TURKISH REPUBLIC
TRAKYA UNIVERSITY
INSTITUTE OF SOCIAL SCIENCES
FOREIGN LANGUAGES TEACHING DEPARTMENT
DIVISION OF ENGLISH LANGUAGE TEACHING
A MASTER'S THESIS

THE EFFECTS OF WRITTEN CORRECTIVE FEEDBACK ON L2 WRITING SKILL AND GRAMMATICAL ACCURACY

SİBEL ÖZÇELİK

ADVISOR
Prof. Dr. MUHLİSE COŞGUN ÖGEYİK

EDİRNE 2022

ACKNOWLEDGMENTS

First of all, I would like to express my deepest appreciation to my advisor, Prof. Dr. Muhlise Coşgun Ögeyik for her invaluable contributions and continuous support in completing this thesis. I would not have completed this thesis without her guidance.

Besides my advisor, I would like to thank Assoc. Prof. Dr. Kutay Uzun. I am grateful to him for his contributions in the process of data analysis. Also, I would like to thank Assoc. Prof. Dr. Emre Güvendir for his valuable suggestions. I am really appreciative of his efforts to teach me how to conduct research.

I would like to thank my students who voluntarily participated in the research. They have never lost their enthusiasm to learn and do their best. Also, I would like to thank my dear colleagues and friends for their continuous support. They have never stopped encouraging me to keep going on.

I am indebted to each of my family members for their love and encouragement. I would like to thank my beloved husband, İsmail Özçelik. He has always been there to give me strength and uplift me in difficult times.

Last but not least, this thesis is dedicated to all my dear teachers who have helped me cultivate a love of learning.

Başlık: Yazılı Geribildirimin İkinci Dilde Yazma Becerisi ve Dilbilgisel Doğruluk Üzerine Etkileri

Yazar: Sibel ÖZÇELİK

ÖZET

Bu çalışma dolaylı yazılı geribildirimin genel yazma performansını iyileştirmedeki etkinliğini araştırmıştır. Bu çalışmaya yabancı dil olarak İngilizce eğitimi gören iki mevcut sınıf katılmıştır. Katılımcılar, Beykent Üniversitesi Yabancı Diller Yüksekokulu Genel İngilizce hazırlık sınıflarında kayıtlı toplam 36 Türk üniversite öğrencisinden oluşmaktadır. Deney grubu hata düzeltme kodları aracılığıyla dolaylı yazılı geribildirim alırken, kontrol grubu ise yazılı geribildirim vermenin yaygın ve geleneksel bir yolu olan doğrudan yazılı geribildirim almıştır. Her bir sınıf yedi farklı paragraf üretmiş ve onları yedi haftalık bir süre boyunca revize etmiştir. 1. Haftada, her iki katılımcı gruba da bir ön test uygulanmış ve sonraki beş hafta, deney grubuna hata düzeltme kodlarını kullanmak üzere eğitim verilmiştir. 6. Haftada, katılımcılara doğrudan veya dolaylı düzeltmeler yapılarak ön testleri geri verilmiş ve gerekli düzeltmeleri yapmaları istenmiştir. 7. Haftada, katılımcıların kazanımlarını yeni bir yazıya aktarıp aktaramayacaklarını görmek için her iki gruba da yeni bir konuyla ikinci bir son test verilmiştir. Sonuç olarak, geri bildirim türünden bağımsız olarak tüm katılımcı grubunun puanları ön testten son teste önemli ölçüde artmıştır. Ancak, bulgular iki geribildirim grubu arasında anlamlı bir fark göstermemiştir. Ayrıca toplam puan ortalamaları açısından birinci ve ikinci son testler anlamlı bir farklılık göstermediği için ikinci son testte tüm katılımcı grubunun kazanımlarını koruduğu görülmektedir. Tüm bunlara ek olarak, her iki grup katılımcılarına geribildirim türüne ilişkin görüşlerini inceleyebilmek adına birer anket uygulanmıştır ve bazı katılımcılarla birebir görüşmeler yapılmıştır. Sonuç olarak, katılımcılar her iki geribildirim türü için de olumlu görüşler bildirmişlerdir.

Anahtar Sözcükler: yazılı geribildirim, dolaylı ve doğrudan yazılı geribildirim, kapsamlı geribildirim

Title: The Effects of Written Corrective Feedback on L2 Writing Skill and Grammatical Accuracy

Author: Sibel ÖZÇELİK

ABSTRACT

This study investigated the effectiveness of indirect written corrective feedback in improving overall writing performance. Two intact English as a foreign language classes participated in the study. The participants were a total of thirty-six Turkish university students enrolled in a General English preparatory class at the School of Foreign Languages, Beykent University. While the experimental group received indirect feedback through error correction codes, the control group received direct feedback, which is a common and traditional way of providing written corrective feedback. Each class produced seven different pieces of writing and revised them over a seven-week period. In Week 1, both groups of participants were given a pre-test, and over the following five weeks, the experimental group was trained to use error correction codes. In Week 6, the participants were given their pre-tests with direct or indirect corrections, and they were asked to revise the same text. In Week 7, both groups were given a second post-test with a new writing prompt to see whether or not the participants could transfer their gains to a new piece of writing. As a result, the scores of the whole group of participants, regardless of the feedback type, significantly increased from the pre-test to the post-test. However, the findings showed no significant difference between the two feedback groups. It was also seen that the first and the second post-tests were not significantly different in terms of the mean Total scores, and this result implied that the whole group of participants retained their gains in the second immediate post-test. In addition, an online survey and interviews were carried out to see the participants' opinions towards each feedback type. The results showed that participants had positive opinions as to both types of corrective feedback.

Keywords: written corrective feedback, direct and indirect feedback, comprehensive corrective feedback

TABLE OF CONTENTS

THE LIST OF TABLES

THE LIST OF FIGURES

THE LIST OF APPENDICES

CHAPTER I

INTRODUCTION

1.1. Introduction

In the first chapter, background of the study, statement of the problem, research questions, aim and scope of the study, the significance of the study, assumptions, limitations, and definitions of terms are given. Subsequently, literature concerning written corrective feedback is presented.

1.2. Background of the Study

Trying to learn to use words, and every attempt
Is a wholly new start, and a new kind of failure by T.S. Eliot

In T.S. Eliot's words, writing is described as a continuous effort put forth to improve one's ability to use words, and it is a never-ending learning process. As stated in Eliot's description, writing is a difficult task even in one's own language. Even for native speakers of English, writing is "a bit of a chore" when they have to do it (Nunan, 2015, p. 77). Comparing speech with writing, Hedge (2000) states that in L2 writing, it is not possible to benefit from devices like gestures, facial expressions, and intonation. Besides, a speaker has the chance of rethinking, revising, and clarifying with the help of listeners who question or oppose the speaker's view. However, the writer has to find other ways to express herself or himself correctly. In effective writing, the organization of ideas or information, the range and accuracy of grammar and vocabulary, and finally creating an appropriate style for the subject matter and the

readers are key issues which make writing demanding for learners of English (Hedge, 2000).

Sokolik (2003) defines the L2 writing process as “a series of contrasts” (p.87). First of all, writing is a physical act as well as a mental act. It requires some manual action, either on a piece of paper or an electronic device like a computer used to commit words with a pencil or a keyboard. Moreover, it has an intellectual aspect, including the development and organization of ideas in order to present them properly. Moreover, according to Sokolik (2003), in writing, it is aimed “both to express and impress” (p.87). Writers have to satisfy both their needs to express an idea or feeling and those of their audience, who have some expectations about the written work. Finally, Sokolik (2003) emphasizes that writing is not only a product but also a process. It is a recursive process that the writer needs to go back and forth among various steps of writing, such as drafting, revising, and rereading until he/she creates a product to present to the audience.

This definition of writing refers to a shifting paradigm in teaching L2 writing. The end product is no longer the only focus of writing activity. While the product approach regards texts as objects that can be examined and construed without taking specific contexts, writers, or readers into consideration, the process approach sees writing as a personal expression of the text owner, a cognitive process which employs planning, drafting, revising, and editing. It is also a situated act that both past experiences of the writer and circumstances surrounding the writer have an effect on. Moreover, writing is an interactive activity through which writers communicate their ideas in ways that they best express themselves to their readers (Hyland K., 2009).

As for the role of the teacher in the L2 writing process, Silva (1990) states that the process approach requires a positive environment that motivates learners to improve their writing by giving sufficient time and assistance. Therefore, the teacher is expected to provide the necessary conditions, but also help learners in different stages of the writing process. As a teacher, it is not sufficient only to assign a writing task and then collect the papers in and mark them. The teacher could help learners find

topics, get ideas, organize their ideas into a text, get feedback on language use and content, write several drafts, revise and edit their papers (Scrivener, 2011).

As mentioned above, in the writing process the teacher is expected to provide feedback and encourage multiple drafts. Although there are some counterarguments to the effectiveness of corrective feedback (Truscott, 1996, 1999), both teachers and students value corrective feedback, and it is also expected by students (Hedgecock & Lefkowitz, 1994; Ferris, 1995b, 1999). Ferris (1999) emphasized the necessity and importance of feedback as follows:

> *... many teachers would respond that the absence of any form of grammar feedback could frustrate students to the point that it might interfere with their motivation and confidence in the writing class, particularly when grading rubrics and writing proficiency examination results tell them that their language errors could prevent them from achieving their educational and professional goals. (p. 8)*

In addition, when the scientific research regarding the benefits of written corrective feedback is analysed, most of them present valuable findings that support its effectiveness (Fathman & Whalley, 1990; Sheppard, 1992; Ferris, 2006; Chandler, 2003; Bitchener & Knoch, 2010).

There are a lot of different ways of providing feedback on the written work. One popular way of providing feedback is conferencing, which involves talking with individual students about their work during or after the writing process. Another way is the use of checklists that give students an opportunity to evaluate their work with the help of questions or statements. Reformulation is also highly preferred by teachers. It helps learners notice the differences between their work and the reformulated version by their teacher. In addition to these techniques, teachers can invite students to self-correct their errors by underlining, highlighting, or using codes referring to different types of errors or simply indicating the error. As an alternative, the teacher can guide

students to give peer feedback. All of these techniques are used to encourage learners to develop their writing skills (Hedge, 2000).

There are various kinds of written corrective feedback such as direct and indirect corrective feedback, and they are mostly combined with other ways of feedback such as brief written explanations, individual conferences and peer feedback. While direct feedback is the explicit correction of errors by the teacher, indirect feedback is a less explicit means which could be in the form of highlighting, underlining the error or using error correction codes corresponding to a particular error, and the student is responsible for the correction. Some researchers have claimed that direct feedback has better results despite their focus on a few particular features (Sheen, 2007; Bitchener, 2008; Bitchener & Knoch, 2008). Chandler (2003) argues that students internalize the correct form better when they see the corrections which was made by the teacher. Moreover, providing indirect corrective feedback causes delay in knowing if their hypothesized corrections are accurate or not, and it requires them to put much more cognitive effort on making their corrections. However, indirect corrective feedback has always been valued by L2 researchers as it provides learners the opportunity and guidance to self-correct their errors, and so enabling more lasting learning (Leki, 1991; Ferris & Roberts, 2001). Bitchener and Knoch (2008) suggest that indirect corrective feedback guides learners by assisting to solve the problems in their works, and thus it promotes acquisition in the long term. The degree of feedback explicitness has been researched for a long time, and the need for further research has always been emphasized.

1.3. Statement of the Problem

Speaking and writing known as productive skills are considered to be the most challenging skills for learners since learners are asked to be active users of the target language. As in the development of L1, errors are part of L2 writing, and corrective feedback provide opportunities to draw learners' attention to their production and comprehension. Schmidt (1990), in the Noticing Hypothesis, emphasizes the

importance of learner awareness of the gap in production, which is necessary for input to become intake. Similarly, Long (1996), in the Interaction Hypothesis, states that negative evidence is as important as positive evidence. In addition, Swain (2000), in the Output Hypothesis, argues that learners need to speak and write so that they discover what they can and cannot do. Therefore, learners need to use the language, and they also need teacher support to notice their gaps in the target language, and thus they can express themselves better.

Providing written corrective feedback is regarded quite valuable in that it supports the L2 writing development of learners but it is also difficult and time-consuming, especially in crowded classrooms. As a teacher, providing comprehensive feedback to the written works of my students take hours. On the other hand, I have always witnessed the negative attitudes of some of my colleagues towards indirect written corrective feedback as it demands more from the students, and it is assumed that only a few students will spend time on solving and correcting their errors when they are asked to do it. Moreover, it is argued that students may never know if their hypothesized corrections are accurate or not. On the other hand, the research that has been done so far is not conclusive. While some studies favour direct feedback, others claim that indirect feedback has better results, especially in the long term. Besides, the studies so far have mostly been limited to quantitative research. There are not much research which examine students' opinions and preferences.

For all these reasons, I have decided to test how indirect feedback works in a real classroom setting. I hope this research on the effectiveness of indirect written corrective feedback will contribute to the field.

1.4. Research Questions

The present study investigates the research questions given below:

RQ 1: Is indirect written corrective feedback effective in improving students' overall L2 writing performance?

RQ 2: What are the participants' opinions about direct and indirect written corrective feedback?

1.5. Aim and Scope of the Study

This study intends to see whether or not indirect written corrective feedback will be effective in improving the overall writing performance of students. As the control group will include students who get direct corrective feedback, it will be possible to examine the differences between the results of both types of corrective feedback. Moreover, the participants' opinions about indirect and direct corrective feedback are going to be examined.

1.6. Significance of the Study

The necessity for further research on written corrective feedback with varying degrees of explicitness has been underlined in previous studies. The present study is believed to contribute to the field as it presents both quantitative and qualitative data regarding the effectiveness of indirect corrective feedback on overall writing performance. Besides, the present study was conducted in a real classroom setting with a group of students who followed their studies via distance education due to the COVID-19 pandemic, and in this respect, the study differs from the previous ones. As a result, this study could help field teachers by providing insights into their online classroom activities.

1.7. Assumptions

It was assumed that both quantitative and qualitative data which were collected within the scope of the present study would provide satisfactory answers regarding the research questions. Also, the participants are assumed to reflect their true opinions regarding the interview questions.

1.8. Hypothesis

The present study aims to test the hypothesis that indirect written corrective feedback is more effective than direct written corrective feedback as it helps learners develop L2 writing skills by prompting autonomous learning behaviour, and so it enables more durable learning.

1.9. Limitations

The major limitation of the present study is that the number of participants, after the elimination of some for several reasons, was narrowed down from 48 to 36. The second limitation is that, due to the language level of students and the school curriculum, the type of writing is limited to opinion paragraphs.

1.10. Abbreviations

ANOVA: Analysis of Variance
ANCOVA: Analysis of Covariance
SLA: Second Language Acquisition
L1: First Language
L2: Second Language
RQ: Research Question
WCF: Written Corrective Feedback
CF: Corrective Feedback
ICF: Indirect Corrective Feedback
DCF: Direct Corrective Feedback
ME: Metalinguistic Explanation

CHAPTER II

LITERATURE REVIEW

2.1. Language and Language Learning Skills

Language is a complex phenomenon which has been defined by a great number of scholars. Sapir, who is one of the most influential linguists, (1921) defines language as "a purely human and non-instinctive method of communicating ideas, emotions, and desires by means of voluntarily produced symbols" (p.8). Another prominent linguist, Noam Chomsky (1957), defines language as "a set (finite) or infinite) of sentences, each finite in length and constructed out of a finite set of elements" (p.13). Halliday (1970) states the different functions of language as follows:

> *Language serves for the expression of 'content': that is, of the speaker's experience of the real world, including the inner world of his own consciousness. [ideational function]...Language serves to establish and maintain social relations. [interpersonal function]...Finally, language has to provide for making links with itself and with features of the situation in which it is used. [textual function]. (p.143)*

For all these vital functions of language, language learning plays an important role in human life. It mainly involves mastering four sub-skills known as listening, speaking, reading and writing. While listening and reading are receptive skills that require learners to extract meaning from discourse, speaking and writing are productive skills that learners need to use language for creative construction (Harmer, The Practice of English Language Teaching, 2007). Nunan (2015, p.78) groups these skills with regard to their mode of communication as follows:

	Productive	Receptive
Visual	Writing	Reading
Aural	Speaking	Listening

Figure 1. The four skills in terms of mode of communication

Although language skills are divided into different categories, they are not taught in an isolated way, and these skills "feed off each other" in many ways as also shown in the figure below (Harmer, 2007; 266).

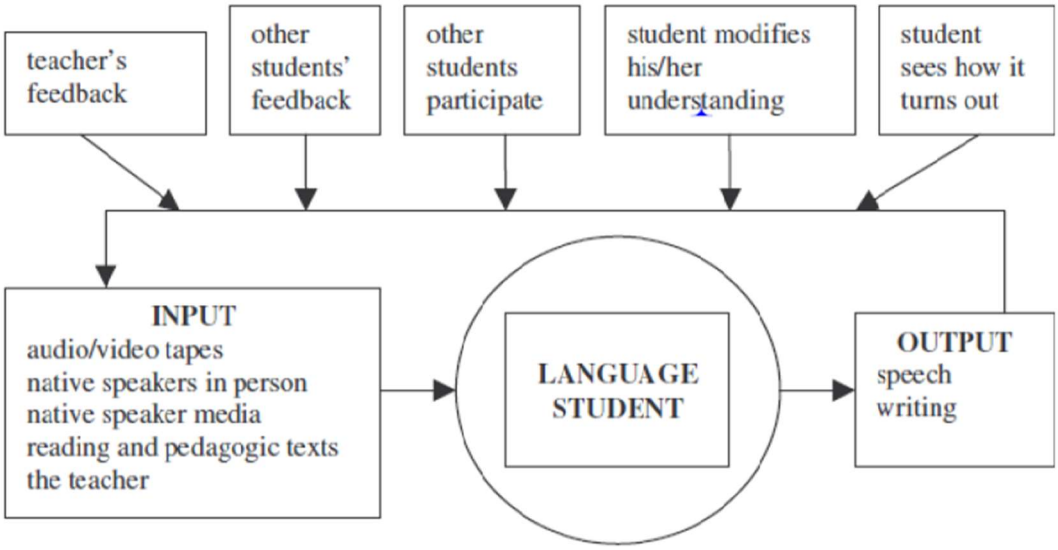

Figure 2. The circle of input and output

As seen in the figure above, learners receive input from many different sources: listening audio, teachers, native speakers, or the reading texts used in class. In addition, learners may get the opportunity to read extensively and do extra listening practice outside of the classroom. Apart from all these, output, which could be in the form of written text or speech, provides input for learners. The output by a learner "is fed back into the acquisition process" as learners revise their own output to communicate their ideas effectively or other people push learners to modify their output (Harmer, 2007; p. 266).

In addition, it is necessary to refer to the processing of input and output. It is frequently divided into two parts: top-down and bottom-up processing. Harmer (2007) uses the metaphors of "looking at a forest" or "studying the individual trees in it" to define the difference between them (p.270). In top-down processing, with regard to receptive skills, the reader or the listener tries to comprehend the text or the audio by looking at the whole, whereas in bottom-up processing, connecting the details such as words, phrases, cohesive devices enables the reader or the listener to comprehend the material and form the overall picture. Learners make use of top-down and/or bottom-up processing regarding productive skills, as well. It may be the details or the general overview that help learners construct meaning (Harmer, 2007).

2.1.2. Writing as a Language Skill

The primacy of speech over writing has long been discussed by philosophers and linguists such as Rousseau, Bloomfield, and Saussure, who adopt a phonocentric approach towards the origin of languages. Saussure (1971) supports the traditional definition of writing as in Aristotle and Plato, and claims that (cited in Derrida, 1976):

> *Language and writing are two distinct systems of signs; the second exists for the sole purpose of representing the first. The linguistic object is not both the written and spoken forms of words; the spoken forms alone constitute the object. (pp. 30, 31)*

Similarly, the linguist Bloomfield (1984) puts forward that the primary language form is spoken language, and says that "for the linguist, writing is, except for certain matters of detail, merely an external device, like the use of the phonograph, which happens to preserve for our observation some features of the speech of past times" (p.299).

Considering speech as a more natural and passionate form of communication, Rousseau (2018) believes that writing causes language change in character, and he believes that writing is just a tool used to secure the spoken language. This process of securing results in considerable changes in language as written language has to conform to its accepted rules.

As opposed to philosophers and linguists like Rousseau and Saussure who, downgrade writing and give priority to speech, Derrida (1976) believes that writing is not a simple supplement to the spoken word and that the written word is equally valuable as speech. For Derrida, writing and speech rely on each other. In his criticism of the "Essay on the Origin of Languages" by Rousseau, he suggests that the oppositions like death/life, evil/good, signifier/signified, writing/speech inhere within each other.

In addition to the discussions about the primacy of speech over writing, some scholars made distinctions between speech and writing to show how they contribute to language in their own ways. Ong (2012), referring to Plato, compares writing to printing and the computer. He says that Plato's age regarded writing "as an external alien technology" as we experienced the same with printing and the computer (p.80). Ong (2012) points out the artificiality of writing by saying that writing naturally is impossible as "the process of putting spoken language into writing is governed by consciously contrived, articulable rules…" (p.81). However, Ong (2012) also emphasizes the significance of such artificial creations, including writing:

> *To say writing is artificial is not to condemn it but to praise it. Like other artificial creations and indeed more than any other, it is utterly invaluable and indeed essential for the realization of fuller, interior, human potentials. Technologies are not mere exterior aids but also interior transformations of consciousness, and never more than when they affect the word. (p.81)*

Reviewing the literature regarding writing as a language skill, it is concluded that writing was regarded as a medium to fix the spoken language by its artificial and

constricting rules. However, today it is acknowledged that writing is unique with its distinctive nature and contributes to thought processes and creativity.

2.2. Teaching L2 Writing

With the acknowledgement of writing as a major language skill like speaking, reading, and listening, more emphasis has been placed on teaching writing. If it is aimed at teaching writing better, understanding its nature is important. According to Weigle (2014), second-language writing requires certain skills and knowledge, and in this respect, it is a cognitive process. Writers need both writing ability and second language proficiency. Furthermore, writing is a means of communication with the members of a particular society and culture, which requires writers to use language in certain ways to achieve their different communicative goals. It is thereby a process in which writers become members of a discourse community (C. Weigle, 2014).

2.2.1. The Distinctive Nature of L2 Writing

Written language differs from spoken language in many ways. To begin with, in writing, writer and the audience are distant from each other, which makes it difficult for the writer to check if the message is communicated or not. While the audience in conversation can ask questions, react or show a lack of understanding even without words, for e.g. using facial expressions, gestures etc., writer is not able to make use of such signs (Horning, 1986). Participants (addresser and addressee) in spoken communication are in close interaction, and during the conversation, it is often possible for them to change roles, and thus the conversation continues in a less structured way in contrast to the structured nature of writing (Harmer, 2004).

Secondly, the process of writing is what distinguishes writing as a different skill from speaking. Speaker has to make decisions regarding what to say and how to express himself/ herself at the time of speaking. Speaker only relies on modifications

and rephrasing that is available to him/ her instantly. On the other hand, writer is able to create multiple drafts before the final product is ready to present (Harmer, 2004). What makes writing cognitively challenging for writer is the need to consider readers' knowledge, goals, and interest and choose words accordingly (C. Weigle, 2014).

The structured nature of language is another difference between writing and speaking. A writer needs to organize his/her words into paragraphs, chapters, etc. to create a clearer structure than speech. Moreover, this structure requires the writer to use organizational features of writing like markers of coherence (Horning, 1986). In addition, "the level of correctness" and "well-formedness" are of great importance to writers. Whereas mispronouncing or grammatical mistakes do not pose a serious problem, correct use of grammar and spelling are crucial to writers so as not to be judged by the audience (Harmer, 2004; p.9). Allen (1966) points out that conventions of writing such as paragraphing, punctuation, and spelling are mastered at school with the help of a teacher. Therefore, he emphasizes that written English, with its unique devices, is like a second language that students learn from their teachers.

2.2.2. Product Approach versus Process Approach

Earlier in the 1940s, formal accuracy and correctness used to take precedence over creativity and originality as writing was thought to be a habit formation and, therefore, the learner was expected to make use of language structures that were already presented (Silva, 1990). The audience and purpose were completely ignored. From the mid-1960s onward, the inadequacy of controlled composition was revealed, and more space was provided for free writing. However, learners were still bound to construct discourse within a form, such as a paragraph with a topic sentence, supporting sentences, and concluding sentences, which would later get more complex and develop into an essay. Students' attention was focused on form, and their papers followed a linear pattern, and thus the emphasis was put on the end product instead of the process of writing and the individual writer (Silva, 1990).

In the following years, with the increasing awareness regarding the discouraging nature of linearity, creative thinking started to gain importance, and it caused researchers to reconsider the approaches towards teaching writing. Drawing attention to the writing process as opposed to the product, Zamel (1983) pointed out the complex and non-linear nature of the composing process. Zamel (1983) suggests that:

> *Understanding that writing may be recursive, non-linear, and convoluted, writers are able to modify or even discard chunks of discourse or original plans as they review their writing, reconsider its function, and distance themselves from it in order to meet their readers' expectations. It is in this way that they approximate more closely in writing what may only have existed on an intuitive level. (p.166)*

Harmer (2004) demonstrates different directions that writers can take in the course of writing by generating a process wheel. According to Harmer, it is not the final version until it "reached its culmination" (p.6).

Figure 3. The process wheel

As seen in the process wheel, the focus is not on the final product but on the writing process. According to process approach, students need assistance to excel in composing. Developing strategies for different stages of writing such as prewriting,

drafting, and revising the texts is important. For all these stages, students need enough time, and the emphasis was placed on the revision process so as to help students discover their opinions. Apart from these, it is necessary to give feedback all through the composing process not only on the final product as it helps students to approximate the meaning they aim at. Conferencing with students during composition process may help achieve this goal. Moreover, peer feedback is a part of process writing as well as teacher feedback (Brown, 2001).

In brief, the process of revision is highly valued since it helps students find out what they would like to say and how to express themselves properly. In order to support students during revision process, giving feedback is one of the key components in terms of giving directions which will lead to the end product.

2.3. Corrective Feedback in Second Language Writing

2.3.1. Definition of Corrective Feedback

Over the last few decades, corrective feedback has been thought to be a popular issue among teachers and researchers in the field of second language acquisition (SLA). 'Error correction', 'error treatment', 'negative evidence' are the other terms which are often used to refer it. Corrective feedback is defined as "an indication to the learners that his or her use of the target language is incorrect" (Lightbown & Spada, 1999; p.172). According to Lightbown and Spada (1999), corrective feedback may vary from implicit to explicit, and metalinguistic information could be added.

Long (1996) uses the terms "positive evidence" and "negative evidence" to identify input. Positive evidence consist of grammatically acceptable use of target language while negative evidence indicates implicitly or explicitly that the utterances by learners are ungrammatical (p.413).

Sheen (2010) points out the differences between oral and written corrective feedback. The main difference concerns the timing of feedback. Oral corrective feedback is provided as soon as the learner has committed the error, whereas the learner is provided with written corrective feedback some time later. Secondly, oral corrective feedback when compared to written corrective feedback is cognitively more demanding as learners depend heavily upon their short-term memory. However, referring to Polio (2001), Sheen (2010) states that oral corrective feedback may not be as complicated as written corrective feedback since written corrective feedback is focused on various aspects of writing such as grammatical accuracy, syntactic complexity, lexical items, content, coherence and fluency. Oral corrective feedback is simpler because it is only focused on the type of error that learners commit at the moment of speaking.

In brief, corrective feedback has an important role in terms of increasing motivation and improving linguistic accuracy (Ellis, 2009). Generally speaking, error correction has been considered essential for writing development, and students have usually expected and liked to be corrected by their teachers (Hyland & Anan, 2006).

2.3.2. The Role of Error and Corrective Feedback in SLA

To understand the role of error in SLA, it is important to distinguish between the two confusing terms: error and mistake. Gass et al. (2013), referring to Corder (1967), clarify the difference between mistake and error. Mistakes are similar to "slips of the tongue" and mostly happen once. Moreover, the learner can identify the mistake and correct it if it is asked for. However, the error is "systematic" and likely to be repeated, and besides the learner may fail to recognize the error due to lack of knowledge (p. 91). Furthermore, as the error is not thought to be an error according to the learner's system of interlanguage, it could only be defined as an error with regard to some external measure such as the target language (Gass et al., 2013).

Despite general view in favour of corrective feedback both by learners and teachers, the early perspectives towards error and error treatment were not as positive as they are today. During 1940s and for the next couple of decades, behaviourist perspective was dominant, and error was something to be prevented rather than being treated, and therefore the right models were needed to be practiced and memorized so as to form target-like habits (Bitchener & Ferris, 2012). Accordingly, as the primary objective was to prevent error, early written corrective feedback research did not raise questions on whether or not written CF support L2 acquisition. The early research focused more on the types of error, how, who and when to fix them. (Shaughnessy, 1977; Hendrickson, 1978,1980). As opposed to the early theorists who claimed that errors need to be prevented, Corder (1967) was one of the first researchers who underlined the importance of learner errors since their errors offer an insight into second language acquisition. In one of his following studies, Corder (1973) reinforces his argument for the importance of learner errors and their correction. He suggests that thanks to corrective feedback, learners can recognize their errors besides they form an opinion regarding both the use and the boundaries of target language in terms of its syntax and lexis.

When the history of SLA is reviewed, the first general theories (Dulay & Burt, 1973; Cancino et al., 1978; Krashen, 1981, 1982, 1985) centre on the idea that second language acquisition is a natural process that learners go through certain stages of development, which is universal and fixed for all. According to nativist views, there is no point in teaching grammar and providing corrective feedback as they do not contribute to creative construction. Krashen (1985), in his Input Hypothesis, claims that enough comprehensible input (i+1) which is a little beyond the learner's current level is the only requirement for language acquisition and "the necessary grammar is automatically provided" (p.2).

The first hypotheses did not refer to the role of learner error and how corrective feedback might contribute to L2 acquisition, however they led to a more detailed investigation due to their inability to explain the continuing existence of learner error despite learners being exposed to a good amount of input. It caused researchers to

direct their attention into the other components such as interaction and output that might be involved with the process of acquisition (Bitchener & Ferris, 2012). Although the researchers formulated their hypotheses based on oral production of language, they are relevant to written language as well. From an interactionist perspective, input is not the only component which promotes L2 acquisition. Output and feedback also play an important role. Long (1996) in his interaction hypothesis suggests that a communication problem causes interactional modifications that learners are supposed to negotiate the meaning and thus provide the required input for L2 learning. Long (1996) states the importance of negative evidence as well as positive evidence:

> *... environmental contributions to acquisition are mediated by selective attention and the learner's developing L2 processing capacity, and that these resources are brought together most usefully, although not exclusively, during negotiation of meaning. Negative feedback obtained during negotiation work or elsewhere may be facilitative of L2 development, at least for vocabulary, morphology and language-specific syntax, and essential for learning certain specifiable L1-L2 contrasts. (p.414)*

Similarly, Swain (2000) in the output hypothesis emphasizes the importance of language use. What caused her to put forward the output hypothesis was her study on children learning French in French immersion programs conducted in Canada. Even though they were exposed to French for several years, their results in different areas related to grammar, discourse, and sociolinguistics were not as satisfactory as they were with native French-speaking children. When both groups of learners were compared, Swain concluded that the difference between native French speaking children and French immersion students was their opportunity to use the target language in productive ways. In conclusion, Swain argued that "in speaking or writing, learners can 'stretch' their interlanguage to meet communicative goals. They need to create linguistic form and meaning, and in so doing, discover what they can and cannot do" (Swain, 2000, p.99).

Schmidt (1990), in the noticing hypothesis, mentions another important element which is essential for acquisition. He claims that learning takes place when

learners consciously pay attention to some certain linguistic forms in the input (Schmidt, 1990). According to Schmidt, “noticing is the necessary and sufficient condition for converting input to intake” (Schmidt, 1990; p.129). It means that corrective feedback with its focus on form may be a part of effective learning as it helps increase salience for the target language, and also prevent fossilization (Doughty & Varela, 1998; Lightbown, 1998).

In addition, DeKeyser (2015), in the skill acquisition theory, emphasizes the importance of explicit knowledge as much as implicit knowledge. In this theory, declarative and procedural knowledge are thought to be equivalent and complementary in that “snergy” is created between the two types of knowledge to learn “a variety of different rules, patterns, or regularities” (DeKeyser, 2015; p.105). While some rules are more likely to be learned explicitly, the others may not be so and implicit learning leads to better results for them (Ferman et al., 2009). Learners may need an expert to directly point out their erroneous uses.

Moreover, Segalowitz and Lightbown (1999) approaches the focus-on-form debate from a psychological perspective, and presents the transfer appropriate processing. The transfer appropriate processing suggests that when compared to language features learned outside a communicative context, learners access knowledge about language features noticed during classroom-based communicative activities more easily, and therefore it will be easier for learners to transfer knowledge to a communicative situation afterwards. In this sense, corrective feedback might enable learners notice the language features and provide the necessary knowledge within a communicative context, and then transfer this knowledge to their future works.

More recently, Lyster and Mori (2006) draw attention to the counterbalance between instructional activities and interactional feedback. Lyster and Mori (2006) put emphasis on the balance relying both on form-focused and meaning-focused practices. Learner’s attentional focus is shifted from meaning to form in a meaning-oriented context, and from form to meaning in a form-oriented context, and so it is thought to

"strengthen connections between changes in long-term memory and actual language use" (p.294).

The Noticing Hypothesis (Schmidt, 1990) the Interaction Hypothesis (Long, 1996), and the Output Hypothesis (Swain, 2000) and the research on Focus on Form(Doughty & Varela, 1998; Lightbown, 1998) draw attention to the usefulness of corrective feedback when focus is both on form and meaning so that learners can see how linguistic form assists to convey meaning. Taking all these approaches into consideration, it could be concluded that corrective feedback plays an important role in creating awareness on form and meaning, and so push learners to modify what they have produced, and eventually promote acquisition.

2.3.3. A Typology of Written Corrective Feedback Types

Direct correction and indirect correction are the two main methods used to provide written corrective feedback (Sheen & Ellis, 2011). As shown in the table below, in direct correction, learner is provided the correct form or the reformulated sentence/ text. In indirect correction, learners are expected to find out the correct form, and the error is not obviously corrected, instead it is highlighted in different ways such as simply underlining the erroneous structure or using some codes referring to specific errors or the error is sometimes only indicated in the text. Both strategies may also be accompanied with a brief explanation concerning the problem. However, it is not possible to make a distinction as explicit or implicit as in oral corrective feedback. WCF is always explicit because the student is aware of the fact that he/she has been corrected (Sheen & Ellis, 2011). Sheen and Ellis (2011) provide a taxonomy of WCF strategies in the table below:

Table 1.

A taxonomy of written corrective feedback strategies

	Direct	Indirect
Metalinguistic information	• Provision of the correct form with brief grammatical explanation.	• Use of error code (e.g., symbols such as VT to signal a verb tense error or WO a word order error are inserted into the text). • Use of brief grammatical explanation (e.g., error types are numbered in the text and then a brief explanation of each type of error is provided at the end of the text).
No metalinguistic information	• Provision of the correct form only. • Reformation of the entire sentence or paragraph.	• Errors are indicated but not located and the correct form is not supplied (e.g., a cross is placed in the margin next to the line where an error has occurred). • Errors are indicated and located but the correct form is not supplied (e.g., an error is underlined in the place in the text in which it occurs).

(p. 594)

2.4. Key Studies on Written Corrective Feedback in L2 Writing

2.4.1. Early Research and Design Issues

Until 1996, there were not many empirical studies on the effectiveness of written corrective feedback (Lalande, 1982; Semke, 1984; Robb et al, 1986; Fathman & Whalley, 1990; Kepner, 1991; Sheppard, 1992). One reason is that writing, when compared to other skills, did not play a major role in second language instruction. Also SLA theories developed by Krashen which emphasize the importance of input rather than output caused writing and written corrective feedback to take a backseat (Ferris, 2010). Moreover, these empirical studies have been criticised because of their methodological problems such as lack of a control group and inconsistencies in their designs, which make it difficult for researchers to compare their results (Ferris, 2004).

Only three of the studies mentioned above test the effectiveness of correction versus no correction (Semke, 1984; Fathman & Whalley 1990; Kepner, 1991). Only one of them presents positive evidence for the effectiveness of correction (Fathman & Whalley, 1990). The three treatment groups who received feedback on grammar only, on content only, and both on grammar and content all benefited from feedback in varying degrees.

Semke (1984) did not find any significant differences in terms of accuracy across four treatment groups. Nevertheless, the results of the study were unclear because he did not measure progress over time by also providing pre-test scores.

Kepner (1991) found positive results regarding the effectiveness of error correction as the treatment group made %15 fewer errors than the control group. However, Kepner interpreted the results in a negative way as he believed error correction was only helpful for low verbal ability learners.

Furthermore, Lalande (1982) concluded that the two treatment groups who received indirect corrective feedback performed better than the control groups, but they were not real control groups as they received direct corrective feedback.

Finally, Robb et al. (1986) and Sheppard (1992) did not have any control groups. While Robb et al. (1986) did not present any significant differences across four treatment groups, Sheppard suggested that both treatment groups showed improvement.

From the 1996 onward, the interest in written corrective feedback increased when Trustcott (1996) published the controversial article “The Case Against Grammar Correction in L2 Writing Classes”. Truscott argued that “grammar correction in L2 writing classes should be abandoned” because it is ineffective and harmful (p. 327). Despite the counterarguments and research evidence for the usefulness of corrective feedback, he reiterated his argument in his following articles (Truscott 1999, 2004,

2007, 2009). Truscott (1996) argued that teachers uncritically accept corrective feedback, and rely on their intuitions on its effectiveness rather than research evidence which reveals theoretical issues and practical problems on the subject. He points out that it is wrong to assume that learners will easily comprehend and remember the right form in the future when it is provided only once before. Truscott (1996) emphasizes that a grammatical structure is acquired over time, and expecting an abrupt change in accuracy level of learners is not realistic.

Secondly, Truscott (1996) puts forward that one single method for all different types of errors, including syntactic, morphological, and lexical ones, is not effective as they are acquired in different ways. According to Truscott, different learning processes for different components of language should be taken into consideration so as to avoid the mere transfer of knowledge which does not contribute to acquisition.

In addition, Truscott (1996) draws attention to learner readiness to acquire a particular linguistic form. In order to benefit from grammar correction, a teacher should consider developmental sequences and be selective in correcting errors (although later, Truscott cites some research evidence and denies the effectiveness of being selective). According to Truscott, correcting all errors and providing comprehensive corrective feedback is not a solution due to the distraction that learners experience when they are exposed to corrections that they are not ready for.

Another point that Truscott (1996) made is “the problem of pseudolearning”. Truscott (1996) doubts the knowledge gained with the help of grammar correction because this type of teaching/ learning will not contribute the development of interlanguage and it does not lead to a thorough and permanent learning, and therefore it does not have any positive effects on actual use of language. To support his point of view, he cites Ellis’ modelled learning (1988) and the distinction he made between implicit and explicit knowledge (1993,1994), Lightbown’s pseudoacquisition (1985), McLauglin’s (1990) and Long’s (1977) concerns about integrating new knowledge into the system, and Krashen’s (1987) concept of learning versus acquisition. What all these theories have in common is that they make a clear distinction between knowledge

of language and competence. If the latter does not develop, the former is of no value for acquisition but only for editing purposes, albeit limited. Truscott (1996) claims that second language acquisition relies on "intuitions of well-formedness, coming from the unconscious language system, than on metalinguistic knowledge of points of grammar" (p.347).

In addition to theoretical issues, Truscott (1996) cites some practical problems concerning written corrective feedback. First of all, he points out teachers' inability to notice and identify errors. He claims that even if they notice an error, it does not guarantee that they are able to offer good and sufficient explanations that reflect actual English grammar. Moreover, teachers may not devote their full attention to students' written works due to time constraints and lack of patience, even if they are qualified to provide high-quality feedback. Another problem that Truscott (1996) underlines is the difficulty of being "consistent" and "systematic" in corrections for a teacher owing to busyness, and therefore it gets more difficult for a learner to understand and remember corrections to apply to future works. Besides, learners may not be motivated enough to rewrite in the long term.

Apart from all these claims, Truscott (1996) argues that grammar correction is harmful because it demotivates learners and increases their stress levels. Moreover, he maintains that correction causes students to write short and simple sentences or texts so as not to make errors, and thus it reduces the complexity of their texts. Besides, Truscott believes that instead of struggling to understand corrections or written comments, which mostly results in misunderstanding or not understanding at all, the time spent on correction both by students and teachers could be used more productively for other activities such as extra writing practice.

Finally, Truscott (1996) presents his opinion against the idea that error correction is necessary to avoid fossilization and that teachers should give feedback because students prefer to have. He argues that a teacher should act in the best interests of students rather than making judgements based on students' preferences.

Upon Truscott's strong claims about the ineffectiveness of grammar correction and its harmful effects, researchers published their responses to state their points of agreement and disagreement (Ferris 1999, 2006; Chandler 2003, 2004, 2009; Bruton 2009, 2010). Ferris (1999) calls Truscott's review "premature and overly strong" (p. 1). She points out three serious problems with Truscott's review of error correction studies, which are also stated in other responses to Truscott. Ferris (1999) draws attention to the fact that the subjects in different studies cannot be compared. Secondly, Ferris (1999) puts forward that various teaching strategies and research patterns are seen in the studies mentioned. Finally, Ferris (1999) states that Truscott overemphasized negative evidence by ignoring the findings against his thesis.

Despite points of disagreement, Ferris (1999) agrees with Truscott (1996) in that different kinds of errors including morphological, syntactic and lexical need to be treated in different ways. On the other hand, Ferris accepts the difficulty of categorizing student errors because not all them are treatable as part of a single system which can be used for all types of errors but there may be a need for a combination of different strategies. Ferris also acknowledges some practical problems that Truscott suggests. Ferris understands his concerns about teachers' varied abilities and enthusiasm to notice and identify errors and give proper feedback, and how unmotivated students could be to deal with feedback. Therefore, Ferris emphasizes the importance of teacher education and opportunities to practice giving feedback. In addition to all these, she states that being realistic and taking student-related factors (such as L1, language proficiency, learning styles) into consideration will help overcome the problems mentioned in Truscott's "gloomy assessment" (Ferris, 1999, p. 7). Besides, underlining student opinion in favour of teacher feedback presented in surveys, Ferris stresses that if a teacher prefers not to provide feedback, it will have unintended consequences both for students and teachers especially when students are faced with their poor proficiency exam results which are full of language errors. Most importantly, Ferris suggests that further empirical research that considers several important factors is needed.

Truscott's abovementioned strong claims and the inadequacy of empirical studies on the issue caused more researchers to study the subject. Ferris (1997) studied on the effects of three different types of teacher commentary that included comments in the form of questions, imperatives and requests, and there was no control group. She found that requests for information given in the margins and grammar comments resulted in much better revisions. Longer comments were more effective than shorter and general ones, such as positive comments.

Lee (1997) conducted a study on the two treatment groups who received indirect CF. The first treatment group's errors were underlined, whereas error-free sentences were ticked in the margins of the text in the other one. The control group did not receive any corrective feedback. The group whose errors were underlined performed better than the other groups whose errors were not marked at all or ticked in the margins only.

Ashwell (2000) provided corrective feedback on content and form to three treatment groups, and one group was not provided with any corrective feedback. Learner errors regarding form were underlined or circled. As a result, no significant differences across the three experimental groups were seen, although all of them outperformed the control group in terms of accuracy.

Ferris and Roberts (2001) studied the effectiveness of coded and uncoded feedback. They showed that both treatment groups' performances were much better than the control group who had no CF. On the other hand, the results were not significantly different between the treatment groups.

Fazio (2001), in contrast to other researchers, did not present positive evidence for the effectiveness of error correction and teacher commentary. The three treatment groups received error correction, teacher commentary, and both of them, respectively. However, no effects on accuracy were reported.

Chandler (2003) carried out research on ICF. Of the two treatment groups, the one who received ICF showed a significant improvement in accuracy, in contrast to the control group, who did not improve in accuracy.

Gascoigne (2004) conducted a replication of Ferris's (1997) study. She concluded that teacher commentary contributed to improvements in writing.

As seen above, most of the early research presents positive effects regarding WCF. However, their weaknesses in design cannot be overlooked. The biggest problem is the lack of a control group because it prevents us from comparing the results of the students who received feedback with the ones who did not (Ferris, 2004; Bitchener, 2008). Moreover, early research measures the effectiveness of corrective feedback on revision accuracy only, which shows how accurately the same error is corrected from one draft to another. Truscott (2007) puts forward that revision studies do not prove that students are able to write accurately due to their short-term nature. Even though it has some positive results in the revision of the same piece of writing, revision accuracy does not show that student have understood the correction and can transfer this knowledge to new pieces of writing. Also, revision accuracy does not contribute to the complexity of future works. The final issue concerns the number of target structures for error correction. The majority of the past studies were comprehensive rather than selective. When the feedback categories are too broad, it is difficult for a learner to see the types of errors he/she needs most help with. Reducing the number of categories may help learners deal with the error when considering its difficulty (Bitchener, 2008).

2.4.2. Recent Research (since 2004)

By reviewing recent research on written corrective feedback, it is seen that researchers approach the subject from two different perspectives: written corrective feedback in SLA and L2 writing. Not only their beginning points and but also the questions they ask are different. While SLA researchers study whether or not written corrective feedback promotes the acquisition of a specific linguistic feature, L2 writing

researchers examine the effectiveness of WCF on improving the overall text quality. The SLA studies are designed studiously to follow a design including a pre-test, post-test and delayed post-test, and include a control group with treatment group(s). Researchers specifically focus on a few error types. The L2 writing studies, however, are conducted in a writing classroom setting, and so may not include a control group for ethical reasons. Also it may not be possible to be selective and limit error types to only a few (Ferris, 2010). As a consequence of this, the distinction between two groups of studies bring into question the "practical applicability" in terms of interpreting the findings of the studies (Ferris, 2010; p. 186). The L2 writing studies are criticised for not being controlled enough, whereas the SLA studies are criticised for not reflecting the realities of an L2 classroom because of their narrow focus. Although a few L2 writing studies have control groups, they still lack a pre-test, post-test and delayed post-test as in SLA studies. Even so, Ferris (2010) emphasizes that "the two lines of research are not in competition; rather they are complementary" (p.191).

The majority of the recent studies have improved in design and addressed the issues of the past studies. The major improvement is that they do not only have treatment group(s) but also a control group. In addition, they have focused on varying degrees of feedback explicitness to investigate the effectiveness of different corrective feedback types. However, there is still an ongoing debate as to the effectiveness of a particular kind of corrective feedback, although there is an agreement that written corrective feedback is effective. Furthermore, researchers have also measured the effectiveness of focused versus unfocused feedback. Recently, some researchers have broadened the scope of their studies, and measured the effectiveness of WCF on new pieces of writing in the long term besides revision accuracy (Karim & Nassaji, 2019).

2.4.2.1. Relevant research for the effectiveness of WCF

Bitchener et al. (2005) conducted research that brought different types of feedback together. Of the three treatment groups, the first group received DCF and a 5-minute student-teacher conference after the completion of writing. The second group received DCF only. The last group received no CF on targeted structures (the past simple tense, the definite article, and prepositions), but for ethical reasons, they were provided feedback concerning the quality and content of their works. Bitchener et al. (2005) concluded that those who received a combination of DCF and individual conference with the teacher performed much better than the other two groups when examining the accuracy of the definite articles and the past simple tense, which were used in new pieces of writing. Bitchener et al. (2005) claimed that it did not apply to prepositions due to their idiosyncratic nature.

Sheen (2007) carried out quasi-experimental research with two treatment groups who received direct correction only, direct metalinguistic correction, and a control group was also included. On the immediate post-tests, the treatment groups who were given feedback on definite and indefinite articles had better results than the control group. On the other hand, in the delayed post-tests, the direct metalinguistic group outperformed the direct-only group and the control group.

Bitchener (2008) studied the effectiveness of DCF. The three treatment groups received DCF with written and oral ME, DCF with written ME, and DCF only. In addition, the control group did not receive any CF. The targeted structures were referential indefinite ‘a’ and referential definite ‘the’. It was found that the treatment groups had better results than the control group, and also the delayed post-test showed that their performance was retained 2 months later.

Bitchener and Knoch (2008) conducted similar research to Bichener’s study (2008) in the same year. The three treatment groups received DCF with written and oral ME, DCF and written ME, and DCF only, and the control group received no CF. The targeted structure was referential indefinite ‘a’ and referential definite ‘the’. As a

result, all three WCF groups performed better than the control group. He also stated that the accuracy level was retained over seven weeks.

Ellis et al. (2008) conducted a quasi-experimental study on the effectiveness of focused versus unfocused feedback. The targeted structures were the English definite and indefinite articles. While the focused group received CF on article errors, the unfocused group received CF on article errors and others. The control group did not receive any CF but general comments or questions. It was seen that both of the treatment groups outperformed the control group, and showed improvement not only in the error correction test but also in the new piece of writing. However, from post-test 1 to post-test 2, the focused group showed consistency in accuracy level, whereas the unfocused group experienced a decline on both the error correction test and the new piece of writing.

Van Beuningen et al. (2008) examined the effects of direct and indirect WCF on L2 writing accuracy. The two experimental groups received DCF and ICF, and the two control groups were asked to have extra writing practice and revision without CF. Direct and indirect CF both have short-term effects on improving accuracy, but only DCF has a significant effect in the long-term. Also, the control treatments were not found to be effective for accuracy.

Bitchener and Knoch (2009) conducted a second study on the relative effectiveness of different types of DCF. The three experimental groups received DCF with written and oral ME, DCF with written ME, and DCF only. The targeted structures were referential definite ‘the’ and referential indefinite ‘a’. There were no differences as to the level of accuracy across the three experimental groups.

Sheen et al. (2009) conducted a quasi-experimental study to examine the differential effects of focused and unfocused CF on the accuracy of grammatical forms. The three treatment groups received focused CF, unfocused CF, and practiced writing, whereas the control group received no CF. The targeted structures were English definite and indefinite articles for the focused CF group, and copular ‘be’,

regular past tense, irregular past tense, and preposition for the unfocused CF group. They concluded that all of the experimental groups improved in grammatical accuracy over time in all the post-tests. Of the four treatment groups, the highest accuracy gain scores were those of the focused CF group, and they were followed by the writing practice group, the unfocused CF group and the control group.

Storch and Wigglesworth (2010) carried out research on the efficacy of direct and indirect WCF. While one group whose errors were reformulated received direct CF, editing symbols were used to give CF to the other one. They concluded that editing feedback evoked more language-related episodes, which directly relate to the feedback given. Moreover, it was found that the level of engagement was higher with editing feedback when compared to the response to reformulation, which was limited to reading the reformulated sentence or text.

Van Beuningen et al. (2012) investigated the effectiveness of direct and indirect comprehensive CF. While the two experimental groups received direct and indirect CF, one control group was asked to self-edit their works, and the other one had writing practice without CF. Three main categories (lexical errors, grammatical errors, and orthographical errors) were used to classify different types of linguistic errors. In conclusion, it was seen that both experimental groups outperformed the control groups in revision and also in new pieces of writing. Moreover, they also concluded that only DCF resulted in accuracy gains in the grammar of the new texts, while ICF was more effective for non-grammatical accuracy.

Shintani and Ellis (2013) researched the comparative effects of DCF and ME. The two groups of participants received DCF, ME, and there was a control group. The targeted structure was the English definite article. They concluded that DCF had no effect on the accuracy level of the targeted structure, and it did not enhance implicit or explicit knowledge. However, metalinguistic explanation caused improvements in the accuracy level of a new piece of writing administered shortly after the treatment, but it was not seen in the second new text administered two weeks later.

Al-Rubai'ey and Nassaji (2013) carried out research with two intact EFL classes who were provided with DCF and indirect metalinguistic CF. The target structure was English articles. As a result, no significant differences across the two treatment groups were observed.

Gholaminia et al. (2014) compared DCF with metalinguistic code correction. The experimental group outperformed the control group who received DCF. Through several weeks of metalinguistic code correction, learners tended to pay more attention to mistakes and errors.

Shintani et al. (2014) measured the effects of DCF and ME on the accuracy of use of indefinite articles and the hypothetical conditional. Five groups of participants received ME, DCF, ME with revision, DCF with revision and there was a comparison group. All types of feedback caused gains in accuracy for the hypothetical condition, but they were not effective for the indefinite article. DCF with revision was found to be the most effective type, besides being longer lasting than ME.

Frear and Chiu (2015) studied the effects of focused and unfocused CF on weak verb accuracy and total accuracy. The experimental groups were given focused and unfocused ICF, and the control group was given no CF. The two experimental groups performed better than the control groups in the post-tests, and there were no differences in the level of accuracy between the two experimental groups, and both types of ICF enabled learners to push their output in the total accuracy of new texts (post-tests).

Lopez et al. (2018) carried out research on the differential effects of comprehensive feedback. The participants were randomly assigned into five groups: DCF for grammatical errors; metalinguistic codes for grammatical errors; DCF for grammatical and non-grammatical errors; metalinguistic codes for grammatical and non-grammatical errors, and a control group. They concluded that learners benefited from direct corrections and codes during text revision, but only direct corrections were effective for accuracy in the long-term.

Karim and Nassaji (2018) measured the effects of comprehensive WCF both on revision accuracy and new pieces of writing. The participants were randomly assigned to one of four groups: DCF, ICF (underlining only), ICF (underlining and metalinguistic cues), and a control group. All the experimental groups performed better than the control group in revision tasks. In addition, direct CF and underlining with metalinguistic cues led to accuracy improvements in new text, albeit largely non-significant.

Suzuki et al. (2019) researched the interactional effects of direct and indirect WCF explicitness on revision accuracy and new pieces of writing. The participants were assigned to four groups: DCF with ME, DCF only, ICF with ME, and ICF only. The targeted structures were the past perfect tense and the English indefinite article. They stated that both types of WCF helped learners improve revision accuracy of both target structures in the short-term, but the past perfect was the only structure that significantly improved in new pieces of writing in the long-term. In addition, partial effects of feedback explicitness on the short term were found for the past perfect tense but not for new pieces of writing.

2.4.2.2. Relevant research against the effectiveness of WCF

Although the majority of the recent research presents evidence for the effectiveness of WCF, a few studies present evidence for Truscott's arguments for the ineffectiveness of WCF in improving accuracy and complexity.

Liu (2008) conducted a quasi-experimental study regarding the effectiveness of error feedback. The participants were divided into two groups: direct correction and indirect correction (underlining the error without providing any correction). The error types were classified into three categories: morphological errors, semantic errors, and syntactic errors. As a result, both types of feedback enabled learners to improve their self-editing skills to some extent. However, although direct correction was effective for reducing the errors in the immediate draft, there was no improvement in accuracy

in a different text. On the other hand, the students benefited from indirect feedback in reducing more morphological errors than semantic errors. Furthermore, it was concluded that some mini-lessons or workshops are also necessary to improve the accuracy level and students' self-editing abilities.

Truscott and Hsu (2008) carried out research on the effectiveness of correction for improving writing ability. The experimental group received ICF that their errors were underlined, and the control group received no CF. As a result, the experimental group performed better than the control group in revision accuracy. However no significant differences across the groups in the new narrative task were observed. They therefore put forward that revision accuracy does not indicate learning and improvements in writing ability.

Hartshorn et al. (2010) examined the effects of dynamic corrective feedback, which is based on the needs of learners and purposeful, immediate, consistent, and feasible writing tasks and feedback. The experimental group received dynamic WCF, and the traditional process approach to writing instruction was used for the control group. The results showed that the dynamic WCF was not effective for rhetorical competence, writing fluency, and writing complexity. However a significant improvement was reported for writing accuracy.

CHAPTER III

METHODOLOGY

3.1. Introduction

The research design of the present study is given in this chapter. Within this context, information regarding the research method, participants, data collection tools and procedure, and lastly data analysis and interpretation are provided. The present study aims to examine the effectiveness of indirect written corrective feedback in improving overall writing performance, and the study seeks answers for the research questions below:

RQ 1: Is indirect written corrective feedback effective in improving students' overall writing performance?
RQ 2: What are the participants' opinions about direct and indirect written corrective feedback?

3.2. Research Method

In order to conduct this study, a mixed-methods research design, which employs both quantitative and qualitative methods in a single study, was preferred. Tashakkori and Creswell (2007) define mixed methods research as "research in which the investigator collects and analyses data, integrates findings, and draws inferences using both qualitative and quantitative approaches or methods in a single study or program of inquiry" (p.4).

The present study involves the mixing of qualitative and quantitative methods both in the procedure of data collection and in the analysis of data gathered. This type of research design possibly enables the researcher to see the research subject from different perspectives and it allows a better understanding of the research (Johnson et al., 2007).

To answer the first research question, a quasi-experimental study was conducted with two groups of participants. Convenient sampling was used to conduct the research. While the experimental group received indirect written corrective feedback, the control group received direct written feedback, which is a common and traditional way of providing written corrective feedback.

The results of the pre-tests, and post-tests which aimed to measure improvement in the participants' overall writing performance by providing indirect written corrective feedback in the experimental group, comprised the quantitative data for the study.

As for the second research question, qualitative research was conducted which involved a survey including open-ended questions to find out the participants' opinions about each feedback type, and semi-structured questions which were used to interview a certain number of participants from each group. All these tools helped to develop a better understanding of the effectiveness of indirect written corrective feedback.

3.3. Participants

The participants in the present study were thirty-six university students enrolled in a General English preparatory class at the School of Foreign Languages at Beykent University, which is a foundation university in Istanbul, Turkey. The students who are able to succeed in the placement test (The Oxford Placement Test), which is a computer-adaptive test for non-native speakers of English, and then in the proficiency exam, are eligible to start in their departments. However, those who are

not able to get the required score on the placement test are required to attend foreign language preparatory classes, and they are assigned to classes depending on their placement test scores.

In the present study, the two groups of participants who were assigned to their classes according to the placement test results were both A2 level. The participants who enrolled in an undergraduate/associate degree program where English is partly or totally used as the medium of instruction, such as Computer Engineering, or Architecture are required to attend English Preparatory Classes and to complete at least the B1+ level. On the other hand, the participants who enrolled in English Language and Literature, Translation and Interpreting, and Applied English and Translation are required to complete at least the B2 level. The participants in both groups were from different programmes. Besides, they were all non-native speakers of English.

To conduct the research, convenient sampling was used by the researcher, who is an English lecturer at the same university. At the beginning of the study, each group included 25 students. However, those who did not or could not follow the lessons regularly for several reasons or those who were unwilling to participate in the study were eliminated. When the treatment was over, there were 18 students from each group whose data could be used for the study. In the experimental group, there were 5 male and 13 female participants whose ages ranged from 19 to 23. On the other hand, there were 9 male and 9 female participants in the control group, and their ages were between 19 and 22. While the experimental group was given indirect written corrective feedback through error correction codes, the control group was given direct written corrective feedback.

As mentioned above, two intact groups of learners participated in the study. However, the intact groups were formed by the institution based on the results of the Oxford Placement Test (Oxford University Press), and according to the results, all participants in both groups were at the A2 level according to the Common European Framework of Reference for Languages. To see if the test scores differed across groups, a t-test was run, considering that the residuals of the placement test data were

normally distributed (*SW* = .044, *df* = 36, *p* > .05) as suggested by Kéry and Hatfield (2003) and Field (2018). The results of the t-test are presented below.

Table 2.

Placement Test Comparison

Group	*n*	*M*	*SD*	*T*	*df*	*p*
1	18	28.44	5.27	-1.48	34	.148
2	18	31.22	5.97			

As seen in the table, the mean placement score of control group was 28.44 (*SD* = 5.27) and that of experimental group was 31.22 (*SD* = 5.97). The t-test results showed that the difference between the mean placement scores of the feedback groups was not statistically significant (*t* = -1.48, *df* = 34, *p* > .05).

Table 3.

Comparison of Pre-test Scores

Variable	Group	*n*	*M*	*SD*	*Mdn*	*t/Z*	*df*	*P*
Task Achievement*	1	18	2.44	0.68	2.50	-1.454	-	.171
	2	18	2.75	0.35	2.75			
Coherence & Cohesion	1	18	2.56	0.48	2.50	-0.297	34	.768
	2	18	2.61	0.63	2.75			
Vocabulary	1	18	3.08	0.67	3.00	-0.147	34	.884
	2	18	3.11	0.47	3.00			
Grammar	1	18	2.69	0.84	2.50	-0.071	34	.716
	2	18	2.61	0.47	2.50			
Total	1	18	10.78	1.90	11.00	-0.361	34	.605
	2	18	11.08	1.60	11.00			

*: Mann-Whitney U (Z Score Reported)

As seen in the table, no statistically significant difference was observed among the feedback groups in terms of the pre-test scores (*p* > .05).

3.4. Data Collection Tools and Procedure

3.4.1. Data Collection Tools

In this study, pre and post-tests were administered to collect the quantitative data. In addition to these, a survey including open-ended questions, semi-structured interviews were included in the qualitative data collection procedure.

As the first data collection instrument, the participants in both experimental and control groups were given a pre-test **(see Appendix 1)** to measure their existing writing performance. The same test was later administered to all participants in both groups as a post-test.

After the treatment, the pre-test was implemented once more as a post-test in Week 6. A writing rubric developed by the institution **(see Appendix 2)** was used for the evaluation of the participants' writing papers. The writing rubric included four main categories: Task Achievement; Coherence / Cohesion; Range and Accuracy of Vocabulary for Purpose; Range and Accuracy of Grammar for Purpose. Each category in the rubric was graded out of 5 points, which made a total score of 20 points. The total number of points was then converted into a score out of 100. The pre-tests and post-tests were assessed twice, first by the researcher and then by another colleague to make sure that the results were accurate and consistent. Therefore, it can be suggested that the scoring rubric helped reduce subjectivity (Moscal & Leydens, 2000).

In addition, a second immediate post-test **(see Appendix 3)** that the participants were asked to write on a different topic was implemented. The aim was to collect the data that allows the researcher to compare the two post-test results and see the short-term transfer effects of corrective feedback on a new piece of writing.

As mentioned above, all the scores obtained by means of the rubric were scored by two different raters to ensure reliability. According to Larsen-Hall (2016), Cronbach's Alpha reliability coefficient should be used in cases where the aim is to see if data coming from different judges agree or not. For this reason, total scores and

each analytical item within the rubric were matched for Rater 1 and Rater 2 to calculate a Cronbach's Alpha value for each pair. The results are given below.

Table 4.

Interrater Reliability Analyses

Pretest	$\alpha_{Pretest}$	$\alpha_{Posttest}$	$\alpha_{Posttest2}$
Task Achievement	.915	.946	.956
Coherence & Cohesion	.887	.929	.959
Vocabulary	.952	.904	.923
Grammar	.941	.939	.950
Total	.955	.971	.978

As shown in the table, the alpha values obtained in the analysis were between .89 and .98, indicating reliable scoring for all parts of the rubric as well as the total scores.

Thirdly, an online survey **(see Appendix 4)** which included open-ended questions was carried out in the experimental group. In addition, another survey including similar questions **(see Appendix 5)** was carried out in the control group. It was aimed at learning the participants' opinions about each feedback type.

Furthermore, six participants from each group were interviewed to gather more detailed data about the participants' opinions about the procedure in all its aspects. Semi-structured questions **(see Appendix 6 and Appendix 7)** were used to interview the participants.

The survey and interview questions were created by the researcher, and the questions were consulted with two experts in the field.

3.4.2. Data Collection Procedure

3.4.2.1. Content of Reading and Writing Lessons

In this study, the participants in both groups had reading lessons in addition to writing lessons. Each week, a total of six hours were allocated for reading and writing lessons. The participants read two reading texts relating to the topic on which they were supposed to write each week. *Q-Skills for Success Reading and Writing 2 (3rd ed.) by Oxford University Press* was used for the reading lessons. The writing topics were therefore created in accordance with the themes of the reading book's units. As Hirvela (2004) emphasized, reading is closely connected to writing as it provides a rich source of input in terms of the organizational pattern, language, and style, all of which help generate a written text. In this study, it was aimed to enrich students' knowledge of vocabulary through the reading texts. Besides, it was a stage that started students' thinking processes on the related writing topic. In addition to reading texts, supplementary writing materials were used to teach how to form an opinion paragraph and draw attention to the important points to consider while writing. During the treatment, only the opinion paragraph genre was addressed.

3.4.2.2. Procedure

As this quasi-experimental study was conducted during the COVID-19 pandemic, the lessons were carried out with the help of online platforms. Google Classroom, which allows the use of Google documents, was used to assign and collect the weekly writing tasks. The same platform was also used to provide feedback for the writing tasks.

The present study was completed in seven weeks. In the first week, the participants in both groups were given the pre-test to see their initial writing performance. The participants were asked to write an opinion paragraph on one of the topics provided (see Appendix 1 for the writing topics). The writing topics were assigned by the researcher. When selecting writing subjects, it was important to keep in mind that the topics should not be too specialized or require much research, but they

should be ones about which all of the participants may have some ideas. The purpose of the pre-test was to determine if the participants could form an opinion paragraph by first expressing their opinions with a topic sentence and then supporting their ideas with details and examples, as well as provide coherence and cohesion in the paragraph by using a variety of cohesive devices such as linking words, sequencers, referencing, and so on. Subsequently, papers from both groups were scored by the researcher and another colleague, and were recorded. The participants were not provided with the feedback immediately after the pre-test was conducted in order to familiarize them with the feedback procedure and the Error Correction Codes.

Moreover, in Week 1, the Error Correction Codes **(see Appendix 8)** were introduced to the experimental group, and the participants were informed about how the feedback procedure was going to take place. As the use of Error Correction Codes as a way of providing written corrective feedback was completely new for the participants, the instructor elaborated on what each code referred to in the following weeks to make sure that they were all clearly understood. In the same week, after the completion of reading and writing lessons, the first writing task was assigned to both groups, and the participants completed the task during the lesson under the supervision of the instructor.

In Week 2, the instructor returned the students' tasks from Week 1. The experimental group had indirect feedback through codes, and their errors were also highlighted in the text, as illustrated by (1). On the other hand, the control group had direct feedback and their errors were explicitly corrected by the instructor, as illustrated by (2).

(1) I think before the pandemic, we are **(VT)** so free and fearless.

(2) Before the pandemic**,** nobody ~~wears~~ **wore /would wear** masks because we ~~don't need to~~ **didn't need to**.

In addition to the feedback for grammatical and lexical errors, the instructor provided some overall feedback for organizational aspects of writing like task achievement and coherence and cohesion in the margins or at the bottom of the text. The participants in both groups were given feedback on paragraph structure, expanding supporting points, or building coherence and cohesion, as illustrated by (3), (4), and (5).

(3) Before going to the job interview, firstly you should research about the company.
You should start your paragraph with a topic sentence.

(4) For example technological inventions make people lazy. *Why and how? You should expand your idea by giving examples.*

(5) *Well- done! It is good to see that you expand your ideas by giving some specific examples. And you use linking words effectively. Please check the use of 'because' again.*

The participants in both groups were then asked to review their papers with corrections, and they were given 30 minutes to write their second drafts during the lesson.

At the beginning of the study, the experimental group had difficulty correcting their errors by themselves, and they needed guidance, especially for their lexical errors. Therefore, both experimental and control groups were taught how to use online dictionaries to find out the correct forms of the words and some collocations.

In the following weeks, the same procedure for each writing task was followed. In total, the participants submitted six writing tasks, except for the pre and post-tests. For all the tasks, they were required to write the first and second drafts. The instructor administered the same instructional materials and conditions to both groups.

In Week 6, the first post-test was carried out. The participants were returned their pre-tests with the feedback provided by the researcher. And they were asked to review their papers with corrections, and write another opinion paragraph on the same topic as the pre-test that they took in the first week.

In addition, in the final week, the participants took a writing exam, and the writing exam results were used as the second post-test with a different topic (see Appendix 3 for the writing topics). Both post-tests were carried out under the supervision of the instructor. The participants' activities during the exam were logged, and before the evaluation of the writing exam, the authenticity of the students' papers was checked through *Turnitin*, a plagiarism detection software.

In addition to post-tests, in Week 7, online surveys and interviews were conducted by the researcher in both groups. Semi-structured interviews with six participants from each group were done via Zoom, an online meeting platform, and each interview took about fifteen minutes. The surveys and interviews presented valuable data to learn the participants' opinions about each feedback type and the procedure.

Table 5.

Research Procedure for Experimental and Control Group

Week	Experimental Group Treatment	Control Group Treatment
Week 1	Pre-test (45 minutes)	Pre-test (45 minutes)
	Unit theme: Marketing	Unit theme: Marketing
	Unit question: Why does something become popular?	Unit Question: Why does something become popular?
	Reading 1: Unusual Ideas to Make a Buzz	Reading 1: Unusual Ideas to Make a Buzz
	Reading 2: How do you decide?	Reading 2: How do you decide?

	Writing Skill: Identifying parts of a paragraph, Paragraph Unity	Writing Skill: Identifying parts of a paragraph, Paragraph Unity
	Vocabulary Skill: Word families	Vocabulary Skill: Word families
	Writing Task 1: Write an opinion paragraph on "Why is social media popular?"	Writing Task 1: Write an opinion paragraph on "Why is social media popular?"
	The introduction of Error Correction Codes	Collecting the first drafts of the first writing task
	Collecting the first drafts of the first writing task	
Week 2	Unit Theme: Social Psychology	Unit Theme: Social Psychology
	Unit Question: What does it mean to be polite?	Unit Question: What does it mean to be polite?
	Reading 1: Being Polite from Culture to Culture	Reading 1: Being Polite from Culture to Culture
	Reading 2: Answers to All Your Travel Questions	Reading 2: Answers to All Your Travel Questions
	Writing Skill: How to write a topic sentence and supporting sentences	Writing Skill: How to write a topic sentence and supporting sentences
	Vocabulary Skill: Suffixes and Prefixes	Vocabulary Skill: Suffixes and Prefixes
	Reviewing the Error Correction Codes	Writing the second draft of task 1
	Writing the second draft of task 1	Writing Task 2: Write and opinion paragraph on "How do you make a good impression before a job interview?"
	Writing Task 2: Write an opinion paragraph on "How do you make a good impression before a job interview?"	Collecting the first drafts of the second writing task
	Collecting the first drafts of the second writing task	
Week 3	Unit Theme: Technology	Unit Theme: Technology

	Unit question: How can technology improve performance? Reading 1: Virtual Reality for Medical Students Reading 2 : The Technology Advantage Writing Skill: Linking words and phrases Vocabulary Skill: Using the dictionary Reviewing the Error Correction Codes Writing the second draft of task 2 Writing Task 3: Write an opinion paragraph on "Technological inventions are not always good for human beings. Do you agree?"	Unit question: How can technology improve performance? Reading 1: Virtual Reality for Medical Students Reading 2 : The Technology Advantage Writing Skill: Linking words and phrases Vocabulary Skill: Using the dictionary Writing the second draft of task 2 Writing Task 3: Write an opinion paragraph on "Technological inventions are not always good for human beings. Do you agree?"
Week 4	Unit theme: Business Unit question: What makes a family business successful? Reading 1: A Successful Family Business Reading 2: The Challenge of Running a Family Business Writing Skill: Avoiding Irrelevant Sentences and Writing Concluding Sentence Vocabulary Skill: Using the dictionary Writing the second draft of task 3 Writing Task 4: "Having your own business is advantageous. Do you agree?"	Unit theme: Business Unit question: What makes a family business successful? Reading 1: A Successful Family Business Reading 2: The Challenge of Running a Family Business Writing Skill: Avoiding Irrelevant Sentences and Writing Concluding Sentence Vocabulary Skill: Using the dictionary Writing the second draft of task 3 Writing Task 4: "Having your own business is advantageous. Do you agree?"

Week 5	Unit theme: Brain Science Unit Question: How can you learn faster and better? Reading 1: You Can Read Faster and Better Reading 2: Brain Secrets of the Most Successful Students Writing Skill: Revision Vocabulary Skill: Collocations Writing the second draft of task 4 Writing Task 5: "People learn faster with the Internet. Do you agree?"	Unit theme: Brain Science Unit Question: How can you learn faster and better? Reading 1: You Can Read Faster and Better Reading 2: Brain Secrets of the Most Successful Students Writing Skill: Revision Vocabulary Skill: Collocations Writing the second draft of task 4 Writing Task 5: "People learn faster with the Internet. Do you agree?"
Week 6	Unit Theme: Public Health Unit Question: How can we prevent diseases? Reading 1: University Health Centre: Cold News Reading 2: Preventing Disease Around the World Writing Skill: Revision Vocabulary Skill: Collocations Writing the second draft of task 5 Post-test 1 (Revision of pre-test) Writing Task 6: "People should stop eating ready-made food to prevent illnesses. Do you agree?"	Unit Theme: Public Health Unit Question: How can we prevent diseases? Reading 1: University Health Centre: Cold News Reading 2: Preventing Disease Around the World Writing Skill: Revision Vocabulary Skill: Collocations Writing the second draft of task 5 Post-test 1 (Revision of pre-test) Writing Task 6: "People should stop eating ready-made food to prevent illnesses. Do you agree?"

Week 7	Writing the second draft of task 6 Post-test 2 (new piece of writing) Survey Interviews	Writing the second draft of task 6 Post-test 2 (new piece of writing) Survey Interviews

3.5. Data Analysis and Interpretation

3.5.1. Quantitative Data Analysis

For data analysis, the assumptions of each parametric test were checked initially. However, a number of studies show that regression models such as ANOVA's, ANCOVA's or t-tests are robust to minor violations of those assumptions (e.g. Blanca, Alarcón, Arnau, Bono & Bendayan, 2017; Oljenik & Algina, 1984; Shielzeth et al., 2020). Even so, parametric results were reported along with their nonparametric counterparts in the case of assumption violations to avoid erroneous conclusions as recommended by Mizumoto and Takeuchi (2018). Group 1 represents the control group, and Group 2 represents the experimental group.

3.5.1.1 Whole-Group Comparisons

Comparing the scores related to the whole group regardless of the feedback type they received necessitated a paired-samples t-test (i.e. dependent t-test). A paired-samples t-test assumes that the differences between pre and post scores, not the scores themselves, are normally distributed (Field, 2018). Therefore, the differences were checked for distribution normality initially.

Table 6.

Normality Tests for Differences in the Paired Comparisons

Score	*SW*	*df*	*p*
Total	.969	36	.387
Task Achievement	.925	36	**.018**
Coherence & Cohesion	.920	36	**.012**
Vocabulary	.929	36	**.023**
Grammar	.950	36	**.000**

As seen in the Table, Shapiro-Wilk Tests indicated that the paired differences were normally distributed only in the total scores ($p > .05$). In Task Achievement, Coherence & Cohesion, Vocabulary and Grammar, the distributions of the differences were seen to be non-normal ($p < .05$). To check the severity of deviations from normality, Q-Q plots (Figures 1, 2, 3, 4) were investigated.

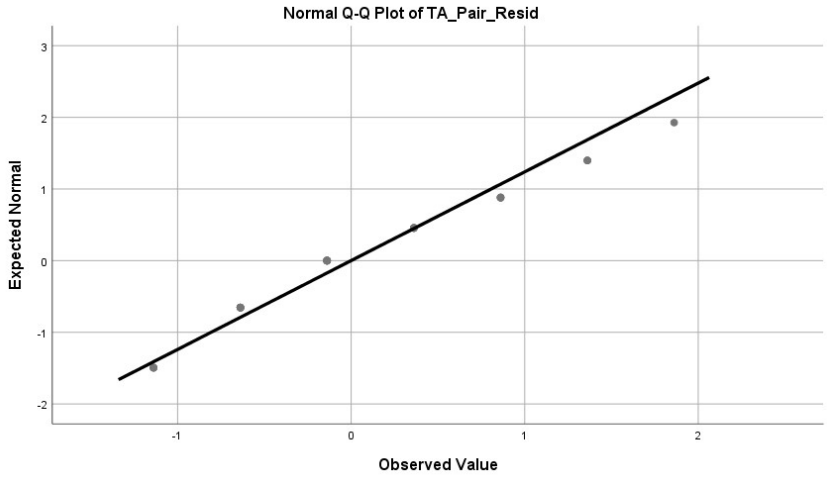

Figure 4. Paired Residuals of Task Achievement

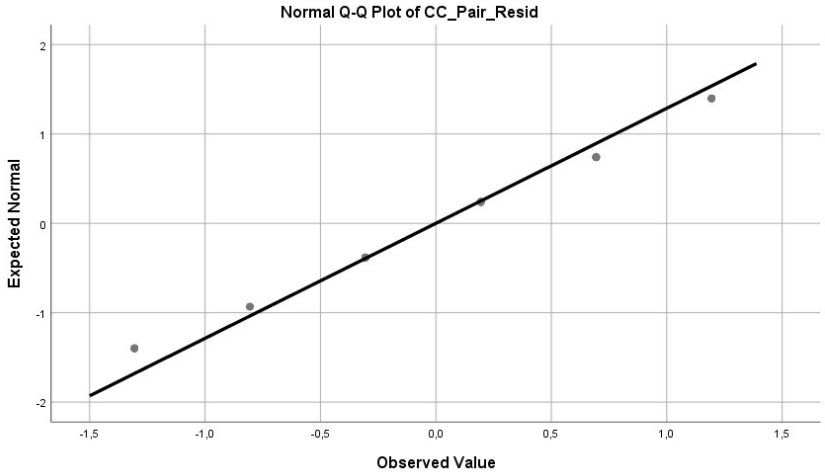

Figure 5. Paired Residuals of Coherence & Cohesion

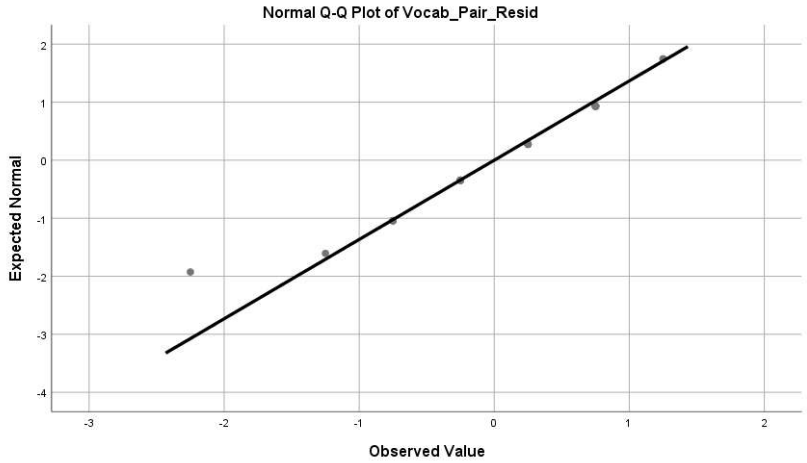

Figure 6. Paired Residuals of Vocabulary

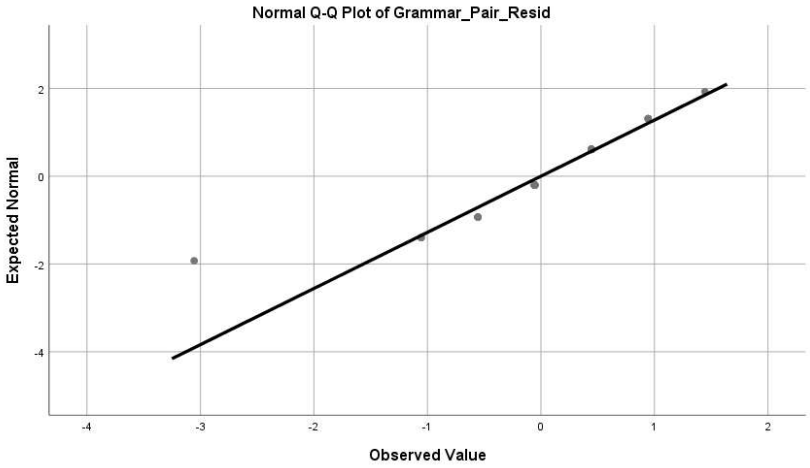

Figure 7. Paired Residuals of Grammar

As seen in the figures, the distributions actually approximated normality, which was sufficient for most parametric analyses. However, a visible outlier was present in Vocabulary and Grammar scores. Although removing outliers is common practice in educational science and applied linguistics, Nicklin and Plonsky (2020) warn against "blindly removing" (p. 26) them unless there is a confirmable error in measurement since such practice results in the loss of legitimate data even though winsorizing and log e transformation produce valid results without any loss of legitimate data points. Also taking into account the robustness of regression models to deviations from normality, parametric results were reported and compared with their nonparametric alternatives for those variables as suggested by Mizumoto and Takeuchi (2018) without any data removal since there was no evidence that the outlier was due to erroneous measurement

3.5.1.2. Comparisons of Feedback Groups

Since the comparison of the pre-test and post-test scores for each written corrective feedback group necessitated an ANCOVA model, the assumptions of ANCOVA, namely approximate normality in residuals, variance homogeneity, homogeneity of regression slopes, linearity and absence of a difference in the pretest scores were initially tested (Field, 2018; Kéry & Hatfield).

Normality of Residual Distributions

Due to its particular power in smaller samples (e.g. $n > 50$), Shapiro-Wilk tests were run to test the residual distribution normality assumption (Larson-Hall, 2016; Ricci, 2005). The results are presented below in different tables due to the large number of variables investigated.

Table 7.

Normality Tests for Residuals in the ANCOVA Model

Score	*SW*	*df*	*p*
Task Achievement	.895	36	**.003**
Coherence & Cohesion	.953	36	.130
Vocabulary	.960	36	.219
Grammar	.973	36	.519
Total	.960	36	.220

As shown in the table, the residuals of the Coherence & Cohesion, Vocabulary, Grammar and Total scores were normally distributed in the ANCOVA model ($p > .05$). However, the residuals of Task Achievement were seen to be non-normally distributed ($p < .05$). Nonetheless, an investigation of the Q-Q plots for the residuals of Task Achievement showed that the distribution approximated normality (Figure 5), which met the requirement for an ANCOVA model. Even so, the analysis regarding Task Achievement were performed both parametrically and non-parametrically in order to check if the results confirmed one another.

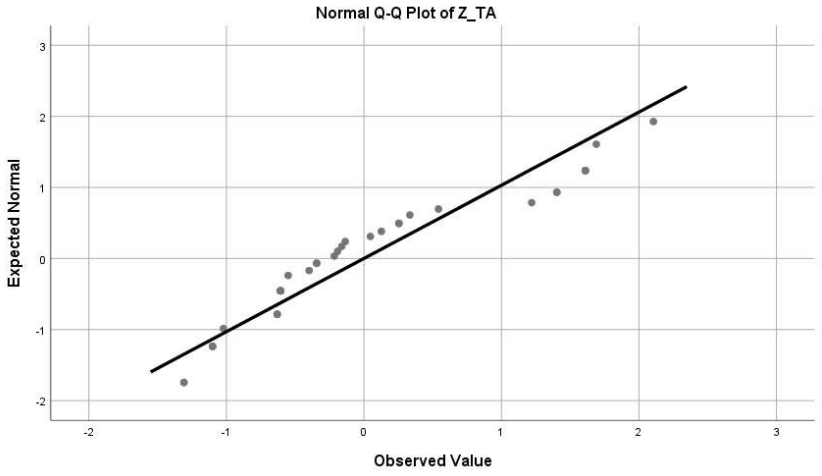

Figure 8. Residual Distribution of Task Achievement

Variance Homogeneity

The *F* test results regarding the assumption of variance homogeneity are given below.

Table 8.

Variance Homogeneity Tests

Score	*F*	df_1	df_2	*p*
Task Achievement	1.776	1	34	.191
Coherence & Cohesion	1.012	1	34	.321
Vocabulary	2.271	1	34	.141
Grammar	.233	1	34	.632
Total	.725	1	34	.400

As seen in the tables, all the *F* tests produced non-significant *p* values, indicating variance homogeneity in all ANCOVA models tested in this thesis ($p > .05$).

Homogeneity of Regression Slopes

According to Larsen-Hall (2016), one of the ways to test the homogeneity of regression slopes in ANCOVA models is to test the interaction effect between the grouping variable and the covariate. The results of the interaction tests are presented below:

Table 9.

Regression Slopes Homogeneity Tests

Score	*F*	*df*	*p*
Task Achievement	0.253	1	.504
Coherence & Cohesion	0.666	1	.268
Vocabulary	0.019	1	.829
Grammar	0.166	1	.519
Total	0.624	1	.731

According to the results, the interaction effects between the grouping variable and each covariate were not statistically significant, indicating homogeneity of regression slopes for all ANCOVA models built in this study ($p > .05$).

Linearity

Deviations from linearity between covariates and dependent variables were also tested to see if they met the assumptions of ANCOVA. The results are tabulated below:

Table 10.

Linearity Tests

Score	*F*	*df*	*p*
Task Achievement	0.687	4	.606
Coherence & Cohesion	0.436	4	.729
Vocabulary	0.710	4	.098
Grammar	3.216	4	**.026**
Total	0.382	4	.820

As seen in the results, only the Grammar scores significantly deviated from linearity in terms of the relationship between the dependent variable and the covariate ($p < .05$) For this reason, non-parametric analyses were also conducted for this variable, too.

Absence of Differences between Pre-test Scores

To test the last assumption of ANCOVA, the pre-test scores of the written corrective feedback groups were compared using t-tests or Mann-Whitney U tests depending on the residual distribution of the variables. The results are tabulated below.

Table 11.

Comparison of Pre-test Scores

Variable	Group	*n*	*M*	*SD*	*Mdn*	*t/Z*	*df*	*P*
Task Achievement*	1	18	2.44	0.68	2.50	-1.454	-	.171
	2	18	2.75	0.35	2.75			
Coherence & Cohesion	1	18	2.56	0.48	2.50	-0.297	34	.768
	2	18	2.61	0.63	2.75			
Vocabulary	1	18	3.08	0.67	3.00	-0.147	34	.884
	2	18	3.11	0.47	3.00			
Grammar	1	18	2.69	0.84	2.50	-0.071	34	.716
	2	18	2.61	0.47	2.50			
Total	1	18	10.78	1.90	11.00	-0.361	34	.605
	2	18	11.08	1.60	11.00			

*: Mann-Whitney U (Z Score Reported)

As seen in the table, no statistically significant difference was observed among the feedback groups in terms of the pretest scores ($p > .05$).

Based on the assumption tests, it was seen that ANCOVA models could be used to test the effectiveness of the written corrective feedback intervention. However, since Task Achievement had a slight deviation from normality and Grammar had a slight deviation from linearity, these variables were analysed in both parametric and non-parametric terms to see if the presence of outliers in the data sets caused different results. Even though ANCOVA models are generally robust to minor deviations from normality as well as linearity and Type I error rate is not seriously affected in minor deviations (Oljenik & Algina, 1984), nonparametric ANCOVA models are known to be stronger than parametric ANCOVA models in serious deviations, producing fewer Type I errors (Rheinheimer & Penfield, 2001). For this reason, the Task Achievement and Grammar variables were analysed by both ANCOVA and Quade's Test, which is one of the non-parametric alternatives to ANCOVA (Cangür, Şungur & Ankaralı, 2018; Oljenik & Algina, 1985).

3.5.1.3. Comparisons of Two Post-tests

In order to compare the first post-test and the second one with a different topic, their differences were initially computed and checked for normality as suggested by Field (2018). The results are given below:

Table 12.

Normality Tests for the Differences in the Paired Post-test Comparisons

Score	*SW*	*df*	*p*
Total	.945	36	.072
Task Achievement	.896	36	**.003**
Coherence & Cohesion	.902	36	**.004**
Vocabulary	.886	36	**.001**
Grammar	.938	36	**.045**

As seen in the table, only the total score residuals were normally distributed in the paired posttest comparisons ($p > .05$). The residuals of Task Achievement, Coherence & Cohesion, Vocabulary and Grammar scores were non-normally distributed ($p < .05$). The severity of deviations from normality for those variables were checked on their Q-Q plots (Figures 6, 7, 8, 9).

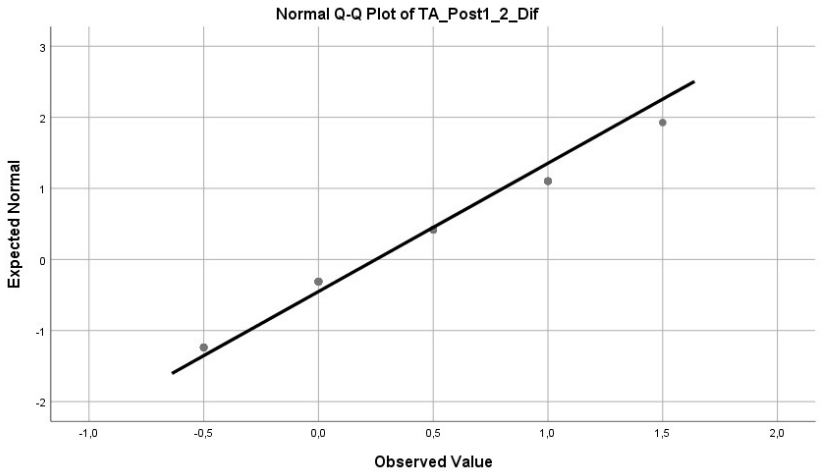

Figure 9. Paired Residuals of Task Achievement Post-tests

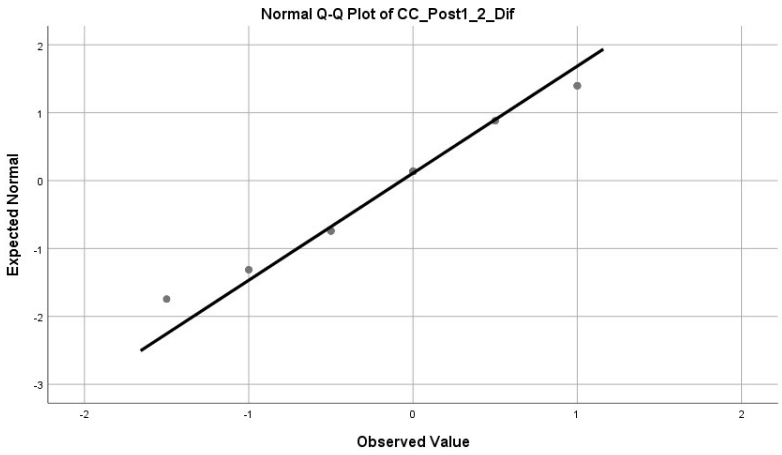

Figure 10. Paired Residuals of Coherence & Cohesion Post-tests

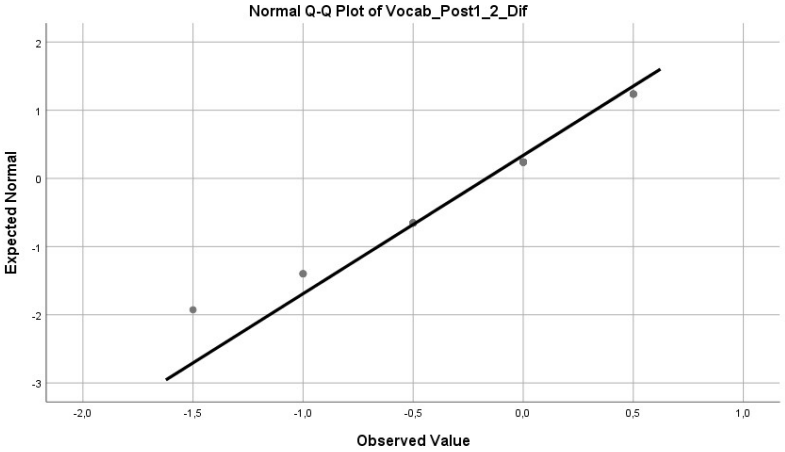

Figure 11. Paired Residuals of Vocabulary Post-tests

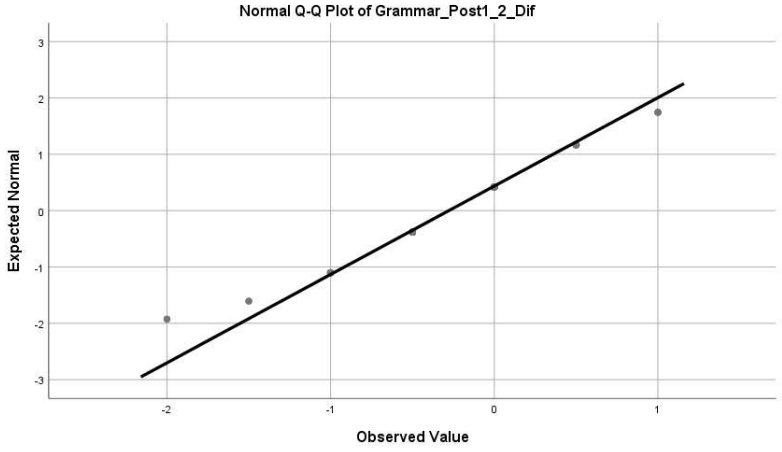

Figure 12. Paired Residuals of Grammar Post-tests

The investigation of the Q-Q plots, as seen in the figures, showed the distributions in the post-test residuals approximated normality. Despite the existence of a few outliers in the data, they were kept as is due to the absence of evidence regarding measurement error (Nicklin & Plonsky, 2020). Therefore, paired samples t-tests were run on the data and reported along with Wilcoxon Signed Rank Test results, or their non-parametric alternatives (Field, 2018; Mizumoto & Takeuchi, 2018).

3.5.2. Qualitative Data Analysis

To analyse the survey data, content analysis, which allows researchers to examine the written contents of a communication, was preferred (Fraenkel et al., 2012). A content analysis was used in this research in order to formulate the major ideas as to the corrective feedback types, and to test the research hypothesis.

A content analysis was conducted on the written accounts of students. Various categories were determined based on the participants' answers, and then the references made to the related categories by each participant were counted. To ensure reliability, the written accounts were analysed again by the researcher a month later. Besides, the content of the survey question "How do you describe the feedback procedure with three adjectives?" was manifest, and therefore no inferences were necessary regarding the underlying meaning.

Another qualitative data collection tool was the interview, including semi-structured questions. The answers were presented as direct quotations to enrich the survey data.

3.6. Findings

The results and interpretation regarding the analysis of the data obtained from the participants were presented in this part.

3.6.1. Findings of the First Question

The results of the pretest and posttest were compared to answer for the first research question, (*RQ 1: Is indirect written corrective feedback effective in improving students' overall writing performance?)*. The pretest and posttest scores and the comparison of post-test scores are presented below.

3.6.1.1. Pre-test and Post-test Comparisons of the Whole Group

The pre-test and post-test results of the whole group of participants, regardless of the type of feedback they received, were compared initially. The results are presented below.

Table 13.

Pre-test and Post-test Comparisons of the Whole Group (N = 36)

Variable	$M_{pretest}$	$SD_{pretest}$	$M_{posttest}$	$SD_{posttest}$	*t*	*df*	*p*	*D*
Total	10.93	1.74	15.18	2.46	-11.039	35	**.000**	2.00
Task Achievement	2.60	0.56	3.74	0.75	-8.465	35	**.000**	1.72
Coherence & Cohesion	2.58	0.55	3.89	0.75	-10.080	35	**.000**	1.99
Vocabulary	3.10	0.56	3.85	0.62	-6.148	35	**.000**	1.27
Grammar	2.65	0.67	3.71	0.65	-8.102	35	**.000**	1.61

According to the paired samples t-test results, there were statistically significant differences with large effects in all pretest and posttest results ($p < .001$), $d > 0.80$). Since the differences between the pretest and posttest scores of Task Achievement ($Z = -4{,}968$, $p < .001$, $r = 0.59$), Coherence & Cohesion ($Z = -4{,}899$, $p < .001$, $r = 0.58$), Vocabulary ($Z = -4{,}531$, $p < .001$, $r = 0.53$) and Grammar ($Z = -4{,}531$, $p < .001$, $r = 0.53$) scores slightly deviated from normality, Wilcoxon Signed Rank Tests were also run on these variables and it was seen that the non-parametric

results were in complete agreement with the parametric ones producing statistically significant results with large effects. Therefore, it was concluded that the scores related to all variables significantly increased from the pretest to the posttest.

3.6.1.2. Pre-test and Post-test Results Based on Feedback Groups

The participants were also divided by their respective feedback groups and compared by means of ANOVA's and Quade's tests to see if one group outperformed the other or not. The descriptive results regarding the scores divided by feedback groups are presented below.

Table 14.

Pre-test and Post-test Results Based on Feedback Groups

Variable	Group	*N*	$M_{pretest}$	$SD_{pretest}$	$Mdn_{pretest}$	$M_{posttest}$	$SD_{posttest}$	$Mdn_{posttest}$
Total	1	18	10.78	1.90	11.00	14.75	2.49	13.50
	2	18	11.08	1.60	11.00	15.61	2.41	15.50
Task Achievement	1	18	2.44	0.68	2.50	3.58	0.69	3.50
	2	18	2.75	0.35	2.75	3.89	0.80	3.75
Coherence & Cohesion	1	18	2.56	0.48	2.50	3.81	0.75	4.00
	2	18	2.61	0.63	2.75	3.97	0.76	4.00
Vocabulary	1	18	3.08	0.65	3.00	3.78	0.71	3.50
	2	18	3.11	0.47	3.00	3.92	0.52	4.00
Grammar	1	18	2.69	0.84	2.50	3.58	0.62	3.50
	2	18	2.61	0.47	2.50	3.83	0.66	4.00

An investigation of the pre-test scores divided by feedback groups showed that slight differences were present in the pre-test scores related to all the variables studied. However, it was previously confirmed in Table 11 as a part of ANCOVA assumption checks that none of those differences were statistically significant ($p > .05$). Similarly, slight differences were visible in the post-test scores of two feedback groups. Descriptive results showed that the mean post-test scores of Group 2 were slightly

higher than those of Group 1 in all variables, by 0.86 point in Total scores, 0.31 point in Task Achievement, 0.16 point in Coherence & Cohesion, 0.14 point in Vocabulary and 0.25 point in Grammar. In order to see if these differences were statistically significant, covariance analyses were run for each of the variables. The results are given below in their separate respective tables.

Table 15.

Group Comparison for Total Scores (ANCOVA)

Source	*Type III SS*	*df*	*MS*	*F*	*p*	η_p^2
Corrected Model	44.069	2	22.035	4.354	**.021**	0.21
Intercept	64.020	1	64.020	12.65	**.001**	0.28
Pretest	37.396	1	37.396	7.389	**.010**	0.18
Feedback Group	4.111	1	4.111	0.812	.374	0.02
Error	167.007	33	5.061			
Total	8507.25	36				
Corrected Total	211.076	35				

R^2 = .21, Adjusted R^2 = .16

ANCOVA results indicated that the Total pre-test scores had a statistically significant effect on the total post-test scores with a large effect ($F_{(1, 33)} = 7.389$, $p < .05$, $\eta_p^2 = 0.18$). When the total pre-test scores were controlled for, it was seen that there was no difference in the Total post-test scores of Group 1 and Group 2 ($F_{(1, 33)} = 0.812$, $p > .05$, $\eta_p^2 = 0.02$).

Table 16.

Group Comparison for Task Achievement Scores (ANCOVA)

Source	*Type III SS*	*df*	*MS*	*F*	*p*	η_p^2
Corrected Model	1.784	2	0.892	1.639	.209	0.09
Intercept	12.396	1	12.396	22.778	**.000**	0.41
Pretest	0.944	1	0.944	1.735	.197	0.05
Feedback Group	0.373	1	0.373	0.686	.413	0.02
Error	17.959	33	0.544			
Total	522.25	36				
Corrected Total	19.743	35				

R^2 = .09, Adjusted R^2 = .04

In Task Achievement, it was seen that the pre-test did not have a statistically significant effect on the post-test scores ($F_{(1, 33)} = 1.735$, $p > .05$, $\eta_p^2 = 0.05$). The post-test scores of the feedback groups were not found to be significantly different, either ($F_{(1, 33)} = 0.686$, $p > .05$, $\eta_p^2 = 0.02$). Since the residuals of Task Achievement scores slightly deviated from normality, a Quade's Test was also run for confirmatory purposes and it was seen that nonparametric results confirmed their parametric counterparts regarding the absence of a statistically significant difference between the Task Achievement post-test scores of Group 1 and Group 2 ($F_{(1, 34)} = 0.548$, $p = .464$).

Table 17.

Group Comparison for Coherence & Cohesion Scores (ANCOVA)

Source	*Type III SS*	*Df*	*MS*	*F*	*p*	η_p^2
Corrected Model	2.139	2	1.069	2.026	.148	0.11
Intercept	12.099	1	12.099	22.923	**.000**	0.41
Pretest	1.889	1	1.889	3.578	.067	0.10
Feedback Group	0.184	1	0.184	0.350	.558	0.01
Error	17.417	33	0.528			
Total	564.000	36				
Corrected Total	19.556	35				

$R^2 = .11$, Adjusted $R^2 = .06$

The analysis of Coherence & Cohesion scores revealed that the pre-test scores did not have a significant effect on the post-test scores ($F_{(1, 33)} = 3.578$, $p > .05$, $\eta_p^2 = 0.10$). The post-test scores did not significantly differ between feedback groups, either ($F_{(1, 33)} = 0.350$, $p > .05$, $\eta_p^2 = 0.01$).

Table 18.

Group Comparison for Vocabulary Scores (ANCOVA)

Source	*Type III SS*	*Df*	*MS*	*F*	*p*	η_p^2
Corrected Model	0.867	2	0.434	1.141	.332	0.07
Intercept	10.357	1	10.357	27.25	**.000**	0.45
Pretest	0.694	1	0.694	1.825	.186	0.05
Feedback Group	0.156	1	0.156	0.412	.526	0.01
Error	12.542	33	0.380			
Total	546.25	36				
Corrected Total	13.41	35				

R^2 = .07, Adjusted R^2 = .01

According to the results, the effect of the Vocabulary pre-test scores on the post-test scores were not statistically significant ($F_{(1, 33)} = 1.825$, $p > .05$, $\eta_p^2 = 0.05$). Furthermore, the differences in the Vocabulary post-test scores were not significant between the feedback groups ($F_{(1, 33)} = 0.412$, $p > .05$, $\eta_p^2 = 0.01$).

Table 19.

Group Comparison for Grammar Scores (ANCOVA)

Source	*Type III SS*	*df*	*MS*	*F*	*p*	η_p^2
Corrected Model	2.011	2	1.006	2.618	.088	0.14
Intercept	17.900	1	17.900	46.600	**.000**	0.59
Pretest	1.449	1	1.449	3.772	.061	0.10
Feedback Group	0.679	1	0.679	1.767	.193	0.05
Error	12.676	33	0.384			
Total	509.750	36				
Corrected Total	14.687	35				

R^2 = .14, Adjusted R^2 = .09

In Grammar scores, it was seen that the pre-test did not have a statistically significant effect on the post-test ($F_{(1, 33)} = 3.772$, $p > .05$, $\eta_p^2 = 0.10$). The post-test scores did not differ according to the feedback group after controlling for the pre-test, either ($F_{(1, 33)} = 1.767$, $p > .05$, $\eta_p^2 = 0.05$). Since the Grammar scores deviated from linearity, the results were compared with the Quade's Test results, which also

confirmed that there was no statistically significant difference between Group 1 and Group 2 in the Grammar post-test scores ($F_{(1, 34)} = 1.370$, $p = .250$).

3.6.1.3. Comparisons of Post-test Scores

Apart from the post-test which was identical to the pre-test, a second post-test which were of the same genre with a different writing prompt was also collected from the participants. The pairwise comparisons related to these two post-tests are presented below. Both parametric and nonparametric pairwise comparison results were reported for Task Achievement, Coherence & Cohesion, Vocabulary and Grammar due to the deviations from normality in their pre-test and post-test differences.

Table 20.

Post-test 1 and Post-test 2 Comparisons (N = 36)

Variable	$M_{posttest1}$	$SD_{posttest1}$	$M_{posttest2}$	$SD_{posttest2}$	t	df	p	D
Total	15.18	2.46	14.92	2.29	0.940	35	.354	0.11
Task Achievement	3.74	0.75	3.99	0.74	-2.707	35	**.010**	0.34
Coherence & Cohesion	3.89	0.75	3.82	0.72	0.657	35	.515	0.10
Vocabulary	3.85	0.62	3.68	0.51	2.029	35	.050	0.30
Grammar	3.70	0.65	3.43	0.62	2.615	35	**.013**	0.43

As seen in the table, the first and the second post-tests were not significantly different in terms of the mean Total scores ($t = 0.940$, $df = 35$, $p > .05$) and the mean scores of Coherence & Cohesion ($t = 0.657$, $df = 35$, $p > .05$, $Z = -0.459$, $p = .646$, $r = 0.05$) and Vocabulary ($t = 2.029$, $df = 35$, $p = .050$, $Z = -1.949$, $p = .051$, $r = 0.23$). On the other hand, Task Achievement ($t = -2.707$, $df = 35$, $p = .01$, $Z = -2.588$, $p = .010$, $r = 0.31$) and Grammar ($t = 2.615$, $df = 35$, $p < .05$, $Z = -2.023$, $p = .043$, $r = 0.24$) scores were significantly different in the first and the second post-tests. It was seen that both parametric and nonparametric analyses yielded the same conclusions regarding the pairwise comparisons.

3.6.2. Findings of the Second Question

Content analysis based on the survey questions and answers was conducted in order to answer the second research question, (*RQ 2: What are the participants' opinions about direct and indirect written corrective feedback?*), and the findings are presented below. Besides, the answers to the interview questions were presented as direct quotations.

3.6.2.1. Content Analyses Based on the Survey and Findings from the Interviews (the experimental group)

The first two questions of the survey concern questions regarding the effectiveness and ineffectiveness of the coded feedback procedure. The participants were asked how they evaluated the feedback procedure in terms of various elements of writing, such as task achievement, coherence and cohesion, range and accuracy of grammar and vocabulary. The findings are presented in the table below.

Table 21.

Content analysis of the areas of effectiveness and ineffectiveness of indirect WCF

Scope of the survey (experimental group)	**Categories**				
Areas of effectiveness of the coded feedback procedure	Task achievement *identifying parts of an opinion paragraph*	Range and Accuracy of Grammar for purpose	Range and Accuracy of Vocabulary for Purpose	Punctuation	Spelling
Mentions	***8***	***9***	***8***	***2***	***1***
Areas of ineffectiveness of the coded feedback procedure	Task achievement *expanding ideas*	Range and Accuracy of Grammar for purpose	Range and Accuracy of Vocabulary for Purpose		
Mentions	***4***	***3***	***4***		

The first category concerns task achievement, in which the participants gave feedback on opinion paragraph structure and developing and expanding ideas. Eight participants expressed that the feedback procedure helped them identify parts of an opinion paragraph, such as the topic sentence, supporting sentences, and concluding sentence(s). On the other hand, four participants stated that they still had difficulty in expanding their ideas by providing supporting points and examples.

"When I look back at my previous texts, I can see how much I have developed myself, especially in providing coherence by using linking words and in giving specific examples to support my ideas."

The second category is related to the range and accuracy of grammar for purpose. Nine of the participants expressed that the coded feedback procedure contributed to their grammatical competence. They stated that they were able to recognize grammatical features and use them more effectively.

"As the time went by and I continued to revise my paragraphs, I started to recognize the errors I had made before, and I tried not to repeat them in the new texts."

"I used to have simple sentences, but I think I make better sentences now."

"My high school teacher used to only underline my errors, and I did not know what kind of error I made. However, this type of feedback helped me recognize my errors. I believe that my use of grammar improved a lot."

However, three participants stated that they could not benefit from the indirect feedback procedure to improve their grammatical competence as much as they expected, and they emphasized that they needed more practice.

"To enhance my grammar, I believe I need more practice. I'm still feeling inept."

The next category is the range and accuracy of vocabulary for purpose. Eight participants commented that the coded feedback procedure was helpful in terms of improving and enlarging their vocabulary.

"I used the previous feedback I got to study for the exams. I revised the feedback concerning my previous texts and also the phrases I used, and thus it enabled me to learn the correct usage. For this reason, I think getting coded feedback is very helpful and I would like to continue it."

However, four of the participants stated that they had difficulty dealing with the lexical errors in their texts. They emphasized that although they understood the error correction code and what it referred to, it was challenging to find the correct word or form, or use a wide range of vocabulary.

"I mostly had trouble with the codes for wrong words and wrong forms, and there were times that I couldn't correct them."

"I still have difficulty using prepositions. I still don't know what word to use instead of the wrong words."

Finally, two of the participants mentioned that the coded feedback procedure enabled them to recognize their errors regarding punctuation and spelling as their errors became more salient.

"I did not use to give importance to punctuation, but now I pay attention to them, and I have learned how to use punctuation marks like comma."

In addition, the participants were addressed a question about their feelings and opinions towards the coded feedback procedure, and they were asked to describe it by using three adjectives. The findings are presented in the table below.

Table 22.

Content analysis regarding students' opinions about indirect WCF

Scope of the survey (experimental group)	**Categories**				
Positive opinions about the coded feedback procedure	Informative /Instructive	Enjoyable	Constructive	Good	Autonomy - supportive
Mentions	***8***	***8***	***7***	***6***	***4***
Positive opinions about the coded feedback procedure	Encouraging	Lasting	Effective/ Useful	Productive	Easy
Mentions	***4***	***4***	***3***	***2***	***1***
Negative opinions about the coded feedback procedure	Challenging	Upsetting	Stressful	Boring	Tiring
Mentions	***6***	***2***	***1***	***1***	***1***

As seen in the table above, in general, the participants keep a positive attitude towards the coded feedback. Eight participants found it informative and instructive.

"Getting coded feedback was highly informative for me."

Eight participants stated that trying to correct their errors was enjoyable. Moreover, seven of the participants commented that the coded feedback was constructive as it allowed them to improve various aspects of their writing.

"I think that the coded feedback procedure that required us to put effort into correcting our errors improved us a lot."

"I was quite nervous in the first week when we started getting feedback with codes, especially when I thought that we were going to repeat the same each week, but I got used to it as time went on. And now I hope that we continue to use them with other teachers."

Six participants commented that it was good to get the feedback with error correction codes rather than direct feedback that their errors were corrected by the teacher. Similarly, four participants agreed that getting coded feedback was much better as it supported their autonomy. Likewise, four of them expressed that the coded feedback encouraged them to write better. In addition, it was said that the coded feedback promoted more long-lasting learning because they put effort into fixing their errors.

"I am really glad that I myself corrected my errors. My high school teacher used to correct my errors directly and give my paper back to me, and then I wouldn't pay attention to my errors. But now what I have learned is more lasting because I myself corrected them."

Three of them found the coded feedback effective and useful, and two of them thought that the procedure was productive for them.

"Before I started to take writing lessons, I assumed that my English was better, but I was puzzled when I saw the number of errors I made, and the feedback helped me to realize my incompetence."

"I did not think that I had made errors before I submitted my paragraph, but I became aware of them when my paper was returned to me with the error correction codes."

Finally, one of the participants mentioned that the error correction codes were easy to understand and apply to the second draft.

"The error correction codes were not difficult to understand because the examples provided for each code were obvious."

Besides the positive aspects, the participants mentioned some negative aspects of the coded feedback. Six of the participants stated that it was challenging for them to learn the error correction codes and figure out what kind of correction needed to be done in their texts. Two of them expressed that they felt upset because of their inability to revise their texts by using codes. One of the participants found the entire process stressful due to its cyclical nature, which required them to revise the text by using codes. For the same reason, one participant found the process boring, and another expressed that it was tiring.

"At the beginning, I could not understand what to do with the error correction codes, but later they became clearer with your explanations."

"The error correction codes were challenging and confusing at the beginning, but I had learned what they referred to by the fourth week, and I did not need to check the list to correct my errors."

"I don't think that I can recall the words I have used in a paragraph before. I'm having trouble remembering the words because they're only used in that subject and I don't use them in my everyday life. Because I don't use them often, I can keep them only in my short-term memory and therefore, I cannot remember and use them in my next writing task. Even though I am asked to write on the same subject again, I don't suppose that I will be able to remember and use them. However, I am trying to understand the words which are often used during the lessons, and so I can keep them in my mind."

The last two questions were relevant to the error correction codes. The participants were asked whether or not there was a code that they found confusing and also if they had ever experienced that they knew what a code referred to but still had difficulty in making the corrections. The findings are presented in the table below.

Table 23.

Content analysis regarding unclear and challenging error correction codes

Scope of the survey (experimental group)	**Categories**			
Unclear error correction codes	Wrong Form (WF)	Unclear meaning (?)		
Mentions	***1***	***1***		
Challenging error correction codes	Add a word (__)	Unclear meaning (?)	Wrong Form (WF)	Wrong Word (WW)
Mentions	***3***	***2***	***2***	***5***

Overall, the participants did not experience any problems related to understanding the error correction codes. Only one participant stated that the code WF, which was used for the wrong form was unclear, and did not know what the teacher asked for as a correction, and similarly, another participant thought the same for the question mark (?), which refers to unclear meaning. In addition to the codes found unclear, some participants found some codes challenging. Three of the participants expressed that the error correction code (_), which means 'add a word', was difficult to deal with as they did not have any ideas about the word to be added. Likewise, two participants emphasized that it is challenging to rewrite a sentence with unclear meaning, and they needed an expert to reformulate the sentence for them. Moreover, two of the participants stated that the errors related to WF (wrong form) were not easy to treat. Lastly, five of the participants agreed that the errors concerning the use of WW (wrong word) were difficult to deal with.

3.6.2.2. Content Analyses Based on the Survey and Findings from the Interviews (the control group)

The control group was asked the same questions as the experimental group, except for the questions concerning error correction codes. The first two questions were about the participants' general views towards the effectiveness of the direct feedback procedure. The findings are presented below.

Table 24.

Content analysis of the areas of effectiveness and ineffectiveness of direct WCF

Scope of the survey (control group)	**Categories**					
Areas of effectiveness of the direct feedback procedure	Task achievement *identifying parts of an opinion paragraph*	Coherence & Cohesion	Range and Accuracy of Vocabulary for Purpose	Range and Accuracy of Grammar for purpose	Punctuation	Spelling
Mentions	***13***	***1***	***7***	***5***	***7***	***3***
Areas of ineffectiveness of the direct feedback procedure	Coherence & Cohesion	Range and Accuracy of Vocabulary for Purpose	Range and Accuracy of Grammar for purpose			
Mentions	***1***	***3***	***5***			

As seen in the table above, the first category was task achievement. The participants were provided some overall feedback on opinion paragraph structure and

developing and expanding ideas. Thirteen participants stated that they benefited from the feedback in identifying parts of an opinion paragraph. They said that they learned what an opinion paragraph is and also how to start and end an opinion paragraph. In addition, one of the participants expressed that s/he learned to use linking words to provide coherence in the paragraph.

"*When I compared my first paragraph with the recent ones, I realized that I had many deficiencies in my writing. I did not know the paragraph structure, and I would not focus on one topic in my paragraphs. I would not use the linking words like firstly, secondly to provide coherence."*

"I did not know that there were different types of paragraphs. I have learned to focus on a single idea in an opinion paragraph. In addition, I have learned how to begin, develop, and end an opinion paragraph. When compared with my previous paragraphs, the recent ones are quite different."

On the other hand, one participant stated that the feedback procedure did not help much to improve cohesion and coherence in the paragraph.

The next two categories were range and accuracy of vocabulary and grammar for purpose. Seven participants stated that the feedback procedure assisted the progress of their grammar and vocabulary as they became aware of their errors.

"Although I knew some vocabulary, I would not pay attention to which form of the word to use. I had the opportunity to examine the correct forms thanks to the feedback, and I believe that I have gained great awareness of their correct uses."

"I improved my writing thanks to the feedback. I did not feel bad at all after the corrections because the corrections helped me realize my errors. I would use a translation app to write most of my sentences, but I don't need to use it as much as I did before."

"I think I improved my writing in terms of grammar, vocabulary, and spelling. I know the appropriate words for particular sentences. On the other hand, it is difficult to write on topics that I don't know any vocabulary about."

"To some degree, I think I have developed my writing in many ways compared to the past. I still have difficulty with the use of some features, such as prepositions. However, the number of errors is smaller. Let's say it has decreased from five incorrect sentences to two incorrect ones."

Besides the positive comments, five participants commented that they still have difficulty with grammar and vocabulary and need more practice to get better. One of the participants emphasized that it would have been better if corrections had not been provided directly because s/he did not need to put any effort into dealing with the errors.

"I still have some problems with grammar. I have been learning English since I was a child, but I haven't been able to improve my grammar."

"I have improved my vocabulary, but I believe that I need more practice to get better."

"I am having difficulty using different tenses. I got a poor writing score due to my incompetence in grammar. I think I could not improve myself because corrections were already provided, and so I did not try to keep them in mind."

"It could be due to distance education or the course book we used, but I expected more improvement in my vocabulary. I had believed that I could improve a bit more in seven weeks, but I may need more time."

Finally, seven participants mentioned that they pay more attention to punctuation, and three participants stated that they think more about the correct spelling of the words.

"I am trying to pay more attention to punctuation."

"I think that I improved my writing in every aspect, including the use of punctuation marks and capitalization."

"I would not use any commas, but now I know where to put them in my sentences."

The participants were also asked about their feelings and opinions towards the feedback procedure, and they were asked to describe it by using three adjectives. The findings are presented in the table below.

Table 25.

Content analysis regarding students' opinions about direct WCF

Scope of the survey (control group)	**Categories**				
Positive opinions about the direct feedback procedure	Good	Effective/ Useful	Informative /Explanatory	Constructive	Systematic /Planned
Mentions	***11***	***7***	***7***	***4***	***3***
Positive opinions about the direct feedback procedure	Fantastic/ Perfect	Awareness - raising	Enjoyable	Encouraging	Comprehensible
Mentions	***3***	***3***	***2***	***2***	***1***
Negative opinions about the direct feedback procedure	Challenging	Upsetting	Embarrassing		
Mentions	***1***	***1***	***1***		

As seen in the table above, the participants have mostly positive feelings and opinions about the feedback procedure. Eleven participants commented that the direct feedback procedure was good, and three of them described it as fantastic and perfect.

"I felt good and happy because I could see my errors and I knew I needed to improve myself in those areas. I believe that revising my paragraph improved my writing."

Seven of the participants found the feedback effective and useful, and three of them expressed that the feedback led to raised awareness of their errors.

"I don't think I could write well at the beginning and I am able to write much better than in the past. I think that it was good to see my errors."

"Before I started writing a new paragraph, I reviewed my previous paragraphs. There were times that I repeated the same error, however, I started writing with the awareness I gained from my previous paragraphs."

Seven participants stated that the entire process was informative, and four participants found it constructive. Moreover, two of them found it enjoyable, and another two participants thought that it was encouraging.

"When I got feedback regarding the incorrect use of tenses, I wrote them down and revised those tenses."

Three participants expressed that the feedback was provided systematically and so it caused them to give more importance to it.

"This feedback procedure was advantageous in many ways. I cannot be considered a perfect student, but I made an effort to complete my assignments regularly, and so I benefited from the feedback. As we knew we were expected to rewrite our paragraphs by reviewing the corrections, and you followed this procedure regularly each week, I paid much more attention to the errors in my assignments, otherwise I would not."

"I think that my improvement in English was slightly hampered by distance education. However, it is not true for my English writing. Conducting writing lessons systematically on a particular platform, being notified by the system when my paper was returned with feedback, and also being able to ask questions through the same system were highly valuable to me. I can say that I notice my improvement in writing."

Finally, one participant stated that the feedback was comprehensible as it allowed her to see the error and the correction together.

"I liked being explicitly corrected. I did not simply copy and paste the corrections, but instead I preferred to rewrite the whole paragraph so as to recall them better next time. In addition, as I could see the error and the correction together, it was possible to compare them and see the difference between them."

Furthermore, a few participants expressed some negative feelings and opinions towards the feedback. One of the participants found it challenging, and another one thought it was upsetting, and finally one of them expressed that being explicitly corrected was embarrassing.

"Despite my progress in writing, I still feel incompetent in terms of expressing my ideas in English."

CHAPTER IV

RESULTS AND DISCUSSION

4.1. Introduction

The findings of the study, which was carried out to investigate the effects of indirect written corrective feedback on overall writing performance, are discussed in this chapter.

Result and Discussion of RQ1. *Is indirect written corrective feedback effective in improving students' overall writing performance?*

To answer the first research question, the results of the pretest and posttest were compared. Initially, the pre-test and post-test results of the whole group of participants, regardless of the type of feedback they received, were compared. As a result, the scores of the whole group of participants related to all variables including Task Achievement, Coherence and Cohesion, Grammar and Vocabulary significantly increased from the pre-test to the post-test ($p < .001$), $d > 0.80$). Therefore, the findings demonstrated that both direct and indirect written corrective feedback seem to have facilitated improvement in revision task.

Furthermore, the participants were also divided by their respective feedback groups and compared by means of ANOVA's and Quade's tests to see if one group outperformed the other or not. According to the descriptive results, the mean post-test

scores of the experimental group were slightly higher than those of the control group in all variables, by 0.86 point in Total scores, 0.31 point in Task Achievement, 0.16 point in Coherence & Cohesion, 0.14 point in Vocabulary and 0.25 point in Grammar. Therefore, covariance analyses were run for each of the variables to see if these differences were statistically significant. ANCOVA results showed that the total pre-test scores had a statistically significant effect on the total post-test scores with a large effect ($F_{(1,\ 33)} = 7.389$, $p < .05$, $\eta_p^2 = 0.18$). When the total pre-test scores were controlled for, it was seen that there was no difference in the Total post-test scores of the control group (Group 1) and the experimental group (Group 2) ($F_{(1,\ 33)} = 0.812$, $p > .05$, $\eta_p^2 = 0.02$). Moreover, it was seen that similar results were achieved for all variables, in Task Achievement ($F_{(1,\ 34)} = 0.548$, $p = .464$), in Coherence and Cohesion ($F_{(1,\ 33)} = 0.350$, $p > .05$, $\eta_p^2 = 0.01$), in Vocabulary ($F_{(1,\ 33)} = 0.412$, $p > .05$, $\eta_p^2 = 0.01$), and finally in Grammar ($F_{(1,\ 34)} = 1.370$, $p = .250$). It could be argued that the small sample size led to this result. Larger sample size may have resulted in different findings, and thus help better understand the effectiveness of indirect written corrective feedback.

Apart from this, this result could be partly due to the fact that the corrections were provided via an online platform in distance education context, and the participants who received direct correction were able to access their first drafts with corrections while revising their texts. If the participants had been allowed to review their corrections only for a short period of time and then their first drafts with corrections had been taken away, the results could have been different since for the participants, the possibility of copying and pasting the corrections would have been prevented, and so it would have yielded more accurate results, and some differences between the feedback groups may have been observed. On the other hand, even if the research had not been conducted in an online context, and the copies of student papers with corrections had been taken away from them after allowing students to review their papers for a short time period, it would have raised questions with regard to the ecological validity of the research since there are not many teachers who provide corrections on the photocopies of the original papers.

Moreover, this research was completed in a relatively short time period, which lasted seven weeks. None of the participants in the experimental group had received indirect written corrective feedback with error correction codes before. Therefore, the indirect corrective feedback group might need more training and time to make sense of error correction codes (Ferris, 2004).

In conclusion, this result adds to the previous research which found both types of written corrective feedback effective in developing revision accuracy (e.g. Ferris & Roberts, 2001; Van Beuningen et al., 2008; Van Beuningen et al., 2012; Al-Rubai'ey and Nassaji, 2013; Lopez et al., 2018; Karim & Nassaji, 2018; Suzuki et al., 2019).

Besides the post-test, which was on the same topic and genre as the pre-test, a second post-test on opinion paragraph with a different writing topic was also collected from the participants to see the effectiveness of indirect corrective feedback on a new piece of writing. It was seen that the first and the second post-tests were not significantly different in terms of the mean Total scores (t = 0.940, df = 35, $p > .05$) and the mean scores of Coherence & Cohesion (t = 0.657, df = 35, $p > .05$, Z = -0.459, p = .646, r = 0.05) and Vocabulary (t = 2.029, df = 35, p = .050, Z = -1.949, p = .051, r = 0.23). On the other hand, Task Achievement (t = -2.707, df = 35, p = .01, Z = -2.588, p = .010, r = 0.31) and Grammar (t = 2.615, df = 35, $p < .05$, Z = -2.023, p = .043, r = 0.24) scores were significantly different in the first and the second post-tests. It revealed that both parametric and nonparametric analyses yielded the same results with regard to the pairwise comparisons. The mean Total scores showed that the whole group of participants retained their gains in the second post-test, which was on a new piece of writing.

However, significant differences between the first and second post-tests in Task Achievement and Grammar should be noted. While the participants performed better in achieving the task in the second post-test, in Grammar, they did not perform as well as they did in the first post-test. The task effect may be the reason why the

participants performed better in achieving the task in the second post-test. Moreover, considering the second post-test scores were the writing quiz results, this result in Grammar may be associated with exam anxiety. The decrease in student motivation through the final weeks of the module was the main reason which caused the researcher to use quiz results in the second post-test. The writing quiz results were used as the second post text since it was assumed that the quiz would encourage the participants to take the task more seriously and do their best, and thus it would reflect a more realistic performance for the research. However, the possible effects of exam anxiety for some participants cannot be disregarded.

To sum up, the results of the present study regarding the first research question and the results of other previous studies imply that indirect written corrective feedback has positive effects in improving students' overall writing performance both in revision task and in new piece of writing. However, the same positive results were also obtained for the direct written corrective feedback group.

Result and Discussion of RQ2. *What are the participants' opinions about direct and indirect written corrective feedback?*

Truscott (1996, 1999, 2004, 2007) has repeatedly claimed that written corrective feedback is ineffective in developing writing accuracy. However, research evidence demonstrates that written corrective feedback is effective in facilitating accuracy development both in revision and in new pieces of writing (e.g. Van Beuningen et al., 2012; Lopez et al., 2018; Suzuki et al., 2019).

Truscott's (1996) other strong claim was that corrective feedback is harmful since it demotivates learners by increasing their stress levels. Furthermore, Truscott argued that due to corrective feedback, students tend to write short and simple sentences to avoid errors, and therefore the complexity of their texts is reduced.

Truscott suggested that other activities like extra writing practice are more beneficial than spending time trying to understand corrections.

The present study also aimed to find out the participants' opinions about direct and indirect written corrective feedback. The direct feedback group was also included in this part of the research so as to provide a more comprehensive view of students' opinions about both types of feedback. Along with the interviews, an online survey was conducted with six participants from each group. The content analyses based on the surveys and the findings from the interviews revealed that students expect to be corrected, and value teacher feedback.

The majority of the participants in both groups stated positive opinions and feelings concerning overall feedback on opinion paragraph structure and developing and expanding ideas. Although this result is not directly related to the effectiveness of any feedback type, it is valuable in terms of showing students' general perception towards feedback.

As for grammatical and lexical errors, the students in the indirect feedback group emphasized that the feedback helped them recognize their errors and also motivated them to avoid the same errors in their following works because they were able to compare and contrast their previous works with the recent ones. A few participants also mentioned the positive effects of coded feedback on spelling and punctuation. Similarly, the students in the direct feedback group stated that they benefited from direct feedback in improving their grammar and vocabulary because they became more aware of their errors. On the other hand, unlike the indirect feedback group, more students in the direct feedback group expressed that they improved in punctuation and spelling. As a result, it is possible to say that written corrective feedback, either direct or indirect, may enhance salience by drawing attention to erroneous uses, which is a necessary condition for acquisition (Schmidt, 1990).

Nevertheless, a few students in both groups stated that they still felt incompetent, and they needed more writing practice to improve their grammar and vocabulary. Moreover, one student especially emphasized that he could keep new words in mind only for the present task and he would not be able to use these words in a different task. In addition, he stated that in order to acquire those newly learned words, he needed repeated exposure to them, and he added that if those words are often used in class, it was easy for him to learn them. What he stated points out "the problem of pseudolearning" put forward by Truscott (1996). As Truscott argued, knowledge acquired through grammar correction may be transient and it may not be sufficient for acquisition to take place, and it may be useful only for editing purposes. As Truscott (1996) emphasized, it could be argued that allocating more time for writing practice and being more selective in correcting errors may help manage time more effectively. Similarly, Liu (2008) concluded that only corrective feedback is not sufficient, and some mini-lessons or workshops are also needed for accuracy development and students' self-editing ability.

Regarding students' opinions and feelings about the indirect feedback, in contrast to Truscott's argument for corrective feedback being harmful as it increases students' stress levels, most of the students found indirect corrective feedback encouraging rather than demotivating. The findings from the survey and interviews indicated the students' positive opinions about the coded feedback. When compared to the number of negative comments (*n=5*) that were mentioned eleven times, more positive comments (*n=10*) that were mentioned forty-seven times were made as to indirect feedback. The students believed that the coded feedback supported their autonomy as they dealt with the corrections on their own. Four students pointed out that the coded feedback resulted in more lasting learning because of its autonomy-supportive nature. In addition, eight students found the coded feedback procedure enjoyable. On the other hand, the mostly used negative expression was challenging, which were mentioned six times. Apart from this, only a few other negative expressions like stressful, boring, tiring, and upsetting each of which were mentioned once were used to define the indirect feedback process.

In addition, the participants in the indirect feedback group were posed questions about whether or not they found any error correction codes unclear and challenging. Although the majority of the students did not experience any problems in understanding what a code refers to, some students found some correction codes such as “add a word (_)”, “unclear meaning (?)”, “wrong form (WF)”, and “wrong word (WW)” challenging. They expressed that they needed instruction and time to interpret error correction codes and revise their texts accordingly. As also stated by Ferris (2004), it proves that students might need more time to comprehend the method used by the teacher.

As for the opinions and feelings of the direct feedback group, it was found that the participants expressed positive opinions and feelings about direct feedback. In total, ten positive expressions mentioned forty-two times were used to define the direct feedback procedure, and only three negative expressions, such as challenging, upsetting, and embarrassing, were each mentioned once. One student commented that she liked and preferred being explicitly corrected so she could see the error and correction together and so it enabled her to see the difference between them. This comment draws attention to how important it is to know students’ expectations. While some students may like being corrected directly, others may favour indirect feedback (Armhein & Nassaji, 2010).

To sum up, it is important to take learners’ preferences and expectations into consideration as it may influence their success and motivation (Storch, 2010; Storch & Wigglesworth, 2010). The present study demonstrates that students have positive opinions about written corrective feedback, either direct or indirect, due to its positive effects in improving students’ overall writing performance.

CHAPTER V

CONCLUSION

5.1. Conclusion

From 1940s to 1980s, second language writing had taken a back seat in comparison with speaking, reading and listening. Ever since 1980s, with the acknowledgement of writing as a crucial skill which serves both to express the self and to communicate with the reader, there has been an increasing interest in the writing research.

The distinctive nature of writing leads to challenges in learning and teaching writing. Writing does not only rely on the mechanics of language, but also has a creative side. In this respect, in order to address various aspects of writing, the researchers have investigated different methods of teaching writing and also the teacher's role. In process writing, giving corrective feedback is determined as one of the roles of teacher in teaching writing. Despite some counter-arguments to the effectiveness of written corrective feedback, the research evidence presents findings for its usefulness in improving the level of accuracy. On the other hand, the debate regarding the effectiveness of different types of written corrective feedback has been going on. While some researchers put forward that direct correction of errors is more useful, other researchers state that indirect correction (e.g. underlining the erroneous structures, or using correction codes) is more effective, especially in the long term as it enables more lasting learning.

The present study investigated the effectiveness of indirect written corrective feedback on students' overall writing performance. It was aimed to see if there were

any differences between the experimental group who received indirect written corrective feedback through error correction codes and the control group who received direct written feedback, which is a common and traditional way of providing feedback. In addition, the participants' opinions about direct and indirect feedback were researched.

The participants were thirty-six university students enrolled in a General English preparatory class at the School of Foreign Languages at Beykent University, which is a foundation university in Istanbul, Turkey. 18 participants were included in each group. The two groups of participants who were assigned to their classes according to the placement test results were both A2 level. The participants in both groups were enrolled in different programmes. Besides, they were all non-native speakers of English.

In this study, pre and post-tests were carried out to collect the quantitative data. In addition to these, an online survey including open-ended questions and semi-structured interviews were included in the qualitative data collection procedure.

Over the course of the treatment, the experimental group received indirect written corrective feedback through error correction codes, and the control group received direct corrective feedback, and their errors were directly corrected by the teacher. In the first week, both groups took a pre-test. The participants in the experimental group were then trained on error correction codes. During the six-week period, participants in both groups were also instructed on how to write an opinion paragraph and were given brief feedback on organizational and content issues in their papers. In week 6, the participants were given a post-test to see the revision effects of indirect corrective feedback. In addition, in week 7, the participants also took a second post-test, which was administered to see the transfer effects of indirect written corrective feedback on a new piece of writing. Moreover, both groups were given an

online survey to find out their opinions about the feedback procedure. And six participants from each group were interviewed to get additional data.

In order to measure the effectiveness of indirect written corrective feedback, pre-test and post-test scores of both groups were statistically compared. Moreover, qualitative content analysis was used to analyse online surveys, and the content analysis was enriched by the interview data.

The major findings of the present study are as follows:

1. There was no statistically significant difference among the feedback groups in terms of the pre-test scores.
2. In comparisons of the whole group (N=36), regardless of the feedback type, there were statistically significant differences with large effects in all pre-test and post-test results. The scores related to all variables significantly increased from the pre-test to the post-test.
3. In comparisons of the groups based on the feedback type, it was seen that there was no difference in the Total post-test scores of the experimental group and the control group. In addition, the analysis of Task Achievement, Coherence & Cohesion, Grammar and Vocabulary scores revealed that the post-test scores did not significantly differ between feedback groups.
4. The first and the second post-tests of the whole group of participants were not significantly different in terms of the mean Total scores. On the other hand, Task Achievement and Grammar scores were significantly different in the first and the second post-tests. While the whole group of participants had better result in Task Achievement according to the second post-test results, their Grammar result was not as good as the first post-test, which was on the revision of same text.
5. It was seen that students are mostly positive towards direct and indirect written corrective feedback.

According to the findings given above, the following conclusions can be drawn:

1. There are no significant effects of indirect written corrective feedback given through error correction codes as both feedback groups improved almost to the same degree in the revision of the same text.
2. In the immediate post-test, it was seen that students in both groups could transfer their gains to a new piece of writing.
3. Students value and expect teacher feedback.
4. It is important to consider students' opinions about written corrective feedback type as it may influence their success and motivation.
5. Students need some time and training to familiarize themselves with the written corrective feedback type.
6. Using error correction codes is not effective for all kinds of errors because students have difficulty understanding and dealing with some error correction codes, especially when the meaning is unclear. A combination of direct and indirect written corrective feedback types could be more useful for students.
7. Besides written corrective feedback, students need extra writing practice and mini-lessons based on their needs.

In conclusion, the type of written corrective feedback is not the only factor to consider regarding the development of writing performance. It is essential to know students' opinions and preferences.

5.2. Limitations of the Study

The small sample size was the major limitation of the present study. Each group included 18 participants after the elimination of some participants due to illness or unwillingness to participate in the study. In addition, they were all students at Beykent University, and convenient sampling was used for the research.

Moreover, the study was administered in an online learning context due to COVID-19 restrictions. The participants were not familiar with the distance education context, and therefore some of them needed assistance with digital learning platforms.

Apart from this, the present study lasted only seven weeks, and after seven weeks, the participants took the final exams and continued their lessons in different classes depending on their levels. It would yield better results to implement a delayed post-test in order to see transfer effects on a new piece of writing over time.

5.3. Suggestions for Further Research

Further research could be designed by increasing the number of participants and by using random sampling. The present study was quasi-experimental and it was not possible to have a real control group for ethical reasons. Therefore, the participants in the control group received direct correction, which is a common and traditional method. Including a control group with no corrective feedback would give more accurate findings. The studies so far have mostly been limited to quantitative research. A mixed-method research approach with improved design could yield valuable findings.

REFERENCES

Allen, R. (1966). Written English is a "Second Language". *The English Journal, 55*(6), 736-746.

Al-Rubai'ey, F., & Nassaji, H. (2013). Direct and indirect metalinguistic feedback: A matter of suitability rather than superiority. In M. Mahmoud, & R. Al Mahrooqi (Eds.), *Issues in TEFL in the Arab World* (pp. 203-231). Muscat, Oman: Sultan Qaboos University Press.

Armhein, H. R., & Nassaji, H. (2010). Written Corrective Feedback: What do students and teachers prefer and why? *Canadian Journal of Applied Linguistics, 13*, 95-127.

Ashwell, T. (2000). Patterns of teacher response to student writing in a multi-draft composition classroom: is content feedback followed by form feedback the best method? *Journal of Second Language Writing, 9*, 227-257.

Bitchener, J. (2008). Evidence in support of written corrective feedback. *Journal of Second Language Writing, 17*, 102-118.

Bitchener, J., & Ferris, D. R. (2012). *Written Corrective Feedback in Second Language Acquisition and Writing.* New York: Routledge.

Bitchener, J., & Knoch, U. (2008). The value of written corrective feedback for migrant and international students. *Language Teaching Research, 12*(3), 409-431.

Bitchener, J., & Knoch, U. (2009). The relative effectiveness of different types of direct written corrective feedback. *System, 37*, 322-329.

Bitchener, J., & Knoch, U. (2010). The Contribution of Written Corrective Feedback to Language Development: A Ten-Month Investigation. *Applied Linguistics, 31*(2), 193-214.

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

Blanca, M. J., Alarcón, R., Arnau, J., Bono, R., & Bendayan, R. (2017). Non-normal data: Is ANOVA still a valid option? *Psicothema, 29*(4), 552-557.

Bloomfield, L. (1984). *Language.* Chicago: University of Chicago Press.

Brown, H. D. (2001). *Teaching by Principles: An Interactive Approach to Language Pedagogy.* White Plains, New York: Longman.

Bruton , A. (2010). Another reply to Truscott on error correction: Improved situated designs over statistics. *System, 38*, 491-498.

Bruton, A. (2009). Designing research into the effects of grammar correction in L2 writing: Not so straightforward. *Journal of Second Language Writing, 18*, 136-140.

C. Weigle, S. (2014). Considerations for Teaching Second Language Writing. In M. Celce-Murcia, M. A. Snow, & D. M. Brinton (Eds.), *Teaching English as a Second or Foreign Language* (pp. 222-237). Boston, USA: National Geographic Learning.

Cancino, H., Rosansky, E. J., & Schuman, J. H. (1978). The acquisition of English negatives and interrogatives by native Spanish speakers. In E. M. Hatch (Ed.), *Second Language Acquisition: A book of readings* (pp. 207-230). Rowley, MA: Newbury House.

Cangür, Ş., Sungur, M., & Ankaralı, H. (2018). The Methods Used in Nonparametric Covariance Analysis. *Duzce Medical Journal, 20*(1), 1-6. doi:10.18678/dtfd.424774

Chandler, J. (2003). The efficacy of various kinds of error feedback for improvement in the accuracy and fluency of L2 student writing. *Journal of Second Lanugage Writing, 12*, 267-296.

Chandler, J. (2004). A response to Truscott. *Journal of Second Language Writing, 13*, 345-348.

Chandler, J. (2009). Response to Truscott. *Journal of Second Language Writing, 18*, 57-58.

Chomsky, N. (1957). *Syntactic Structures.* the Hague: Mouton.

Corder, S. P. (1967). The significance of learners' errors. *International Review of Applied Linguistics, 5*, 161-170.

Corder, S. P. (1973). *Introducing applied linguistics.* Harmondsworth, UK: Penguin.

DeKeyser, R. (2015). Skill acquisition theory. In B. VanPatten, & J. Williams (Eds.), *Theories in Second Language Acquisition: An Introduction* (pp. 94-112). New York: Routledge.

Derrida, J. (1976). *Of Grammatology.* (G. Chakravorty Spivak, Trans.) Baltimore: The Johns Hopkins University Press.

Doughty, C., & Varela, E. (1998). Communicative focus on form. In C. Doughty, & J. Williams (Eds.), *Focus on form in classroom second language acquisition* (pp. 114-138). Cambridge: Cambridge University Press.

Dulay, H., & Burt, M. (1973). Should we teach children syntax? *Language Learning, 24*, 245-258.

Ellis, R. (2009). Corrective Feedback and Teacher Development. *L2 Journal, 1*(1), 2-18.

Ellis, R., Sheen, Y., Murakami, M., & Takashima, H. (2008). The effects of focused and unfocused written corrective feedback in an English as a foreign language context. *System, 36*, 353-371.

Fathman, A. K., & Whalley, E. (1990). Teacher response to student writing: focus on form versus content. In B. Kroll (Ed.), *Second Language Writing: Research Insights for the Classroom* (pp. 178-190). Cambridge: Cambridge University Press.

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

Ferman, S., Olshtain, E., Schectman, E., & Karni, A. (2009). The acquisition of a linguistic skill by adults: Procedural and declarative memory interact in the learning of an artificial morphological rule. *Journal of Neurolinguistics, 22*, 384-412.

Ferris, D. (1995b). Student Reactions to Teacher Response in Multiple-Draft Composition Classrooms. *TESOL Quarterly, 29*(1), 33-53.

Ferris, D. (1997). The Influence of Teacher Commentary on Student Revision. *TESOL Quarterly, 31*(2), 315-339.

Ferris, D. (1999). The Case for Grammar Correction in L2 Writing Classes: A Response to Truscott (1996). *Journal of Second Language Writing, 8*(1), 1-11.

Ferris, D. (2004). The ‘‘Grammar Correction’’ Debate in L2 Writing: Where are we, and where do we go from here? (and what do we do in the meantime . . .?). *Journal of Second Language Writing, 13*, 49-62.

Ferris, D. (2006). Does error feedback help student writers? New evidence on the short- and long-term effects of written error correction. In K. Hyland, & F. Hyland (Eds.), *Feedback in Second Language Writing* (pp. 81-104). Cambridge: Cambridge University Press.

Ferris, D. (2010). Second Language Writing Research and Written Corrective Feedback in SLA: Intersections and Practical Applications. *Studies in Second Language Acquisition, 32*, 181-201.

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

Field, A. (2018). *Discovery statistics using IBM SPSS Statistics.* London, UK: Sage Publications LTD.

Fraenkel, J. R., Wallen, N. E., & Hyun, H. H. (2012). *How to Design and Evaluate Research in Education.* New York: McGraw-Hill.

Frear, C., & Chiu, Y. (2015). The effect of focused and unfocused indirect written corrective feedback on EFL learners' accuracy in new pieces of writing. *System, 53*, 24-34.

Gascoigne, C. (2004). Examining the effect of feedback in beginning L2 composition. *Foreign Language Annals, 37*, 71-76.

Gass, S. M., Behney, J., & Plonsky, L. (2013). *Second Language Acquisition: An Introductory Course.* New York: Routledge.

Gholaminia, I., Gholaminia, A., & Marzban, A. (2014). An investigation of meta-linguistic corrective feedback in writing performance. *Procedia - Social and Behavioral Sciences, 116*, 316-320.

Halliday, M. A. (1970). Language Structure and Language Function. In J. Lyons (Ed.), *New horizons in linguistics* (pp. 128-169). Harmondsworth: Penguin.

Harmer, J. (2004). *How to teach writing.* Harlow, UK: Pearson Education Limited.

Harmer, J. (2007). *The Practice of English Language Teaching.* Harlow, United Kingdom: Pearson Education Limited.

Hartshorn, K. J., Evans, N. W., Merril, P. F., Sudweeks, R. R., Strong-Krause, D., & Anderson, N. J. (2010). Effects of Dynamic Corrective Feedback on ESL Writing Accuracy. *TESOL Quarterly, 44*(1), 84-109.

Hedge, T. (2000). *Teaching and Learning in the Language Classroom.* New York: Oxford University Press.

Hedgecock, J., & Lefkowitz, N. (1994). Feedback on feedback: Assessing learner receptivity in second language learning. *Journal of Second Language Writing, 3*(2), 141-163.

Hendrickson, J. M. (1978). Error correction in foreign language teaching: Recent theory, research, and practice. *Modern Language Journal, 62*, 387-398.

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

Hirvela, A. R. (2004). *Connecting Reading & Writing in Second Language Writing Instruction.* Ann Arbor, MI: The University of Michigan Press.

Horning, A. S. (1986). *Teaching Writing as a Second Language.* Carbondale: Southern Illinois University Press.

Hyland, K. (2009). *Teaching and Researching Writing.* Harlow, England: Pearson.

Hyland, K., & Anan, E. (2006). Teachers' perceptions of error: The effects of first language and experience. *System, 34*(4), 509-519.

Johnson, R. B., Onwuegbuzie, A. J., & Turner, L. A. (2007). Toward a Definition of Mixed Methods Research. *Journal of Mixed Methods Research, 1*(2), 112-133.

Karim, K., & Nassaji, H. (2018). The revision and transfer effects of direct and indirect comprehensive corrective feedback on ESL students' writing. *Language Teaching Research, 24*(4), 519-539.

Karim, K., & Nassaji, H. (2019). The effects of written corrective feedback: a critical synthesis of past and present research. *Instructed Second Language Acquisition, 3*(1), 28-52.

Kepner, C. G. (1991). An Experiment in the Relationship of Types of Written Feedback to the Development of Second-Language Writing Skills. *The Modern Language Journal, 75*(3), 305-313.

Kéry, M., & Hatfield, J. S. (2003). Normality of raw data in general linear models: The most widespread myth in statistics. *Bulletin of the Ecological Society of America, 84*(2), 92-94.

Krashen, S. D. (1981). *Second language acquisition and second language learning.* Oxford: Pergamon.

Krashen, S. D. (1982). *Pricnciples and practices in second language acquisition.* Oxford: Pergamon.

Krashen, S. D. (1985). *The input hypothesis: Issues and implications.* London: Longman.

Lalande, J. F. (1982). Reducing Composition Errors: An experiment. *Modern Language Journal, 66*, 140-149.

Larson-Hall, J. (2016). *A guide to doing research in second language acquisition with SPSS and R.* New York, NY: Routledge.

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

Leki, I. (1991). The Preferences of ESL Students for Error Correction in College-Level Writing Classes. *Foreign Language Annals, 24*(3), 203-218.

Lightbown, P. M. (1998). The importance of timing in focus on form. In C. Doughty, & J. Williams (Eds.), *Focus on form in classroom second language acquisition* (pp. 177-196). Cambridge: Cambridge University Press.

Lightbown, P. M., & Spada, N. (1999). *How languages are learned.* Oxford, UK: Oxford University Press.

Liu, Y. (2008). The Effects of Error Feedback in Second Language Writing. *Arizona Working Papers in SLA & Teaching, 15*, 65-79.

Long, M. H. (1996). The role of the linguistic environment in second language acquisition. In W. Ritchie, & T. Bhatia (Eds.), *Handbook of Second Language Acquisition* (pp. 413-468). San Diego, CA: Academic Press.

Lopez, M. B., Steendam, E. V., Speelman, D., & Buyse, K. (2018). The Differential Effects of Comprehensive Feedback Forms in the Second Language Writing Class. *Language Learning, 68*(3), 813-850.

Lyster, R., & Mori, H. (2006). Interactional feedback and instructional counterbalance. *Studies in Second Language Acquisition, 28*, 269-300.

Mizumoto, A., & Takeuchi, O. (2018). Modeling a prototypical use of language learning strategies: Decision tree-based methods in multiple contexts. In R. L. Oxford, & C. M. Amerstorfer (Eds.), *Language Learning strategies and individual learner characteristics: Situation strategy use in diverse contexts* (pp. 99-122). New York, NY: Bloomsbury.

Moscal, B. M., & Leydens, J. A. (2000). Scoring Rubric Development: Validity and Reliability. *Practical Assessment, Research and Evaluation, 7*(10). doi:https://doi.org/10.7275/q7rm-gg74

Nicklin , C., & Plonsky, L. (2020). Outliers in L2 research in applied linguistics: A synthesis and data re-analysis. *Annual Review of Applied Linguistics, 40*, 26-55. doi:https://doi.org/10.1017/S0267190520000057

Nunan, D. (2015). *Teaching English to Speakers of Other Languages: An Introduction.* New York: Routledge.

Olejnik, S. F., & Algina , J. (1984). Parametric ANCOVA and the rank transform ANCOVA When the data are conditionally non-normal and heteroscedastic. *Journal of Behavioral and Educational Statistics, 9*(2), 129-149.

Olejnik, S. F., & Algina, J. (1985). A Review of Nonparametric Alternatives To Analysis of Covariance. *Evaluation Review, 9*(1), 51-83. doi:10.1177/0193841X8500900104

Ong, W. (2012). *Orality and Literacy.* Oxon: Routledge.

Polio , C. (2001). Research methodology in second language writing research: The case of text-based studies. In T. Silva, & P. K. Matsuda (Eds.), *On Second Language Writing* (pp. 91-116). Mahwah, NJ: Erlbaum.

Rheinheimer, D. C., & Penfield, D. A. (2001). The effects of type I error rate and power of ANCOVA F test and selected alternatives under nonnormality and variance heterogeneity. *The Journal of Experimental Education, 69*(4), 373-391.

Ricci, V. (2005). *Fitting Distributions with R. Vienna: R Foundation for Statistical Computing.* Retrieved from https://cran.r-project.org/doc/contrib/Ricci-distributions-en.pdf

Robb, T., Ross, S., & Shortreed, I. (1986). Saliennce of Feedback on Error and Its Effect on EFL Writing Quality. *TESOL Quarterly, 20*(1), 83-95.

Rousseau, J. -J. (2018). Essay on the Origin of Languages. In J. Rousseau, & V. Gourevitch (Ed.), *Rousseau: The Discourses and Other Early Political Writings* (pp. 257-310). Cambridge: Cambridge University Press.

Sapir, E. (1921). *Language: An Introduction to the Study of Speech.* New York: Harcourt, Brace.

Schmidt, R. (1990). The role of consciousness in second language learning. *Applied Linguistics, 11*, 129-158.

Scrivener, J. (2011). *Learning Teaching: The Essential Guide to English Language Teaching.* Oxford, UK: Macmillan Education.

Segalowitz, N., & Lightbown, P. (1999). Psycholinguistic approaches to SLA. *Annual Review of Applied Linguistics, 19*, 43-63.

Semke, H. D. (1984). Effects of the Red Pen. *Foreign Language Annals, 17*(3), 195-202.

Shaughnessy, M. P. (1977). *Errors and expectations.* New York: Oxford University Press.

Sheen, Y. (2007). The effect of focused written corrective feedback and language aptitude on ESL learners' acquisition of articles. *TESOL Quarterly, 41*(2), 255-283.

Sheen, Y. (2010). The Role of Oral and Written Corrective Feedback in SLA. *Studies in Second Language Acquisition, 32*, 169-179.

Sheen, Y., & Ellis, R. (2011). Corrective Feedback in Language Teaching. In E. Hinkel (Ed.), *Handbook of Research in Second Language Teaching and Learning.* New York: Routledge.

Sheen, Y., Wright, D., & Moldawa, A. (2009). Differential effects of focused and unfocused written correction on the accurate use of grammatical forms by adult ESL learners. *System, 37*, 556-569.

Sheppard, K. (1992). Two Feedback Types: Do They Make a Difference? *RELC Journal, 23*(1), 103-110.

Shielzeth, H., Dingemanse, N. J., Nakagawa, S., Westneat, D. F., Allegue, H., Teplitsky, C., & Araya-Ajoy, Y. G. (2020). Robustness of linear mixed-effects models to violations of distributional assumptions. *Methods in Ecology and Evolution, 11*(9), 1141-1152.

Shintani, N., & Ellis, R. (2013). The comparative effect of direct written corrective feedback and metalinguistic explanation on learners' explicit and implicit knowledge of the English indefinite article. *Journal of Second Language Writing, 22*, 286-306.

Shintani, N., Ellis, R., & Suzuki, W. (2014). Effects of Written Feedback and Revision on Learners' Accuracy in Using Two English Grammatical Structures. *Language Learning, 64*, 103-131.

Silva, T. (1990). Second language composition instruction: developments,issues, and directions in ESL. In B. Kroll (Ed.), *Second Language Writing; Research insights for the classroom* (pp. 11-23). New York: Cambridge University Press.

Sokolik, M. (2003). Writing. In D. Nunan (Ed.), *Practical English Language Teaching* (pp. 87-108). New York: McGraw Hill.

Storch, N. (2010). Critical feedback on written corrective feedback research. *IJES, 10*, 29-46.

Storch, N., & Wigglesworth, G. (2010). Learners' Processing, Uptake, and Retention of Corrective Feedback on Writing. *Studies in Second Language Acquisition, 32*, 303-334.

Suzuki, W., Nassaji, H., & Sato, K. (2019). The effects of feedback explicitness and type of target structure on accuracy in revision and new pieces of writing. *System, 81*, 135-145.

Swain, M. (2000). The Output Hypothesis and Beyond: Mediating Acquisition through Collaborative Dialogue. In J. P. Lantolf (Ed.), *Sociocultural Theory and Second Language Learning* (pp. 97-114). Oxford: Oxford University Press.

Tashakkori, A., & Creswell, J. W. (2007). The New Era of Mixed Methods. *Journal of Mixed Methods Research, 1*(1), 3-7.

Truscott, J. (1996). The Case Against Grammar Correction in L2 Writing Classes. *Language Learning, 46*(2), 327-369.

Truscott, J. (1999). The Case for "The Case Against Grammar Correction in L2 Writing Classes": A Response to Ferris. *8*(2), pp. 111-122.

Truscott, J. (2004). Evidence and conjecture on the effects of correction: a response to Chandler. *Journal of Second Language Writing, 13*, 337-343.

Truscott, J. (2007). The effect of error correction on learners' ability to write accurately. *Journal of Second Language Writing, 16*, 255-272.

Truscott, J. (2009). Arguments and appearances: A response to Chandler. *Journal of Second Language Writing, 19*, 59-60.

Truscott, J., & Hsu, A. Y.-p. (2008). Error correction, revision, and learning. *Journal of Second Language Writing, 17*, 292-305.

Van Beuningen, C. G., De Jong, N. H., & Kuiken , F. (2012). Evidence on the Effectiveness of Comprehensive Error Correction in Second Language Writing. *Language Learning, 62*, 1-41.

Van Beuningen, C. G., De Jong, N. H., & Kuiken, F. (2008). The Effect of Direct and Indirect Corrective Feedback on L2 Learners' Written Accuracy. *International Journal of Applied Linguistics, 156*, 279-296.

Zamel, V. (1983). The Composing Processes of Advanced ESL Students: Six Case Studies. *TESOL Quarterly, 17*(2), 165-187.

APPENDICES

Appendix 1

Pretest and Posttest Questions

Choose <u>one</u> of the topics below and write an opinion paragraph of 80 - 100 words.

1. "The Covid-19 pandemic has changed our lives." Do you agree with this statement?
2. "Distance education is better than face to face education." Do you agree with this statement?

Appendix 2

Writing rubric

TASK ACHIEVEMENT: 5 POINTS

ONE POINT	**TASK NOT ACHIEVED; too short** and/or **not** on topic, **no** topic sentence, **no** supporting sentences, **no** conclusion. Writing **may be** in **wrong** format.
1.5 POINTS	Writing has elements of a and b.
TWO POINTS	**TASK PARTLY ACHIEVED; almost right** length with **few** ideas on topic. Writing has a topic sentence, **poor** supporting sentences, **no** conclusion. Writing **may be** in **wrong** format.
2.5 POINTS	Writing has elements of b and c.
THREE POINTS	**TASK ACHIEVED; correct** length with **some simple** ideas on topic. Writing has a topic sentence and **simple** supporting sentences **although the paragraph is not well concluded**. Writing is in **correct** format.
3.5 POINTS	Writing has elements of c and d.
FOUR POINTS	**TASK ACHIEVED; correct** length with **some good** ideas on topic which are **well supported and well concluded**. Writing is in **correct** format.
4.5 POINTS	Writing has elements of d and e.
FIVE POINTS	**TASK ACHIEVED; correct** length with **original** ideas on topic which are **well supported and well concluded**. Writing is in **correct** format.

COHERENCE / COHESION: 5 POINTS

ONE POINT	**No** organizational features. Writing has **no cohesive devices** i.e. linking words, sequencers, referencing, etc.
1.5 POINTS	Writing has elements of a and b.
TWO POINTS	**Some** information/ideas but **not arranged coherently**. Writing has **inaccurate / repetitive / overused** cohesive devices.
2.5 POINTS	Writing has elements of b and c.
THREE POINTS	**Some** info/ideas with **some** organization and **overall** progression. Writing has **limited** cohesive devices used **accurately**, though there **may be some under/over use**.
3.5 POINTS	Writing has elements of c and d.
FOUR POINTS	Info/ideas arranged **correctly** with **clear** progression. **Sufficient** cohesive devices used **accurately**.
4.5 POINTS	Writing has elements of d and e.
FIVE POINTS	Info/ideas sequenced logically. A **wide range** of cohesive devices used **effectively**; **almost all aspects of cohesion** displayed i.e. linking words, sequencers, referencing, etc.

RANGE AND ACCURACY OF VOCABULARY FOR PURPOSE: 5 POINTS

ONE POINT	Writing contains **extremely limited** range of vocabulary and has **no control** of word formation or spelling.
1.5 POINTS	Writing has elements of a and b.
TWO POINTS	Writing contains **very limited** range of vocabulary **(basic words/repetition of words)** and has **very limited** control of word formation and spelling.
2.5 POINTS	Writing has elements of b and c.
THREE POINTS	Writing contains **limited** range of vocabulary and **may** have **noticeable** errors in word formation and spelling.
3.5 POINTS	Writing has elements of c and d.
FOUR POINTS	Writing contains **sufficient** range of vocabulary, **phrases and collocations** with **occasiona**l errors in word choice/formation and spelling.
4.5 POINTS	Writing has elements of d and e.
FIVE POINTS	Writing contains a **wide range of vocabulary**, **phrases and collocations** with **rare minor** errors.

RANGE AND ACCURACY OF GRAMMAR FOR PURPOSE: 5 POINTS

ONE POINT	Writing contains **no sentence** forms except for memorized phrases. **No** awareness of sentence formation and punctuation.
1.5 POINTS	Writing has elements of a and b.
TWO POINTS	Writing contains **limited range** of structures. **Basic SVO** sentence structure, punctuation is **faulty**.
2.5 POINTS	Writing has elements of b and c.
THREE POINTS	Writing contains **some variety** in structure use **with inconsistent accuracy**. **Basic SVO** sentence structure**.** Errors occur in punctuation.
3.5 POINTS	Writing has elements of c and d.
FOUR POINTS	Writing contains a **sufficient variety** of structures **mostly accurate**. **Good SVO** sentence structure. **Some** errors occur in punctuation.
4.5 POINTS	Writing has elements of d and e.
FIVE POINTS	Writing contains a **wide variety** of structures used **accurately** and **effectively**. **Good SVO** sentence structure. **Good control** of punctuation.

Appendix 3

Second Posttest Questions

Choose <u>one</u> of the topics below and write an opinion paragraph of 80 - 100 words.

1. “Having a part-time job has many advantages.” Do you agree with this statement?
2. “Your family is more important than your friends.” Do you agree with this statement?

Appendix 4

Online Survey Questions (the experimental group)

1) What do you think of the feedback procedure through error correction codes? Do you think it led to any improvements in your writing? If so, in what areas do you think it helped you improve your writing most? (Paragraph structure, Grammar, Vocabulary, Punctuation, Spelling etc.)

2) Are there any areas that you think you have not improved on?

3) How did you feel about trying to find out your errors and correct them through error correction codes?

4) Were there any error correction codes that you could not understand what type of an error they referred to? If so, what were they?

5) Were there any error correction codes that you understood what type of an error they referred to, but you had difficulty in revising?

6) How do you describe the feedback procedure with three adjectives?

Appendix 5

Online Survey Questions (the control group)

1) What do you think of the feedback procedure? Do you think it led to any improvements in your writing? If so, in what areas do you think it helped you improve your writing most? (Paragraph structure, Grammar, Vocabulary, Punctuation, Spelling etc.)

2) Are there any areas that you think you have not improved on?

3) How did you feel about being explicitly corrected?

4) How do you describe the feedback procedure with three adjectives?

Appendix 6

Interview Questions (the experimental group)

1) What do you think of the feedback procedure through error correction codes? Do you think it led to any improvements in your writing? If so, in what areas do you think it helped you improve your writing most? (Paragraph structure, Grammar, Vocabulary, Punctuation, Spelling etc.)

2) Are there any areas that you think you have not improved on? If there are any, why do you think you have not improved on them?

3) How did you feel about trying to find out your errors and correct them through error correction codes?

4) Were there any error correction codes that you could not understand what type of an error they referred to? If so, what were they?

5) Were there any error correction codes that you understood what type of an error they referred to, but you had difficulty in revising?

Appendix 7

Interview Questions (the control group)

1) What do you think of the feedback procedure? Do you think it caused improvements in your writing? If so, in what areas do you think it helped you improve your writing most? (Paragraph structure, Grammar, Vocabulary, Punctuation, Spelling etc.)

2) Are there any areas that you think you have not improved on? If there are any, why do you think you have not improved on them?

3) How did you feel about being explicitly corrected?

Appendix 8
Error Correction Codes

SYMBOL	KIND OF ERROR	EXAMPLE	CORRECT SENTENCE
S	Spelling	The woman has **eihgt** children.	The woman **has eight** children.
P	Punctuation	We go to school everyday	We go to school everyday**.**
C	Capitalization	I love to speak **english.**	I love to speak **English.**
WW	Wrong Word	He lives **at** Vista.	He lives **in** Vista.
__	Add a Word	You should speak to__teacher.	You should speak to **the** teacher.
X	Take Out a Word	Do you go to **the** work?	Do you go to work?
SV	Subject - Verb Error	She **live** in a big house.	She **lives** in a big house.
VT	Verb Tense	I **go** to the beach yesterday.	I **went** to the beach yesterday.
S/P	Singular / Plural	**Three boy** went to school.	**Three boys** went to school.
?	Unclear Meaning	**She likes very today.**	**She likes to play tennis every day.**
WO	Word Order	They went **yesterday to school.**	They went **to school yesterday**.
WF	Wrong Form	She is a **beauty** woman.	She is a **beautiful** woman.
‿	Going Together	I **some times** attend classes.	I **sometimes** attend classes.
LC	Lower case letter	Why **İs** he there?	Why **is** he there?
IF	Informal	You **wanna** play with us.	You **want to** play with us.