttest compared to the pretest even though the length of the posttest essay is slightly longer than the pretest essay. This is an indication of improvement in the spelling error category. The word *listen* is misspelt in the pretest and in the posttest essay, *listened* is misspelt, showing that the student had not benefited from the feedback on this word although some progress is shown as the student remembered the *t*. The student seemed to have problems with the high front vowels and diphthongs, as shown in two other instances: *feel asleep* instead of *fell asleep* and also *live my room* instead of *leave my room*. For both genres the student made fewer errors in the posttests and this is an indication of improvement in the spelling category.

For the coded feedback group, C9 showed the most improvement in the spelling error category and was selected and the errors made are discussed below, beginning with the narrative essay.

The pretest narrative essay errors were:

… *draggy* = … dragging.

…*fantisies* =…fantasies

… *laying on my back* =…. lying on my back ….

… *my self* = … myself

…*this ideas* = …these ideas

…*foot steps* = …footsteps

… *tlee* =… tree

…*any thing* = …anything

…*signes* =…signs

I stood steel = … I stood still.

… *fling* = … flying

For the posttest narrative essay C9 made fewer spelling errors in comparison with the pretest essay. These were:

… *spacking* = … sparking

….*lay on my back* = …lie on my back.

All this ideas… = All these ideas…

This student too, like the above student in the direct feedback group (D8), had a problem with differentiating between the words *this* and *these* which could, as mentioned be due to pronunciation problems, with many African language speakers not distinguishing between them in their own pronunciation. This error was a repeat of what had been made in the pretest, indicating that the student had not benefited from the feedback correction that was given as regards this error. In the pretest the word *myself* is misspelt, but in the posttest it was correctly spelt showing improvement as regards this word.

For the pretest descriptive essay C9 made the following spelling errors.

… *left overs* = … leftovers

… *leave the house daddy* = … leave the house dirty.

…*any one* = … anyone

… *they were laying on the floor* = … They were lying on the floor.

… *The door of my wall drop* = … the door of my wardrobe

…*we jumping* = …were jumping.

… *injoing* = … enjoying…

… *apologies* = … apologised

… *excepted their kindness* = … accepted their kindness

… *leaving in a pretty room* = … living in a pretty room

For the posttest essay the C9 made fewer spelling errors in comparison with the ones made in the pretest. These were:

I scrimmed =… I screamed….

… *at a perticular time* = … at a particular time.

It was noted that in the posttest this student was able to use the word *live* correctly. Apparently the student from the direct feedback group and the one from the coded group seemed to have two common spelling errors. Both students missspelt *scream* and *lie*, suggesting that these were problematic words

that needed attention to help the students master them. The first probably relates to the pronunciation problems in distinguishing between the two high front vowels of English and the second could be the result of confusion between the words *lie* and *lay*.

Student C9 reflects the finding for the coded group that it showed the highest significance in improvement in the spelling error category for both the narrative and descriptive genres. It indeed can be seen from this student's result that there was a considerable drop in spelling errors from the pretest to the posttest. The data for each student is limited, but the coded feedback mechanism seems to have assisted this student improve in spelling.

In the uncoded feedback group, the student with the code UN6 showed the most improvement in the spelling error category. The errors that UN6 made are discussed below, beginning with the narrative essay. The following words were misspelt in the narrative essay:

…*they* =… there

… *where* ×2 = were

… *sight* = … side

… *beatiful* ×2 = … beautiful

… *behatiful* = … beautiful

… *u* = … you

… *my self* = … myself

… *vomite* = … vomit

… *their* = …there

… *sea side* = … seaside

… *oclock* ×3 = … o'clock

This particular student had numerous spelling errors in the narrative pretest with some being repeated. In this essay this student spelt *vomited* correctly but misspelt the word *vomit*, showing that this spelling had not been mastered. The student also has used the shortened version of *you* by just using the letter *u*, which, like the use of *nite* instead of *night* in her posttest, is a problem that may have arisen due to the use of such forms in short text messages in cellphones and so the student assumes that they are acceptable in formal writing or just forgets to use the standard form, possibly because a lot of her writing is now done on cellphone.

For the posttest narrative essay, the following words were misspelt. These were:

…*entairtnment* = … entertainment

… *their* = … there

…. *Nite* = … night

… *tryed* = … tried

… *they* = … there

… *oclock* ×5 = … o'clock

In this essay, this particular student wrote a longer essay in the posttest and made fewer errors compared to the pretest. However, the student repeats the errors *they*, *their* and *oclock* in the posttest, showing that she had not learnt from the feedback that had been given. The word *tried* is misspelt as *tryed*, showing that the student has not mastered some basic spelling rules such as the change from the infinitive form to past tense. The same error was noted with student D8 in the descriptive pretest essay.

For the pretest descriptive essay the following words were misspelt.

…*where*=… were

… *their* = … there

… *they* = … there

… *oclock* ×4 = … o'clock

These four errors also appeared in the narrative pretest essay, showing that they were problematic words to this particular student. The student however spelt the word *beautiful* correctly thrice, unlike in the narrative essay where the word had been misspelt three times. The feedback that had been given in the narrative pretest since it was given first had clearly helped this student.

For the posttest essay the student made the following errors:

they ×2 =… there

where = … were

…*my self* ×3 = …myself

… *their* = …the

This student repeated the errors *they* and *where* that appeared in both narrative and descriptive pretests while *my self* had also appeared in the narrative pretest. This is a clear indication that the three words remained problematic and the student had not learnt from the feedback given.

This illustrative account of individual spelling errors reveals that the selected students sometimes repeated errors made in the pretest, showing that they had not mastered the correct spelling of those words. However, in some instances the students appeared to have benefited from the feedback, suggesting that some learning had taken place. There were some common errors that were noted amongst the three students discussed above. The first one was the confusion of the words *their* and *there*, which are homophones. The two words are used interchangeably and the difference has not been understood fully. There is also confusion of the word *they* and the possessive form *their*. Another common error was the splitting of single words into two words, such as *myself*, which was misspelt as *my self* by all the three students above. This is an indication that the students had not yet mastered such spelling basics of the language.

There were instances where students learnt from the feedback that was given. For example, studentC9 misspelt the word *myself* in the narrative essay pretest but spelt it correctly in the posstest. Student Un6 misspelt the word *beautiful* in the narrative pretest essay but spelt it correctly in the descriptive pretest essay that was written after the former. This is an indication that some learning had actually taken place.

4.5 Students' opinion of teacher error correction feedback

The study sought to find out the students' opinions about teacher error correction. The students were given a questionnaire to complete that comprised five questions and the responses were assessed on a five-point Likert scale (see Appendix 3). A total of 32 students from the experimental groups filled in the questionnaire. The direct feedback group comprised 10 students while the coded and uncoded groups comprised 11 students each. The first three questions sought information on the same area but were different to some extent. For Question 1 on whether students were able to see the kind of errors they made, the aim was to find out whether they were able to identify errors that were corrected by the teacher, since the feedback mechanisms used were not all very explicit. The second question on whether the feedback given helped the students to do corrections was to find out whether feedback

helped them recognise the kind of errors made and so make the corrections. The third question on 'understanding' sought to find out whether the students actually understood the kind of feedback that the teacher used.

The questionnaire was given to the experimental groups only since they were the groups that received teacher correction feedback and it was important to find out the extent to which the students were affected by the kind of feedback that they received. This kind of questionnaire would not have been relevant to the control group as this group received no teacher correction.

The responses to the questions have been arranged on a question-to-question basis so as to make comparisons across groups easy. In terms of the five-point Likert scale, 'always' scored 4, 'most of the time' scored 3, 'sometimes' scored 2, 'rarely' scored 1, while 'never' scored 0. However, Question 4, which related to discouragement, had its values reversed because here the students' opinion was being assessed on a negative matter. All values for individual responses for each group were calculated, added and the result was divided by the number of members in each group. The result was the mean value of each group per question. Then the overall totals of the three groups for all the five questions were compared using the Chi-square statistic. The results of the questionnaire are discussed below.

The first question required the students to indicate whether it was easy to see the kind of errors that they made in their writing. The students were required to rewrite their essays after they had received teacher correction feedback and so it was essential to find out if they noticed the errors that they made. The responses to this question are presented in Table 4.23.

Table 4.23: Was it easy to see the kind of errors that you made?

Students' Response							
Feedback group	**Always (4)**	**Most of the time (3)**	**Sometimes (2)**	**Rarely (1)**	**Never (0)**	**Total**	**Mean**
Direct (n=10)	4	3	3	0	0	31	3.1
Coded (n=11)	3	2	3	3	0	27	2.5
Uncoded (n=11)	3	5	2	1	0	32	2.9

In response to this question the direct group had the highest mean of 3.1 showing that in this group the students felt that that they were able to notice the errors on average 'most of the time'. This would be

expected since the students in this group were given the correct forms of the errors made, hence they had only to rewrite their work filling in the correct forms already given by the teacher. The uncoded group had a mean of 2.9, which also indicated that the students felt it was easy to correct their errors at least 'sometimes' and mostly 'most of the time'. The coded group had the lowest mean in this question, the mean of 2.5 being an indication that the students in this group on average found it easy to see the kind of errors they made 'sometimes'. In the correction success hypothesis analysis, the direct feedback showed the highest corrections in both genres and this result confirms that students who received direct error correction found it easy to see their errors.

The second question required the students to indicate whether the correction they received helped them to correct the errors they made. The results are presented in Table 4.24 below.

Table 4.24: Did correction help you correct the errors you made?

Students' Response							
Feedback group	Always (4)	Most of the time (3)	Sometimes (2)	Rarely (1)	Never (0)	Total	Mean
Direct (n=10)	6	4	1	0	0	38	3.8
Coded (n=11)	6	4	1	0	0	38	3.5
Uncoded (n=11)	6	3	1	0	1	35	3.2

In response to this question, all the groups had very high means. For all the groups the students found the correction helpful to correct the errors they made most of the time. The direct feedback had the highest mean of 3.8 while the uncoded feedback had the lowest mean of 3.2. The responses indicate that the students felt that the teacher error correction was helpful. Again when the students rewrote their work, the direct feedback group had the highest error percentage correction of errors while the uncoded feedback had the least in both genres, and it appears that the more explicit the feedback received, the more the students also appreciated the feedback.

For the third question the students were asked if they understood the kind of correction that they received from the teacher. Since the students in the experimental groups received different types of feedback, it was essential to find out if they understood the type of feedback that they received. The responses are presented in Table 4.25.

Table 4.25: Did you understand the correction that the teacher gave you?

Students' Response							
Feedback group	Always(4)	Most of the time (3)	Sometimes (2)	Rarely (1)	Never (0)	Total	Mean
Direct(n=10)	5	1	2	2	0	29	2.9
Coded (=11)	7	3	1	0	0	39	3.5
Uncoded (n=11)	5	3	3	0	0	35	3.2

The highest mean for this question was from the coded feedback group, namely 3.5, showing that the students in this group understood the correction that the teacher gave most of the time. The uncoded group had a similar response with a mean of 3.2. The direct group had the lowest mean of 2.9 showing that the students understood the feedback that they were given only sometimes. This was an interesting response because for this feedback group the students were given the correct forms for all the errors that they made. It then becomes contradictory that the students in this group say that they understood the feedback that they got only sometimes. Maybe the coded and uncoded groups thought they understood more because they had to think harder about the feedback.

The fourth question required the students to say whether they got discouraged by the way the teachers corrected their errors. Bearing in mind that the students wrote eight original essays, with six rewrites apart from the posttest essays, within one school term and each received the same kind of feedback over the period, it was important to find out if they got discouraged. The results on the fourth question are presented in Table 4.26.

Table 4.26: Did you feel discouraged by the way the teacher corrected your errors?

Students' Response							
Feedback	Always (0)	Most of the time (1)	Sometimes (2)	Rarely (3)	Never (4)	Total	Mean
Direct	1	1	2	1	5	28	2.8
Coded	1	0	1	0	9	38	3.5
Uncoded	0	2	2	1	6	33	3.0

In response to this question on whether students got discouraged when they received teacher correction, the coded group showed the highest totals as well as means for this question while the direct feedback group had the lowest. The coded group had nine out of 11 students saying 'never', the direct and coded group had half of the students saying they never got discouraged. The fact that the coded group, which often showed the most positive results in terms of accuracy development, also felt least discouraged, points to a positive link between students' feelings about the different feedback types and the success the different types delivered. The overall finding here is similar to that of Lalande (1982), where students did not get discouraged by the kind of feedback that they received, and in Chandler (2003), where only one out of all the participants felt discouraged after receiving correction feedback. The last question required the students to state how often they would like to have their errors corrected. The response is shown in Table 4.27.

Table 4.27: How often would you like your teacher to correct your errors?

Students' Response							
Feedback group	Always (4)	Most of the time (3)	Sometimes(2)	Rarely (1)	Never (0)	Total	Mean
Direct (n=10)	5	1	3	0	1	29	2.9
Coded (n=11)	5	4	2	0	0	36	3.3
Uncoded (n= 11)	2	1	6	2	0	25	2.3

The results for this question showed that the coded feedback group had the highest mean (3.3). The students of this group would like to have their errors corrected most of the time. Response from the direct and uncoded feedback showed that both groups would have liked to have their errors corrected sometimes, though the direct feedback group had a mean of 2.9 which was quite close to 'most of the time'. The uncoded group had a mean of 2.3.indicating that they wanted feedback only sometimes. This could be attributed to the fact that because this is the least explicit feedback type, some of the students never understood the correction that was expected of them. This response is similar to that in Greenslade and Felix-Brasdefer (2006), where learners who received error feedback that was underlined only were not sure about the type of correction that was expected.

The results from the student questionnaire reveal a strong implication from all the groups that they wanted their errors corrected, as they felt that it was beneficial to them and that they rarely got

discouraged. This concurs with Lalande (1982) and Ferris and Roberts (2001), whose studies revealed that students wanted corrective feedback of some kind from their teachers.

To see whether there were overall differences between the three groups, the totals were calculated for all the five questions and the result was tested for significance using the Chi-square. The totals are shown in Table 4.28.

Table4.28: Overall totals for the questionnaire questions

Group	Question 1	Question 2	Question 3	Question 4	Question 5	Total	Mean
Direct	31	38	29	28	29	155	3.1
Coded	27	38	39	38	36	178	3.2
Uncoded	32	35	35	33	25	160	2.9

The Chi-square statistic was calculated using VassarStats and it was revealed that there was no significant difference between the groups (Chi-square = 3.86 for 8df: p=.8695). It is worth noting that the coded feedback group performed best in three of the questions (Questions 3, 4 and5). This group seems to have performed somewhat better than the other groups in terms of accuracy development, as revealed in the spelling error category for example, but any link between positiveness about correction and accuracy development should be treated with caution, as differentiation with the other groups is not significant.

4.6 Conclusion

In terms of the fluency hypothesis (H1), for the three experimental groups the results revealed that the students did not write longer texts in the posttests. Rather, in both the narrative and descriptive essay they actually wrote shorter texts in the posttests, though the differences were not statistically significant. The control group showed a decline in the mean number of words in the narrative essay, while in the descriptive essay the same group wrote a higher number of words although this was not significant. Thus the students in this study did not show any improvement in fluency with or without feedback and so the study does not support Hypothesis 1, namely that there is a relationship between the type of feedback given and the development of fluency in student writing. The result of this study

concur with the findings of Frantzen (1995) and Robb et al. (1986), where subjects in the two studies did not improve in fluency, but contradicts that of Chandler (2003), where there was a significant improvement in fluency.

Hypothesis 2 was concerned with the extent to which students succeeded in correcting their errors when doing revision. The results reveal that students managed to correct most of the errors they made in the original essays. This is similar to the findings of Lee (1997), Ferris and Roberts (2001) and Fathman and Whalley (1990), who found that students' writing improved in the short term. When the data was tested for statistical significance, it was found that there was a relationship between the feedback given by the teacher and students' improvement in correcting their work, with direct feedback significantly higher than the others in both genres and coded feedback next highest in both genres and significantly higher than uncoded feedback in the descriptive essay. This study thus supports Hypothesis 2, which states that there is a relationship between the type of teacher correction feedback provided on student errors and student success in correcting errors, indicating that the more explicit the feedback, the more success students have in correcting during revision.

Hypothesis 3 sought to find out the effect of feedback mechanisms on the development of students' writing accuracy over time. Overall, when all error categories were considered together all three experimental groups revealed that all the errors decreased in both the narrative and descriptive essays. In the narrative essay the decrease in errors for the direct and coded feedback essays was significant and nearly significant ($p \leq .10$) respectively. The control group however showed an increase in the number of errors in the posttest in both the narrative and descriptive essay and this was nearly significant in the descriptive essay.

When the individual errors were studied the results revealed verb errors increased for all groups in both the narrative and the descriptive essays but this was significant only for the coded group (narrative essay). This error category has been found to be problematic in other studies as well (Schlig 2005, Lee 1997). It is a very broadly defined category and this study, like nearly all the others, did not look at subcategories of verb error such as tense, mood, voice, aspect and modality. Probably if the study had looked at specific aspects of the verb error, more definite findings may have resulted, and future studies should take this issue into consideration.

In the article error category, the errors increased for the direct and coded group in the narrative essay while this was the case for the uncoded group in the descriptive essay. For all the other groups the errors declined in both the narrative and descriptive essay. For the control group, in the narrative essay the result was significant (p=.039). This finding is difficult to explain. This particular category was problematic to the students, as was the case in Hong (2004) although the same was easy to correct in the study by Lee (1997).

For the preposition error category, there was a decline for all the groups for both genres except for the control group in the descriptive essay, which had its error density increase, but this was not significant. In the narrative essay, the direct group's error decrease inclined towards significance. The decline in the preposition error category sends a positive signal in that even if the finding was not significant, at least the errors declined in all cases. A study carried out for a longer period could help confirm if there would be improvement in this category or not.

In the punctuation error category, errors declined in the direct feedback and uncoded feedback groups for both genres in the descriptive essay, however none of these declines was significant.

For the spelling category, all the experimental groups showed a decline in both genres and the results were that the coded feedback showed very significant improvement while the direct and the uncoded feedback groups had nearly significant results for both genres. As for the control group, the spelling errors increased in the narrative genre, but there was a decline in this category in the descriptive essay. For both cases, though the results were not significant. This result suggests that when students receive feedback they relatively easily notice their spelling errors and so often correct them but without feedback they do not notice the errors. There is an implication here that students need to be helped to notice their errors.

Lastly, for the wrong word error category, in the narrative essay there was a decrease of errors for all groups and the direct group showed a significant decrease while the coded group and the control group revealed a nearly significant decrease. Interestingly, in the descriptive essay wrong word errors increased for all groups, though only the control group showed a significant increase.

The results for all error categories other than for the spelling error category are not consistent. For instance, for the wrong word error category, the control group shows a nearly significant decline in the narrative essay and on the other hand a significant increase in the descriptive essay. The spelling error category for both genres was the only result that clearly supported Hypothesis 3, which postulates that there is a relationship between the type of feedback given and students' development in accuracy, as it is here that we see significant improvement, as well as definite contrast with the control group.

With regard to student opinions, the direct feedback group found it easiest to see the kind of errors made and this was reflected in the narrative essay, where there was a significant improvement as regards error correction. Again, as regards whether correction helped the students correct their errors, the direct feedback group had the highest mean, which seems to reflect the fact that this group had the highest correction percentage in both genres. The coded group said they understood the correction that was given, they never got discouraged by the way the teacher corrected their errors and they wanted their errors corrected more often compared to the other groups. Students in general showed a need to be given feedback by their teachers and this is similar to findings in other studies, such as those of Lee (1997), Lalande (1982) and Ferris and Roberts (2001). Overall, there was no significant difference about this between the groups.

CHAPTER 5

CONCLUSION

5.0 Introduction

This concluding chapter comprises four main sections. Section 5.1 reviews the key findings of the study and discusses its contribution, while Section 5.2 discusses some of its limitations. This is followed by Section 5.3, which deals with implications and suggestions for further research, while section 5.4 concludes the study.

5.1 Contribution of the study

This section starts with a few points about the contribution of the study to existing research with regard to existing research and then it provides a final review of the findings of the study, briefly comparing them with previous research.

As regards the methodology, Ferris (2004, 51) emphasises the point that "very few studies of error correction in L2 writing actually 'compare the writing of students who have received grammar correction over a period of time with that of students who have not"'. The present study thus makes a contribution towards the limited group of quasi-experimental longitudinal studies that have been carried out. It has a pretest-posttest design and a control group as well as three experimental groups. Most of the other studies that have been done have lacked one or more of these features, which are important for reliability and validity.

With regard to Hypothesis 1, the study found that rewriting after receiving teacher correction feedback did not lead to students improving in their fluency. For all the feedback mechanisms except for the uncoded and the control group in the narrative essay, students actually wrote fewer words in the posttests compared to the pretests. There were no statistically significant changes at all, with the only tendency toward significance shown by the decrease in the posttest word numbers for the descriptive essays of the control group. This finding is similar to that of Robb et al. (1986) and Frantzen (1995), where subjects did not improve in fluency either. The present study also showed that there was no relationship between the feedback mechanism used and fluency. One possible reason for these findings

could be the fact that the students, having had English as the LoLT for many years, were already quite fluent writers at the start of the study and so had less room for improvement (a 'ceiling' effect). Hypothesis 1 was therefore not supported by the study.

With regard to Hypothesis 2, the study revealed that there was a strong relationship between the feedback mechanism provided on student errors and students' success in correcting errors. In both genres, the direct feedback mechanism led to the highest percentage of corrections in the revised versions of the essays. The most explicit form of error correction (direct feedback) helped the students to correct most of their errors. Hypothesis 2 was therefore supported by the study. This finding concurs with those of Chandler (2003), Fathman and Whalley (1990), Hong (2004) and Ferris and Roberts (2002), where rewriting led to more accurate written texts.

For Hypothesis 3, on whether students who received teacher correction feedback improved the accuracy of their writing over time more than those who did not, the study revealed that there was a decline in the overall error densities of all the experimental groups in both genres, but it was only in the narrative essays of the direct feedback group that a significant decline resulted, while those of the coded group were nearly significant (i.e. $p \leq 0.10$). The results of these two feedback mechanisms thus give some support to Hypothesis 3 for the narrative genre, namely that there is a relationship between the feedback mechanism used and the development of accuracy in student writing. The results also revealed that for the control group in both genres, the error densities of the students actually increased and in the descriptive essay the result was nearly significant. This seems to suggest that lack of feedback led to the students writing less accurately, an indication that students need teacher correction feedback in order to improve in their writing. However, the result should be treated with caution because for the descriptive essays none of the feedback types led to significant improvement in students' accuracy over time.

For the individual error categories, the verb error density increased in all the feedback mechanisms in both genres but these were not significant except for the coded feedback (narrative essay) where the increase was very significant. It is difficult to explain these results, though one factor could be the nature of the topics in the pretests and posttests. The topics for the posttests were possibly more unusual, more interesting and more challenging and students could have therefore concentrated more

on the content than grammar, so making more errors. This finding is similar to that of Lee (1997), where students had most problems with verb errors but contradicts Sheppard (1992), where students improved in the verb category significantly.

The study showed that some errors were easier to reduce over time than others. The verb error category was problematic as all groups showed an increase in error densities which concurs with Schlig (2005) and Lee (1997) where verb errors were difficult to correct. For the article errors, there was a decline for all groups except for the control group (narrative essay) that was nearly significant concurring with Lee (1997) where article errors were found to be easy to correct. For the punctuation and the preposition error categories, there was no significant result for any of the groups in both genres.

The wrong word error category also showed improvements in accuracy to some extent, as the coded group improved significantly or nearly significantly in both genres, while the direct group showed significant improvement in the narrative essay. It is interesting, though difficult to explain, that in the narrative genre students in the control group showed near significant improvement in the wrong word category, while in the descriptive essay they got very significantly worse. The narrative genre finding was similar to that of Ferris and Roberts (2001), where students who received no feedback improved in correcting wrong word errors compared to article, sentence structure, verb and noun errors. Students seem to edit wrong word errors without feedback on their own.

Spelling was the only error category where there were significant or nearly significant results for all the experimental groups in both genres, while the control groups in both genres did not improve in spelling errors at all. Teacher correction seems to have helped the students to notice some of their spelling errors and improve spelling accuracy over time concurring with Lee (1997) where spelling errors were easily noticed compared to other error categories. This result however contrasts with that of Fazio (2001), where students who received error correction had their spelling error ratios remaining the same while those receiving commentaries or both corrections and commentaries had their spelling ratios get worse at the end of the study.

5.2 Limitations

This study was a mainly quantitative one, and complementary qualitative exploration was limited, though it did include a teacher interview and a more qualitative follow-up on some individual changes in spelling that was carried out on the most improved students in this category. If I had not had to return to Kenya at the end of the term after I collected my data, I could have been in a position to interview selected students so as to gain further qualitative perspectives on their error correcting and language learning strategies, though this is not the main focus of my research.

Similarly to many related studies, the present one focused on a limited number of error types. Thus it does not provide a comprehensive account of errors in the scripts, though it does cover the majority. This also means that the students did not receive feedback on all their errors, though it is not likely that this would have had major effects on their writing progress or their attitudes to error correction. Also, the error categories needed to be quite broad, and so finer analysis of subcategories of verb errors, for example, was beyond the scope of the study.

This study was carried out within one school term, which was effectively the relatively short period of nine weeks. Ferris (2004) investigated the period within which most studies have been carried out and she found that the period varied from 10 weeks to nine months, while quite a number of studies were not longitudinal at all. The period of this study could well not have been long enough to allow the students time to improve on the error categories that were studied and a study carried out over a longer period of time could have given a more conclusive result.

During the experiment, it was difficult to have all the students write all eight essays, and it could be argued that this group was made up only of its relatively more committed students, as some of them were absent at one time or the other and this could have affected the means of some of the groups, especially the control group, where only about half of the students completed all the eight essays, and it could be argued that this group was made up only of its relatively more committed students, so influencing the overall results.

Another limitation of the study was that the students were not used to rewriting their essays in their usual class situations with their teacher and so this could have negatively affected their correction

success. In addition, the students had more writing sessions compared with what they normally had and also they had not had any follow up of rewriting their work with their teacher. This could have led to students becoming overwhelmed by the whole exercise, as perhaps suggested by the results showing that the students wrote shorter texts in the posttests compared to the pretest essays.

The number of students in the experimental groups was quite small and so the results of this study may not be representative of what could happen if a larger number of students were used in a similar study. The subjects were also of a relatively low academic proficiency, thus making it difficult to generalise the results to, for example, other learners at this level in Botswana schools.

5.3 Implications and suggestions for further research

Findings of this study are quite relevant with regard to teachers of English who have been giving their students feedback without asking them to carry out any form of revision, such as rewriting their work. The generally better improvements in accuracy of the experimental groups over the control group means that the study can be interpreted as implying that to some extent at least, revision of composition writing leads to improvement in students' accuracy. This was shown in terms also of findings for Hypothesis 2, where all the experimental group students managed to correct most of the errors they made after teacher correction had been done. As Ferris (2004, 54) argues "the cognitive investment of editing one's text after receiving error feedback is likely a necessary, or at least helpful, step on the road to a longer term improvement in accuracy". Giving feedback without following up with revision is, according to Lalande (1982), as good as giving none.

The study also revealed that the direct feedback mechanism led to a significantly higher percentage of error correction compared to the other mechanisms and this might be interpreted to mean that direct feedback should be the preferred method in the classroom. However, findings in respect of accuracy development over time point to the superiority of coded feedback, probably because of the deeper "cognitive investment" it requires. More research should be undertaken in this connection, but the implication is that teachers should prefer this type of indirect mechanism.

Teachers of English will also realise that students' responses to errors differ from one category to another. The study found that some error categories were improved on greatly, such as spellings and to

some extent wrong word choice, while students made more errors in others such as the verb category, so concurring with Lee (1997). Teachers of English therefore need to treat certain errors cautiously as some errors need more attention than others.

Students expressed their need for teacher correction feedback. Even though Truscott (1996) feels that students do not always have to be given what they want, students have shown that they also need it in order to notice errors and so have a base for improving accuracy.

With regard to future research, a similar study could be carried out to cover a longer period of time and not just three months, as was the case in this study. Similarly, given that the number of students in this study was quite small, a future study involving a larger number of students should provide a clearer picture of the value of error correction feedback.

More studies should be undertaken to find out what exactly leads students to improve in their writing fluency, as this study did not discover any relationship between giving teacher correction feedback and fluency, contrary to Chandler (2003) and Fathman and Whalley (1990), where fluency in student writing improved.

This study only investigated six error categories and more studies in the future should investigate others, while there is also a need to look at more narrowly defined subtypes of error categories, particularly in studies involving large numbers of students, where substantial number of errors within subtypes could be compared statistically.

Finally, most error correction feedback studies have been predominantly quantitative, as the present one is, but there is a need in future studies to make better use of a combination of quantitative and qualitative methods in order to arrive at better understanding of the role of error correction feedback in the development of writing accuracy by individuals over time.

5.4 Conclusion

As a conclusion to the study, I would like to refer to the words that helped stimulate it:

> … we need to think of ways to carry out longitudinal, carefully, designed, replicable studies that compare the writing of students receiving error feedback with that of students who receive none, as well as comparing and controlling for other aspects of error treatment. (Ferris 2004, 60)

I hope that my study, in spite of its limitations, has made some contribution in this direction.

REFERENCES

Anusiene, L., Kaminskiene, L. and Kavaliauskiene, G. 2009. Classroom practice: Error correction at tertiary level. *Studies about Languages (Kalbu Studijes)* 14, 65-72 on www.ceeeol.com. Accessed: 17/8/2009

Arthur, J. 1997. There must be something undiscovered which prevents us from doing our work well: Botswana primary teachers' views on Educational Language Policy. *Language and Education* 11, No.4, 1-2.

Ashwell, T. 2000. Patterns of teachers' response to student writing in a multiple-draft composition classroom: Is content feedback followed by form feedback the best method? *Journal of Second Language Writing* 9, 227-258.

Bitchener, J., Young, S., and Cameron, D. 2005. The effect of different types of corrective feedback on ESL student writing. *Journal of Second Language Writing* 14 (3), 191-205.

Botswana Languages. www.botswanalanguages,com Accessed: 17-04-2007

Brown, H. D. 1994. *Principles of language learning and teaching*. Englewood Cliffs, NJ: Prentice-Hall.

Brown, H. D. 1991. *Breaking the language barrier*. Yarmouth, ME: Intercultural Press.

Chandler, J. 2003. The efficacy of various kinds of error feedback for improvement in accuracy and fluency in 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.

Chaudron, C. 1998. *Second language classrooms*: *Research on teaching and learning*. Cambridge: Cambridge University Press.

Cumming, A. and So, S. 1996. Tutoring second language text revision: Does the approach to instruction or the language of communication make a difference? *Journal of Second Language Writing* 5 (3), 197-226.

Delgado, R. 2007. Effects of different error feedback: Approaches in students' ability to self-edit their writing. *Divergencias. Revista de Estudios Linguisticos y Literarios*. Volumen 5 Numero 2, 1-16.

Fathman, A. K. and Whalley, E. 1990. Teacher response to student writing: Focus on form versus focus on content. In B. Kroll (ed.). *Second Language Writing: Research insights for the classroom*. Cambridge: Cambridge University Press, 178-190.

Fazio, L. 2001. The effect of correction and commentaries on the journal writing accuracy of minority- and majority- language students. *Journal of Second Language Writing* 10, 235-249.

Ferris, D. R. 1997. The influence of teacher commentary on student revision. *TESOL Quarterly* 31, 315-339.

Ferris, D. R. 2000. Chaney,S. J., and Mackee, S. Perspectives, problems, and practices in treating written error. Colloquium presented at TESOL Convention, Vancouver, BC.

Ferris, D. R. 2002. *Treatment of error in second language writing classes*. Ann Arbor: University of Michigan Press.

Ferris, D. R. 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. R. and Roberts, B. 2001. Error feedback in L2 writing classes. How explicit does it need to be? *Journal of Second Language Writing* 10, 161-184.

Frankfort-Nachmias, C. and Nachmias, D. 2004. *Research methods in the social sciences*. London: Arnold.

Frantzen, D. 1995. The effects of grammar supplementation on written accuracy in an intermediate Spanish content course. *Modern Language Journal* 79, 329-344.

Greenslade, T. A. and Felix-Brasdefer, J. C. 2006. Error correction and learner perceptions in L_2 Spanish writing. In C. A. Klee and T. L. Face (eds) *Selected proceedings of the 7th conference on the acquisition of Spanish and Portuguese as first and foreign languages.* Somerville, MA: Cascadilla Proceedings Project, 185-195.

Guenette, D. 2007. Research design issues in studies of feedback on writing. *Journal of Second Language Writing* 16, 40-53.

Havranek, G. 2002. When is corrective feedback most likely to succeed? *International Journal of Education Research* 37, 255-270.

Hedgcock, J. and Lefkowitz, N. 1994. Feedback on feedback: Assessing learner receptivity to teacher response in L2 composing. *Journal of Second Language Writing* 3, 141-163.

Hedgcock, J. and Lefkowitz, N. 1996. Some input on input: Two analyses of student response to students' feedback in L_2 writing. *Modern Language Journal* 80, 287-308.

Hendrickson, J. M. 1978. Error correction in foreign language teaching: Recent theory, research, and practice. *Modern Language Journal* 62, 379-398.

Hendrickson, J. M. 1980. The treatment of error in written work. *Modern Language Journal* 64, 216-221.

Hong, Y. 2004. *The effect of teachers' error feedback on international students' self-correction ability*. M.A. thesis, Brigham Young University, Provo, Utah.

Horner, D. 1988. Classroom correction: Is it correct? *System* 16, 213-220.

Hule, M. R. and Mooko, T. 2002. An examination of teacher's feedback on students' writing in Agriculture, English and Social Studies in selected Botswana junior secondary schools. www.bera.co.bw/attatchments/mosenodi. Accessed: 24/02/2009.

IELT-International English Language Testing system. www.ielts.org Accessed: 17-04-2007

Kepner, C. G. 1991. An experiment in the relationship of types of written feedback to the development of second language writing skills. *Modern Language Journal* 75, 305-313.

Kombo, D. K. and Tromp, D. L. A. 2006. *Proposal and thesis writing*. Nairobi: Don Bosco Printing Press.

Lalande, J. F. 1982. Reducing composition errors: An experiment. *Modern Language Journal* 66, 140-149.

Lee, I. 1997. ESL learner's performance in error correction in writing: Some implications for college level teaching. *System* 25, 465-477.

Lee, I. 2004. Error correction in L2 secondary writing classrooms: The case of Hong Kong. *Department of Education studies, Hong Kong Baptist University*. Hong Kong, China.

Liu, Y. 2008. The effects of error correction feedback in second language writing. *Arizona Working Papers in SLA and Teaching* 15, 65-79.

Ngoh, G. L. and Fereira, D. P. 1980. *Understand and communicate 4*: *An English course for secondary schools*. FEB International (Pak) Limited.

Nunan, D. 1992. *Research methods in language learning.* Cambridge: Cambridge University Press.

Nyati–Ramohobo, L. 1991. Language development for literacy. The case of Shiyeju in Botswana. http://literacyonline.org/products/ili/webdocs/ramahobo.html. Accessed: 2/6/2007.

Pica, T. 1994. Questions from the language classroom: Research perspectives. *TESOL Quarterly* 28, 49-79.

Polio, C. G. 1997. Measures of linguistic accuracy in second language research. *Journal of Second Language Writing* 47, 101-143.

Polio, C., Fleck, N. and Leder, N. 1998. "If only I had more time": ESL learners' change in linguistic accuracy on essay revisions. *Journal of Second Language Writing* 7, 43-68.

Richards, J. C., Platt, J. and Platt, H. 1992. *Dictionary of language teaching and applied linguistics*. Malaysia. Longman Group UK Limited.

Robb, T., Ross, S. and Shortreed, I. 1986. Salience of feedback on error and its effect on EFL writing quality. *TESOL Quarterly* 20, 83-93.

Schlig, C. 2005. Improving foreign language writing competence. Georgia State University. http://education.gsu.edu/ctl/flu/schlig-FL writing-revising.htm Accessed: 03/09/2006.

Semke, H. D. 1984. Effects of the red pen. *Foreign Language Annals* 17, 195-202.

Sheppard, K. 1992. Two feedback types: Do they make a difference? *RELC Journal* 23, 103-110.

Truscott, J. 1996. The case against grammar correction in L2 writing classes. *Language Learning* 46, 327-69.

Truscott, J. 1999. The case for "The case against grammar correction in L2 writing classes". A response to Ferris. *Journal of Second Language Writing* 8, 111-122.

UNESDOC Database United Nations Educational, Scientific and cultural organisation. http://unescodoc.unesco.org/images/0014/001422/14225e.pdf Accessed: 17-04-2007

University of Botswana. 2006. *General information on general regulations for Bachelor's Degree programme*. Gaborone: University of Botswana.

VassarStats. Vassar College Public Access Employee Directory. http://faculty.vassar.edu/lowry/vassarStats.html Accessed: 10-01-2009

Wolfe-Quintero, K., Ingaki, S., and Kim, H. Y.1998. Second Language Development in Writing: Measures of fluency, accuracy, and complexity (Tech. Rep. No. 17). Honolulu: National Foreign Language Resource Centre.

APPENDICES

Appendix 1

Topics given for the writing sessions

Time 1 hour 20 minutes

Instructions: In about 300 words write a composition on the following topic. Indicate the number of words at the end of your essay.

Descriptive writing (the four compositions were sequenced as below)

1. Describe your room after a party is over and the guests have left (pretest).
2. Describe your school.
3. Describe a bus stop in town during the rush hour.
4. Describe a natural disaster you have witnessed (posttest).

Narrative writing (the four compositions were sequenced as below)

1. The most interesting holiday I have ever had (pretest).
2. Describe a day that you felt disappointed.
3. A bus is going round a bend, suddenly it skids off the road. Relate what happens.
4. Write a short story on the following topic. A midnight visitor (posttest).

Appendix 2

Instructions for rewriting sessions

Corrections were done within 1 lesson which was 40 minutes long. This was done after students received teacher correction feedback that had been done by the researcher and the scripts given back to the students via the participating teacher. The following were the instructions that were given to the students for each feedback group.

1. Direct feedback

Please go through your essay carefully again. I have given the correct forms of the errors that you made. Then rewrite your essay.

2. Coded feedback

Please go through your essay carefully again. I have marked the errors that you made with codes, explained below. Then rewrite your essay.

ww – wrong word error

p – punctuation error

V – verb error

prep – preposition error

sp – spelling error

art – article error

3. Uncoded Feedback

Please go through your essay carefully again. I have underlined the errors you made. Then rewrite your essay.

4. No feedback

Students were not asked to rewrite their essays.

Appendix 3

Students' questionnaire

Instructions: This questionnaire aims at finding how you feel about the kind of corrections made in your compositions. Please answer them as sincerely as possible. Do not write your name on the paper. Your response will be treated with confidentiality.

1. Was it easy for you to see the kind of mistakes that you made?
   (a) Always (b) Most of the time (c) Sometimes (d) Rarely (e) Never

2. Did correction by the teacher help you to correct the errors that you made?
   (a) Always (b) Most of the time (c) Sometimes (d) Rarely (e) Never

3. Did you understand the kind of correction that the teacher gave you?
   (a) Always (b) Most of the time (c) Sometimes (d) Rarely (e) Never

4. Did you feel discouraged by the way the teacher corrected your errors?
   (a) Always (b) Most of the time (c) Sometimes (d) Rarely (e) Never

5. How often would you like the teacher to correct your errors?
   (a) Always (b) Most of the time (c) Sometimes (d) Rarely (e) Never

Appendix 4

Teacher interview

The researcher explained to the teacher the purpose of the interview. The researcher then sought permission from the participating teacher to use whatever information that was elicited from the interview for the purposes of a dissertation. The researcher made use of open ended questions. The researcher took notes as the interview took place.

Part 1

Personal Details

1. Name of teacher, sex and age
2. Highest level of education
3. Professional qualification
4. Teaching experience
5. Any formal training in giving feedback in composition

Part 2

1. What type of writing do you give your students at this level?
2. How often do you give composition writing in a term.?
3. What type of error correction do you give your students?
4. Of these feedback mechanisms which do you use often?
5. What form of revision do your students carry out after you have given feedback?
6. Do you think that revision is necessary in student writing?
7. What follow-up measures do you take to ensure that students attend to the feedback that you give them?
8. What is your opinion about the effects of different feedback mechanisms on student writing?

Appendix 5

A sample of students' scripts and error correction feedback used in the study

A sample of essays of the most improved student in the spelling error category was selected from each feedback group. This sample illustrates the feedback in respect of **all (and only) the six error types analysed**, but because of the focus on spelling errors in the discussion in Chapter 4, they are presented in bold print. The following students were selected: D8 from the direct feedback group, C9 from the coded feedback group and UN6 from the uncoded feedback group.

DIRECT FEEDBACK

Teachers' remarks (at the end of all direct feedback essays)

Please go through your essay carefully again. I have given the correct forms of the errors that you made. Then rewrite your essay.

Pretest narrative essay (direct feedback)

D8

THE MOST INTERESTING HOLIDAY I HAVE EVER HAD

It was on the 16th of June 2003 when we were celebrating [the (African child) African Child] day with my friends and people in the village.

It was in the morning, and those who were going to participate at the community hall were preparing [**themselfs**themselves]. I was taking part in drama as an old man and my school mates liked to watch when it was my chance to act, while we [arewere] [**practicing** practising].

It was my first time to take part and be watched by hundreds of people [,˙] I was so scared but my coach told me just to [**relux**relax]. Some other groups were doing their work and when they came out they told us that it [iswas] [scaringscary] when you see those eyes looking at you. The coach told me to look on [onat] the top of their heads and not [thetheir] eyes for me to do my part wonderfully [, ˙] [immediatelyImmediately]our group was called to get inside the hall and take our chances.

We entered the hall singing loudly and there were a lot of people but they [**well come**welcomed] us with a hand. We started the action immediately after the song. I was the first person to take part and viewers were [**quete**quiet][ˌ] waiting for me to [deliverydeliver] the [**massage**message] to them and I was so scared [bybut] I [holdheld] myself until I finished the [**massage**message] and they gave me a hand, showing that my part was perfectly done. At the end of the action people were amazed [ˌ] telling me that I was so confident and I was given the [presentprize] [offor] the best actor of the entire group that were taking part. I was so happy and this holiday was the most interesting to me.

Posttest narrative essay (direct feedback)

D8

A MIDNIGHT VISITOR

It was on the 25th of December 2000 at night. The night was cool and [**quete**quiet] [ˌ] the sound of owl far away.

My sister went to Gantsi to bought [boughtbuy] me some clothes for [new yearNew Year], so I was sleeping alone in the house. While sleepingˌ I dreamed awakened by [thea] noise at the window, when I woke up I saw [thea] shadow [inon] the window and it moved towards the door and I opened it. I sat up straight on

[athe] bed and waited for that thing to switch the lights on [, ˙] [immediatelyImmediately] it did so and I was so scared and shocked.

This thing looked like a person [ˌ] but short [ˌ] with a lot of hair around it and a big head. It looked like a female person moving towards me and I was trying to [**scrim**scream] but the sound was not there. It touched me with its rough hands and pulled off my pants and immediately I collapsed.

While still [onin] bed, my friends came and [wakewoke] me up. I was like a running person, I was so wet all over the body and they laughed at me [,] asking me what was going on [,] I told them that I was having a scary dream. They said ["] we have visitors you [**now**know] [ˌ] [letslet's] have some fun ["] [, ˙] I was confused, by this time 12:00 midnight having fun.

When I [go[went]] outside the house and there [was[were]] a lot of beers and drinks. ["]Don't forget that it is Christmas today ["], they [say[said]] to me. We started drinking all [**this**[these]] drinks and beers until [in [delete]] the morning. I said to [**my self**[myself]], sometimes midnight visitors can help if you are attacked in the house as you can see that they helped me from that scary and terrible dream.

Pretest descriptive essay (direct feedback)

D8

DESCRIBE YOUR ROOM AFTER A PARTY IS OVER AND THE GUESTS HAVE LEFT

The room was like where the dogs sleep, there were things all over the room and the guests just left and I had to make sure that the room [is[was]] clean.

I was so tired because I [have[had]] been busy since the party started. Some of my friends came over and they took me to their party, I [**tryed**[tried]] to explain that I was just from the party but they didn't [**lessen**[listen]] to me, so I [have[had]] to [**live** [**leave**]] my room untidy.

When I came back it was just to shoot straight to the bed [,] no more to think about the room how it looks and I fell on the rubbish and I [**feel**[fell]] asleep. When I [wake[woke]] up it was ten on Sunday morning and I [have[had]] to go to church and the room was just like [yesterday[the previous day]]. I [**desided**[decided]] to go to church [.] I washed my clothes and cooked.

I was still feeling like going to bed and after finishing the clothes and cooking I [go[went]] to sleep again and when I [wake[woke]] up it was at six o'clock [[on]] Monday morning and I [have[had]] to go to work but the room was just the same and I left it like that and [go[went]] to work

I will never [**plane**[plan]] to [do[have]] parties in my room because my room was dirty and it is still dirty by [by[up to]] now.

Posttest descriptive essay (direct feedback)

D8

DESCRIBE A NATURAL DISASTER YOU HAVE WITNESSED

It was on the first day of October when I was preparing myself to go to the cattle post and there [was[were]] some dark clouds in the sky.

When I was about to leave, the wind started to blow hard and the clouds were producing a [scaring [scary]] sound, and the lightning [[was]] [**stricking**[striking]**]** everywhere. My parents told me not to go but I said ["] [that [that's]] nothing [,] I will just walk to the cattle post ["].

I told them that I [am[was]] going and they said "be careful." When I was about to reach the cattle post a heavy rain started [,] with a powerful wind blowing from all directions. I tried to run but it was difficult. The wind was so powerful and it [break[broke]] trees down and there was no need for me to go under the trees. It was so dark but I [**prefered**[preferred]] to walk and I took a different direction and I did not realize that I took [a[the]] wrong direction.

The rain stopped and I found myself where we used to collect some firewood and it was about four kilometers away from our cattle post. When I realized, this I was so angry that if I [[had]] [**lestened**[listened]] [[to]] my parents I would not have been [here[there]]. The trees on the way were lying down, broken by a strong wind and at the cattle post the house was taken by the wind and there was nothing left.

CODED FEEDBACK

Teacher's remarks (at the end of all coded feedback essays):

Please go through your essay carefully again. I have marked the errors that you made with codes explained below. Then rewrite the essay.

ww – wrong word error

p – punctuation error

V – verb error

prep – preposition error

sp – spelling error

art – article error

Pretest narrative essay (coded feedback)

C9

THE MOST INTERESTING HOLIDAY I HAVE EVER HAD**.**

It was sunset when I was at the lands on the exciting day of my life. It was on [art] [eighteenww] of April 2006. I was with my father, mother, little brother and little sister.

Halfway before we arrive [arrivev] I had bad dreams and I thought they [willv] come therefore the next day. The holiday was so dramatic in the afternoon but during the day was [**draggy**sp]. That means that my dreams [wasv] just [**fantisies**sp].

I was [**laying**sp] on my back on top of my bed when I [hearv] some sweet melodies outside. I went out exhausted, [frighteningww] and with anger [toprep] [**my self**sp] I stood outside in the middle of nowhere to go in the opening door. I looked up [inprep] the sky [p] it was dark, then I looked [onprep] [art] top of trees. I found branches, trees and something else. It was difficult to recognize what it was because it was dark [p] I decided to go closer. When I [comev] closer and closer, I also [becomeww,v] shivering.

The pitch of the sound also [becomev] loud. That sound was nice that [makesv] me [to havev] [**this**sp] ideas of finding what was that which [makesv] sweet voices in the forest. I didn't [madev] any sound of my [**foot steps**sp]. When I reached the [**tlee**sp] I looked [p] searched and searched but didn't [foundv] [**any thing**sp]. I saw some [**signes**sp] of leaves and branches. They were moving but there was no blowing wind.

I stood [**steel**sp] and [shoutv] out loud [p] said [p] one is here [P] The response [, p] I found many [of prep] birds [**fling**sp] away. They made a loud sound that made me run away from them too. [Inprep] [myww] halfway to the house I stopped [, p] I laughed at myself. I said "[p] am I stupid [p] how I can run away from pony birds [.p] [p] That cannot even hurt me".

During the next day when I woke up, I thought it was a dream. I looked at my shoes and I found them with someone soil. There were my footprints on the floor. I said to myself, "that was [anart] interesting holiday I [hadv] ever had".

Posttest narrative essay (coded feedback)

C9

A MIDNIGHT VISITOR

The sun [was[vj]] already set and it was dark. The stars were [**spacking**[sp]] like diamonds in the sky. The night was [on[prep]] ninth of March 2005 and it was so dramatic that made it exciting.

Before I went to sleep, I had some premonitions about somebody's life in danger. I started to become [shuddered[ww]] [[p]] wondering what [is[v]] happening. I muttered to myself and said, "What is this all about?" I did not get an answer. There was no one to ask because [there[ww]] were all asleep.

I went to my bedroom and started to [**lay**[sp]] on my back on top of my bed. All [**this**[sp]] ideas of sleeping [was[v]] to relax and rest my mind. I heard a voice of someone calling for help. I [ignore[v]] [they[ww]] voice but the more I [ignore[v]] it, the [more[ww]] the pitch of the sound became. I woke up and stood still. The sound disappeared and I asked myself questions like [[p]] who is that[[p]] what is happening? [[p]]

I went back to bed and it started again. I listened carefully and it [seems[v]] that [is[v]] the voice of a man. I thought is those people who [likes[v]] to go to pubs and [there[ww]] are going back to their cribs. Some thoughts were that they [are[v]] just drunk and they [are[v]] playing. The voice became louder and nearer to my house. I tried to look through the window[[p]] [[p]] because it was so dark outside I did not find anything or any sign. I just [wait[v]] and [wait[v]].

While I was waiting, I heard someone knocking at the door. I went to the door and asked who [is[v]] this knocking at this time. The visitor just said, "Open the door, I need help from you, someone is wanting to kill me".

I opened the door slowly and I found that it [is[v]] my cousin. I became [dramatically[ww]] and excited.

Descriptive pretest essay (coded feedback)

C9

DESCRIBE YOUR ROOM AFTER A PARTY IS OVER AND THE GUESTS HAVE LEFT

It was a bright, [glory[w]] and lovely day. It was the first Sunday of January 2004. We had a family party at home. We were celebrating for the New Year we [are[v]] still together and happy.

After the party [close[v, w]] I decided not to sleep at home because there [was[v]] some guest that [come[v]] from far places. They needed [where[ww]] to rest until the next day. They slept in my room and I slept at my friend's house.

In the morning I decided to go back home to eat some [**left overs**[sp]] because I was very hungry. I wanted to go and clean outside where the party was held. After cleaning I went inside, when I opened the sitting room door I felt excited and I said how could they leave the house [**daddy**[sp]] like this? I started cleaning without [**any one**[sp]] to come and help me because they were all sleepy tired.

I even [picture[v, ww]] about my own bedroom. I thought that they would respect it but I was wrong. They were [**laying**[sp]] on the floor [putting[ww]] under my blanket and some on top. Some of them were vomiting on them.

They were drunk. My furniture [were[v]] broken [[p]] [like[ww]] my mirror [were[v]] cracked, the door of my [**wall drop**[sp]] [were[v]] removed. It looked like [there[ww]] [**we**[sp]] jumping on my bed and [**injoing**[sp]] with little ones.

I didn't say even a word to them. I murmured and said, "I did not think an elder [can[v]] do a mess like a kid or animals". I waited and waited until they woke up. When they woke up I found them shivering and they came to me and [**apologies**[v, sp]] I [**excepted**[sp ww]] their kindness and they promised to repair their damage.

They kept their promise. After a week it was repaired. I did not [forgot[v]] that bad day of my life. Now I am [**leaving**[sp]] in a pretty room like I used to. Only those who forgive can be forgiven.

Posttest descriptive essay (coded feedback)

C9

DESCRIBE A NATURAL DISASTER YOU HAVE WITNESSED

The night was dark that made difficult to recognize anything. It was the second day of March 2004. It was the [worse[ww]] night for my neighbour.

Our church organized a trip to Kyalagadi. We decided to camp in the forest because we were many. To prevent [to be[v]] overcrowded we decided to make some tent, for everyone to live with his/her family

When it was time to sleep, to be prepared to go back the next day because we camped for five days. Me and my church mates, we planned to sleep outside [in[prep]] the fire. We were singing, dancing because we were enjoying the warmth of the fire. When times [goes[v]] on we felt tired and we slept [in[prep]] the fire.

I [failed[ww]] [of sleep[ww]]. The moment I woke up, I found that there [is[v]] a large light. I just thought it was from the fire we were [in[prep]]. I woke up and looked everywhere. The forest was burning [[p]] the fire [were[v]] getting closer and closer. It was coming from my neighbours [[p]] tent and it was getting close to it. I [**scrimmed**[sp]] and said, [[p]] run, run for your life. There is fire coming to our side [[p]]. It was too late for our neighbour.

The time they tried to get out, the fire was already surrounding the tent. They called for [a[art]] help. We tried and tried [[p]] we making it worse. A person can have many ideas at a [**perticular**[sp]] time. We decided to use branches. It [stops[v]].

It is good to fight fire with tree branches because water makes it worse than before.

UNCODED FEEDBACK

Teacher's remarks (at the end of all uncoded essays):

Please go through your essay carefully again. I have underlined the errors you made. Then rewrite your essay.

Pretest narrative essay (uncoded feedback)

UN6

THE MOST INTERESTING HOLIDAY I HAVE EVER HAD.

Wow!! It was on the nineteen of December 2006 when I had a very interesting holiday and enjoyable one, with my family at Cape Town.

On the fourth of December my mother told me that we will be touring as a family. It was a happy mood on my **sight** to hear this good news. She gave me a paper in which **they where** names of hotels at cape town and she told me to choose, these hotels were displayed in the paper they were all **beatiful** and I did not know which one to choose but at the final I chose _ Cape town hotel_ it was the most **behatyful** when **u** look clearly. I keep on pray so that days can pass fast so that I can go_ I do not know.

As if God or what was with me and it was on the eighteen when my mum told me that I should pack everything in the bag as we are going to leave with the aeroplane which departure very early. In that night I did not sleep_ it was just to rest_ as I asked **my self** when is the sun going to rise but it was not six **oclock**. When we have finished to do everything and we went to Sir Seretse Khama Airport (SSKA). We stayed about six minutes and we were told to get into the aeroplane and put the belt on ourselves and it departure. As it was my first time to be in an aeroplane i vomited and I was given a red bucket to **vomite** inside it.

It was ten **oclock** when we arrived in Cape town airport and we took a taxi to cape town hotel. It was so **beatiful** that I thought I was in heaven__ it was a green pasture. At 3 **oclock** mummy and daddy told me that we would be leaving and going to the game reserves and parks_ **their** I saws kangaroos, elephant**,** lions and many more_ after that we went to the sea and I was playing in water while mother and father **where** resting. It was a daily exercise after eating__ we went to the **sea side** playing with sand and going shopping.

On the twenty-seven my mother told me that on the twenty eight we will be leaving. I was so bored. I was bored at the end because I was coming back to Botswana, but I enjoyed Cape town more than the word.

Posttest narrative essay (uncoded feedback)

UN6

THE MIDNIGHT VISITOR

What a shame!! Happened to me that day and I will never forget it in my life. It was on the first of July 2006.

Just like other days I told my mother and daddy that I would be attending a friend party on Saturday. they agreed and gave me money to buy a present on Friday. On Saturday I wake up in the morning wash my laundry and cleaned my bedroom, after I have finished to do everything I went straight to the bathroom and bathed. The party was starting at two **oclock**, at exactly twelve **oclock** noon I finished to do everything and went to say goodbye to my family_ before I even told them that I am leaving my father told me that by eleven **oclock** I should be home. I said goodbye to everybody and went to the bus stop, the party was at extension two.

When I got at the party it was so good. We have got a braai, drinks, alcohol, beer and salads and about **entairtnment** I do not need to say much_ it was cool. I even forgot at eleven I should be home, the party finished at ten **oclock** instead of going home I went to calabash **nite** club we got **their** it was nice. It was twelve **oclock** midnight when I departed to go home_ when I got home the gate was locked. I **tryed** to call mum but her cell phone was off_ I did not want to call dad because I know that he cannot open up for me_ just because I was afraid waiting at the gate alone I end up calling him, he did not answer the phone. I did not know if he was asleep or he did not want to answer the phone. I ended up give up and take a walk to my friend's place at Tsholofelo, from Maruapula to Tsholofelo is along way_ I was afraid but **they** was no way to go.

We shared the gate and I told her my problem_ instead of her admitting my problem she said to me "How can you ask a toad to give you a chair while you see it squatting"_ She told me that I know very well that her bedroom is small and she is sharing it with her younger sister,_ I did not know what to do. I said goodbye to her and when I got near the gate, she called me and told me that she will ask her mother to sleep with her younger sister. She went to her mother's room and knocked and she opened up and she told her everything and I could hear her mother's voice "what a midnight visitor_" she

cooled her down and she agreed. I went straight to bed and unwear shoes only then went to bed. It was a boring weekend I have never experiment.

Pretest descriptive essay (uncoded feedback)

UN6

DESCRIBE YOUR ROOM AFTER THE PARTY IS OVER AND THE GUESTS HAVE LEFT

What A_ mess!_ happened in my room, it was on the twenty-two of November when we had a party at home,_ it was my mother's party graduating for being a mother for about twenty-eight years.

The party started on Friday at around 5 **oclock** at **they** were many people at home. My mother's relatives **where** **their** and father's relatives. Indeed it was a very good party I have never attended. My mother was so beautiful_ she was wearing a two piece of pink trouser well designed and a pink top, at the head I don't need much to say because she was_ beautiful the person who did her did it very slow she/he did not rush, even father was wearing_ very smart black suit and that day he was wearing a beautiful black hat which match with the clothes_ he was so handsome. On Friday it was just to make things in order so that on Saturday **they** will be no time to start preparing salads. I went to bed at twelve **oclock** after I had finished to do everything.

On Saturday the party started at two **oclock** in the afternoon_ it was good and I was very happy. People were coming with presents and they were unwrapped in my room, elder people were ululating_ others dancing and my father's brother made a poem and perform it during the party.

The party finished up at five **clock** and guests left to their places_ when I came to my room, at first when I opened the door, I found shoes on **their** way out laundry on top of the bed everything fighting the party started_ it was like I can cry. I called my mum to saw what happened and she told me to go and sleep at the other room she will sweep it tomorrow. Happiness covered my face but not that much.

Posttest descriptive essay (uncoded feedback)

UN6

DESCRIBE A NATURAL DISASTER YOU HAVE WITNESSED

What a shame!! I could not believe what my eyes had seen. It was on the twenty-fourth of August 2004 when the thunder burn my neighbour's house.

It was on the afternoon when clouds gather they **where** very black, they **where** from the north going to the West they just gather fast and the earth was covered with black clouds. Anyway the weather was cool and I even enjoyed it with a cousin when this terrible thing happen when were sitting on the veranda chatting and laughing the news **where** good. My neighbours does not like rainfall so when it its rainy she call us to come and stay with her or she come to my place but that day we went to her home. We were in the sitting room watching television.

The program that was on television was so nice so we kept silent at watch my African dream. I remember the television switch off and all the necessary light switch off and I heard the loudest noise and the house was on fire. After that I do not remember anything. I just remember **my self** at the hospital and mummy and daddy **where** on my side I asked them why I am in the hospital they could not answer they looked each other and my mother was crying. I asked **my self** what could had happened but **they** were no answers.

I remember being taken by the police to a secret room asking me what had happened and I told them that I saw the television switch off and all necessary lights switch off and I last I saw **my self** at the hospital. They told me that the house had been burnt by the thunder and my neighbour is no longer alive and my cousin is better even if she had scars on her face. I was the only one who was not burnt I was admitted for being afraid very much.

I was back home and family members **where** happy but not that much because of what happened to our neighbour and even my cousin. I said "Thanks god I am alive."