

Written corrective feedback, individual differences and second  
language acquisition of the English passive voice

Saeed Roshan

A thesis submitted to Auckland University of Technology  
in fulfilment of the requirements for the degree of  
Doctor of Philosophy (PhD)

2017

School of Language and Culture

***Dedication***

*To my father and my father-in-law who were eagerly awaiting the completion of this thesis but did not live to see it.*

***And***

*To my mum and my wife*

# TABLE OF CONTENTS

<b>DEDICATION</b>	ii
<b>TABLE OF CONTENTS</b>	iii
<b>LIST OF TABLES</b>	vi
<b>LISTS OF FIGURES</b>	viii
<b>LIST OF ABBREVIATIONS</b>	ix
<b>ABSTRACT</b>	x
<b>ATTESTATION OF AUTHORSHIP</b>	xii
<b>ACKNOWLEDGEMENTS</b>	xiii

## CHAPTER 1: INTRODUCTIN

1.1 Introduction	1
1.2 Can written CF facilitate L2 development immediately and overtime?	2
1.2.1 Is revision following feedback evidence of L2 development?	4
1.3 Are some types of written CF more effective in treating error types than others?	6
1.4 Is written CF more effective for some linguistic forms/structures than others?	8
1.5 Does working memory moderate the effectiveness of written CF?	9
1.6 Conclusion	10
1.7 The overall design and focus of the present thesis	11
1.8 Thesis outline	12

## CHAPTER 2: THEORETICAL EXPLANATION OF THE EFFECTIVENESS OF WRITTEN CF FOR L2 DEVELOPMWNT

2.1 Introduction	14
2.2 An overview of L2 development	16
2.3 Skill-based theories of second language acquisition	17
2.3.1 McLaughlin's skill acquisition model	18
2.3.2 Anderson's adaptive control theory	19
2.4 Types of written CF as input	20
2.5 Cognitive processing of input for explicit declarative knowledge	21
2.6 Theoretical accounts of the efficacy of different types of written CF	23
2.7 Theoretical accounts of the efficacy of revision following feedback on L2 development	25
2.8 Theoretical explanations of the efficacy of focused and unfocused written CF on L2 development	27
2.9 Theoretical explanations of the differential efficacy of written CF types for treating different types of linguistic error	28
2.10 The potentially moderating influence of working memory on cognitive processing	30
2.11 Summary	33

## CHAPTER 3: A REVIEW OF WRITTEN CORRECTIVE FEEDBACK STUDIES

3.1 Introduction	35
3.2 Research on whether or not written CF can facilitate L2 development	37
3.2.1 The efficacy of written CF on a revised version of a text	38
3.2.2 The efficacy of written CF on new pieces of writing	41
3.2.2.1 Studies on unfocused new texts	41
3.2.2.2 Studies on focused new text	47
3.2.2.3 Studies comparing focused and unfocused written CF	49

3.3 The relative effectiveness of different types of written CF	51
3.3.1 Studies comparing direct CF and less explicit types of written CF	52
3.3.2 Studies comparing direct CF and direct CF with more explicit types of written CF	53
3.3.3 Studies comparing metalinguistic explanation and other types of written CF	56
3.4 The effect of feedback on different grammatical structures	61
3.5 Individual Factors and Written Corrective Feedback	66
3.5.1 Studies on working memory, phonological short-term memory and oral corrective feedback	68
3.6 Rationale of the present study and research questions	70

## CHAPTER 4: METHODOLOGY

4.1. Introduction	73
4.2 Scope of the Research	74
4.3 Philosophical Approach	74
4.4 Study design	75
4.5 Context and Participants	77
4.5.1 Context	77
4.5.2 Participants	78
4.6 Target Structure	78
4.7 Instruments	80
4.7.1 Student Writing Task	81
4.7.2 Reading Span Test	83
4.7.3 Phonological Short-term Memory	84
4.8 Treatment Procedure	85
4.9 Data Analysis	89
4.9.1 Writing Tasks	90
4.9.1.1 Scoring	90
4.9.2 Statistical analyses	94
4.9.2.1 Statistical analyses for writing tasks	94
4.9.2.2 Statistical analyses for reading-span test	96
4.9.2.3 Statistical analyses for non-word-span test	98
4.10 Validity and reliability	98
4.11 Ethical Approval	99

## CHAPTER 5: Results

5.1 Introduction	101
5.2 The findings for RQ1: What effect does focused DCF and ME with and without revision have on learners' use of the English passive voice in an immediate text revision and in new texts overtime?	102
5.3 The findings for RQ2: Does the opportunity for the revision influence the efficacy of DC and ME?	109
5.4 The findings for RQ3: Is there any difference in the effect of DC and ME regardless of whether there is an opportunity for revision?	116
5.5 The findings for RQ 4: To what extent do working memory (WM) and phonological short-term memory (PSTM) moderate the effectiveness of focused DCF and ME with and without revision in new texts over time?	122
5.5.1 Correlation and Regression Findings for Experimental Groups	123
5.5.2 Correlation and Regression Findings for DMWR (direct CF and metalinguistic explanation) and DMPR (direct CF plus revision and metalinguistic explanation plus revision)	125

5.5.3 Correlation and Regression Findings for DCO (direct CF and direct CF plus revision) and MEO (metalinguistic explanation and metalinguistic explanation plus revision)	128
5.6 Conclusion	131

## **CHAPTER 6: Discussion**

6.1 Introduction	134
6.2 RQ1: What effect does focused DCF and ME with and without revision have on learners' use of the English passive voice in an immediate text revision and in new texts over time?	134
6.2.1 Discussion of within group comparison of findings immediately and over time	135
6.2.2 Discussion of between group comparison of findings immediately and over time	136
6.2.2.1 Discussion comparing the effectiveness of the treatment groups and the control group immediately and over time	137
6.2.2.2 Discussion comparing the difference between individual treatment groups immediately and over time	138
6.2.3 The findings on the impact of written CF on the passive voice as a complex structure	141
6.3 RQ2: Does the opportunity for revision influence the efficacy of DC and ME?	143
6.3.1 Discussion of the within group comparison of findings immediately and over time	143
6.3.2 Discussion of between group comparison of findings immediately and over time	146
6.4 RQ3: Is there any difference in the effect of DC and ME regardless of whether there is an opportunity for revision?	148
6.5 RQ4: To what extent do working memory and phonological short-term memory moderate the effects of different types of feedback?	151
6.6 Conclusion	157

## **CHAPTER 7: Conclusion**

7.1 Introduction	159
7.2 The aims of the study	159
7.3 Summary of key findings	161
7.4 Contributions of the study	164
7.4.1 Contributions to new empirical knowledge	166
7.4.2 Contributions to theory	166
7.4.2.1 Theoretically, the value of explicit input, in the form of written CF, for L2 development	167
7.4.2.2 The value of revision for improving the effectiveness of written CF	172
7.4.2.3 Contributions of written CF to the development of a complex linguistic structure	173
7.4.2.4 The moderating effect of working memory in written CF	177
7.4.3 Contributions to research methodology advancement	181
7.4.4 Contributions to pedagogy	183
7.5 Limitations of the study	185
7.6 Recommendations for future research	188

<b>REFERENCES</b>	190
-------------------	-----

<b>APPENDIXES</b>	
<b>Appendix A</b>	203
<b>Appendix B</b>	204
<b>Appendix C</b>	207
<b>Appendix D</b>	209
<b>Appendix E</b>	211

## LIST OF TABLES

Table 3.1 Studies of written CF on the effectiveness of text revision	38
Table 3.2 Studies of written CF on the effectiveness of text revision and new text writing	38
Table 3.3 Early unfocused studies on the effectiveness of written CF on new texts	41
Table 3.4 Recent studies on unfocused written CF	43
Table 3.5 Recent focused studies investigating the efficacy of written CF in new texts	47
Table 3.6 Studies comparing direct error correction and less explicit types of written CF	51
Table 3.7 Studies comparing direct error correction and direct error correction plus more explicit forms of written CF	52
Table 3.8 Studies comparing metalinguistic feedback and other types of written CF	55
Table 3.9 The efficacy of written CF for developing accuracy in the use of linguistic form and structure	61
Table 4.1 Study Design	74
Table 4.2 Regular Past Participle	90
Table 4.3 Irregular Past Participle	90
Table 4.4 Irregular Past Participle – No Change in Form	91
Table 4.5 Cohen’s Guideline to Interpret Partial Eta Squared and Cohen’s d	93
Table 5.1 Descriptive statistics for the accuracy scores in the written tasks	99
Table 5.2 Descriptive statistics of gain scores for the accuracy scores in the written tasks	102
Table 5.3 Effect sizes in the form of Cohen’s d for the accuracy scores between groups	103
Table 5.4 Effect sizes in the form of Cohen’s d for the accuracy scores overtime	105
Table 5.5 Descriptive statistics for the accuracy scores in the written task	107
Table 5.6 Descriptive statistics of gain scores for the accuracy scores in the written tasks	108
Table 5.7 Effect sizes in the form of Cohen’s d for the accuracy scores between groups	110
Table 5.8 Effect sizes in the form of Cohen’s d for the accuracy scores overtime	112
Table 5.9 Descriptive statistics for the accuracy scores in the written task	113
Table 5.10 Descriptive statistics for the gain scores in the written tasks	115
Table 5.11 Effect sizes in the form of Cohen’s d for the accuracy scores between groups	116
Table 5.12 Effect Sizes in the form of Cohen’s d for the accuracy scores overtime	118

Table 5.13 Descriptive statistics of the feedback groups' working memory and phonological short-term memory	120
Table 5.14 Correlation results for experimental groups	121
Table 5.15 Regression results for the effects of experimental groups and contributions of WM and PSTM	122
Table 5.16 Descriptive statistics of the feedback groups' working memory and phonological short-term memory	123
Table 5.17 Correlation results for experimental groups	124
Table 5.18 Regression results for the effects of DMWR and DMPR and contributions of WM and PSTM	125
Table 5.19 Descriptive statistics of the feedback groups' working memory and phonological short-term memory	126
Table 5.20 Correlation results for DCO and MEO	126
Table 5.21 Regression Results for the Effects of DCO and MEO and Contributions of WM and PSTM	127

## LIST OF FIGURES

Figure 2.1 Development of the working memory model (Baddeley, 2000)	32
Figure 5.1 Group means of different types of feedback over time	103
Figure 5.2 Group means of different types of feedback with gain scores	105
Figure 5.3 Group means of DMWR (direct CF and metalinguistic explanation) and DMPR (direct CF plus revision and metalinguistic explanation plus revision) feedback over time	110
Figure 5.4 Group means of DMWR (direct CF and metalinguistic explanation) and DMPR (direct CF plus revision and metalinguistic explanation plus revision) feedback with gain scores	112
Figure 5.5 Group means of DCO (direct CF and direct CF plus revision) and MEO (i.e., Figure metalinguistic explanation and metalinguistic explanation plus revision) feedback over time	117
Figure 5.6 Group means of DCO (direct CF and direct CF plus revision) and MEO (i.e., metalinguistic explanation and metalinguistic explanation plus revision) feedback over time	118



## **LIST OF ABBREVIATIONS**

CF: Corrective Feedback

DC: Direct CF

EFL: English as a Foreign Language

ESL: English as a Second language

ME: Metalinguistic explanation

PSTM: Phonological short term memory

WM: Working memory

# ABSTRACT

The extent to which written CF plays a role in learners acquiring the target language is a question that has received a lot of attention over the last 20-30 years. This thesis, by drawing on a cognitive view, continued with that focus, exploring not only the efficacy of written CF on the improved accuracy of learners but also the extent to which working memory and phonological short-term memory may moderate the effects of different types of feedback.

The study was undertaken with 100 university students in Iran. Firstly, a quasi-experimental study was used, with a pre-test, treatment, immediate and delayed post-tests, to investigate the effectiveness of four types of written CF (direct CF, direct CF plus revision, metalinguistic explanation, metalinguistic explanation plus revision) on a complex linguistic structure, the English passive voice. Additionally, the learners' working memory was measured using a reading span test using DMDX software and their phonological short term memory was measured using a non-word span test.

The findings showed that a single episode of written CF improved accuracy immediately and over time (after two weeks) for all the experimental groups, but not for the control group. They also revealed that direct CF relatively led to more improved accuracy compared to the other groups. Additionally, the results showed that the non-revision groups (i.e., direct CF and metalinguistic explanation) had greater accuracy in the immediate post-test, but that the accuracy of the revision groups (i.e., direct CF plus revision and metalinguistic plus revision) was retained over a longer period of time.


The results regarding the moderating effect of working memory and phonological short term memory revealed that (1) working memory moderated the impact of the metalinguistic explanation and combined metalinguistic explanation groups (i.e. metalinguistic explanation and metalinguistic explanation plus revision) and the combined metalinguistic and direct CF groups both immediately and over time; and (2) working memory moderated the direct CF plus revision and combined revision groups (metalinguistic explanation plus revision and direct CF plus

revision) only in the long term. Additionally, phonological short-term memory negatively moderated the impact of direct CF plus revision only in the long term.

These results support the role that explicit knowledge in the context of writing can play in L2 development because the learners sustained the explicit knowledge gained from written CF on a complex structure (e.g. the English passive voice) in new texts over time. Additionally, the findings suggest that more explicit types of written CF may be more effective than less explicit types of written CF for targeting a complex structure in a single processing episode.

## **Attestation of Authorship**

I hereby declare that this submission is my own work and that, to the best of my knowledge and belief, it contains no material previously published or written by another person (except where explicitly defined in the acknowledgements), nor material which to a substantial extent has been submitted for the award of any other degree or diploma of a university or other institution of higher learning.

Saeed Roshan: 

Date: 20.10.2017

## Acknowledgments

The completion of this PhD thesis would not have been possible without the support and help from the following people.

First and foremost, I would like to express my sincere gratitude to my primary supervisor, Prof. John Bitchener, for his unending patience, support and encouragement. As an excellent researcher and teacher, John provided me with his insightful comments and constructive recommendations throughout this thesis. I have learned so much from his extensive expertise and conscientious supervision and this knowledge will surely continue to benefit me in my future career.

I would like to express my deepest gratitude to my second supervisor Dr Shaofeng Li from the University of Auckland. His timely detailed feedback and encouragement helped me keep pace and motivated me during this journey.

My special thanks also goes to Dr Jonathan Forster at the University of Arizona in the USA for his insightful guidance on writing a program for the DMDX system to measure the participants' working memory.

Moreover, I would like to express my heart-felt gratefulness to Annelies Roskvist, the deputy head of the School of Language and Culture, and to Elizabeth Brugh, the programme leader of the Intensive Literacy and Numeracy programme at AUT, for their kind support throughout my PhD journey and for providing me with research and teaching assistant jobs. This has provided me with valuable experience in the context of New Zealand.

I would like to acknowledge the Auckland University of Technology Ethics Committee (AUTEC) for approving this thesis (Ethics Application Number 15/184).

I would also like to thank Jeremy Cooper for the recording of his voice for the PSTM test and for being a real friend over this journey.

Last but not least, my thanks goes to my family. First, a special thanks goes to my mum who has been missing me during this journey, especially after the death of my dad. Her kindness, unending patience and encouragement have been valuable during my PhD journey.

I am greatly indebted to my wife, Zahra, who has always supported me. Zahra has been working hard both inside and outside the home to support me and allow me to focus on my study. Even though she had a very hard time after the death of her dad, she has kindly and patiently stayed in New Zealand to support me to finish this journey. This thesis would not have been completed without her generous support. Thank you very much for supporting me to achieve my dream.

Finally, I would like to thank my sisters for their encouragement and support. My special thanks goes to my kind sister, Elahe, for taking care of mum while I have been in New Zealand and for her unending support.

# CHAPTER 1

## INTRODUCTION

### 1.1 Introduction

The contribution of written CF to second language (L2) development has been the focus of a growing number of studies over the last 20 years. So far, written CF research has given much time to investigate whether a learner's accuracy improves after receiving written CF, and whether the effectiveness of written CF varies due to different types of written CF and differences in the error types that are targeted. Learners' improved accuracy after the provision of focused and unfocused written CF has also been stated in new writing texts immediately and over time in written CF studies (e.g., Bitchener, 2008; Rummel & Bitchener, 2015; Van Beuningen, De Jong, & Kuilken 2008, 2012).

However, the extent to which written CF plays a role in learners' acquisition of a target language is a problem that researchers and teachers of second language writing and acquisition have sought to address. For instance, questions remain as to whether focused written CF facilitates L2 development immediately and over time and whether certain types of written CF are more effective in improving accuracy than other types. Also requiring further investigation is whether written CF is useful for treating complex structures such as the English passive voice, because to date most of the research has focused on simple linguistic structures such as English articles and the past tense. Further questions have to do with revision, and whether it may have an impact on improved accuracy in the short and long term along with whether learners' individual differences (especially in working memory) may moderate the effectiveness of written CF types.

In order to address these areas where questions remain, the present study investigates the potential role of written CF within a cognitive view. Within the cognitive framework

developed by Gass (1997) written CF is considered as a form of input, which may be noticed, attended to, and processed. As such, it is likely to facilitate the development of *explicit knowledge*, or the type of knowledge that learners can access through controlled and conscious processing and that they typically achieve in an instructional/educational context. Furthermore, written CF is likely to draw a learner's attention to his/her stored explicit knowledge, and through controlled processing of such knowledge, the knowledge may be consolidated. Additionally, a learner's cognitive variables may affect the efficacy of written CF. Therefore, written CF type, linguistic error type, and individual differences (e.g., working memory) may potentially moderate the effectiveness of written CF (Bitchener, 2012). Thus, by drawing on cognitive processing, this thesis not only examines the efficacy of written CF on L2 development, but also investigates the potentially moderating effect of (1) text revision following feedback, (2) written CF type, (3) linguistic error type, and (4) individual differences in working memory.

Using both theory and empirical research respectively, the following section explains the potential of written CF to facilitate L2 development immediately and over time. It also explains why the efficacy of written CF may vary as a result of text revision following feedback, error type, feedback type, and a learner's working memory.

## **1.2 Can written CF facilitate L2 development immediately and overtime?**

This section presents a theoretical and empirical argument for why written CF can facilitate L2 development immediately and over time. In order to theoretically address this important question, cognitive information processing theories (Gass 1997; Anderson, 1983; MacLaughlin, 1990) are combined to illuminate the potential role of written CF in facilitating explicit knowledge. In her framework of theories, Gass (1997) reported that an episode of cognitive information processing begins when the learners notice input.



Written CF, as a form of input, has the potential to make the learners notice any difference between their existing knowledge and the information provided by the written CF. When learners understand a linguistic form/structure, it is processed in the short term memory and matched against their existing knowledge. If the written CF input is similar to their existing hypothesis, the hypothesis is confirmed, but if it is different from their existing hypothesis, it is rejected. When learners reject a hypothesis, they have an opportunity to make a new hypothesis and when further input is provided, they can confirm or reject the input. At this stage, learners may restructure or modify the new hypothesis. If learners do not accurately produce the target structure, they can be provided with modified written CF input and the initial episode will need to be repeated. If learners produce accurate output on one occasion, they are then able to consolidate the new explicit knowledge over time and with more practice by retrieving the stored information from the long-term memory and by producing the accurate form/structure. Each time learners, through practice, retrieve the new explicit knowledge from their long-term memory and process it in their working memory, it becomes more likely that the new information will be consolidated. Thus, learners are likely to be able to convert their explicit knowledge through *practice* to implicit knowledge. Skill acquisition theory (Anderson, 2000; McLaughlin, 1990) will be drawn upon to explain how the conversion occurs (see Chapter 2, Section 2.3).

As discussed above, cognitive information processing theories reveal how written CF can facilitate L2 development. The effectiveness of written CF has also been empirically investigated. In most studies, learners' output that was produced when writing a new text over time was investigated. Accuracy performance in the written output was measured employing a pre-test, written CF treatment, immediate post-test, and delayed post-test design. To determine whether there was improvement, learners' accuracy performance prior to the provision of written CF on linguistic errors in the pre-test task was compared

with accuracy performance in the post-test tasks after written CF had been provided. If learners showed a significant improvement in accuracy between the pre-test and post-tests, and if this was significantly higher than that of the control group who had not been provided with written CF treatment, it was understood that written CF had facilitated the improvement in accuracy. If the improvement in accuracy was maintained over time, then it was understood that the knowledge had been consolidated. A number of studies have indicated that written CF can improve accuracy over time. These studies have intensively targeted a narrow set of simple rule-based categories such as English articles and the past tense (e.g., Bitchener & Knoch, 2008; Frear, 2012; Sheen, 2007a). In simple rule-based categories, the error is a result of the lack of knowledge of one aspect of the feature. Simple rule-based categories are a good place to start if they are a problematic area, because when a target structure is focused, simple and rule-based it is easy to learn (Bitchener, 2016). However, so far, there are only two studies that have targeted complex structures: Shintani, Ellis and Suzuki (2014) and Rummel (2014) targeted the hypothetical conditional and present perfect tense respectively.

Therefore, there is a need for further research on the effectiveness of written CF on the improved accuracy of complex linguistic error types (Bitchener & Ferris, 2012; Bitchener & Storch, 2016). Thus, this thesis examines the efficacy of different types of written CF for improving the accuracy of the English passive voice in new texts over time. In order to address this key question, one issue that is investigated is the extent to which learners' output as a result of revision may facilitate L2 development. This issue is discussed in the following section.

### **1.2.1 Is revision following feedback evidence of L2 development?**

This section provides a theoretical and empirical overview of the extent to which revision following feedback may facilitate L2 development. Theoretically, providing learners with

an opportunity to revise a text plays an important role in the development process. This opportunity *pushes* learners to attend to the feedback they have been provided with; that is, they “notice a gap between what they want to say and what they can say, leading them to recognize what they do not know, or know only partially” (Swain, 1995, pp. 125-126)]. They then process this feedback across the cognitive processing stages as determined by Gass (1997), before making a hypothesis on an accurate modification. In particular, when learners are pushed, they go from semantic (i.e., meaning) to syntactic (i.e., form) processing. In other words, Swain argued that *pushed output* (i.e. noticing a problem) can encourage learners to modify their output (Swain & Lapkin, 1995) and may contribute to the learners noticing the grammatical forms.

Although revision may facilitate L2 development, Truscott (2007) argued that a learners’ improvement in producing an accurate revision of a text does not mean that they are able to produce these target forms in a new piece of writing. Truscott’s claim may be true because there is a distinction between revision and writing a new text. In revision, learners are generally only required to focus on and revise the errors in the same linguistic context while in writing a new text, learners may be required to undertake deep processing and focus both on form and meaning in a new linguistic context. This is especially so when learners are provided with a type of written CF (e.g., direct CF) where they only need to copy the correct form in the revised text. Thus, the effectiveness of written CF – either with revision or without revision – needs to be investigated through the writing of a new text.

A number of studies (e.g., Truscott & Hsu, 2008; Van Beuningen et al., 2008, 2012) have examined whether revision results in improved accuracy in writing new texts. Truscott and Hsu (2008) reported that the increased accuracy revealed by their experimental group in the revision of their texts was not shown in their writing of new texts. However, the findings of Van Beuningen et al. in two different studies (2008, 2012) contradicted the

findings of Truscott and Hsu (2008). The main difference between Van Beuningen et al.'s (2008, 2012) and Truscott and Hsu's (2008) studies is the degree of explicitness of the written CF, and this may have led to the contradictory findings. In contrast to Truscott and Hsu's (2008) study in which one type of written CF with a low level of explicitness (underlining) was employed, in Van Beuningen et al.'s (2008, 2012) study the researchers used two more explicit written CF types (direct correction and error codes). This may indicate that the degree of explicitness of written CF can impact the effectiveness of written CF. Therefore, these conflicting findings reveal a need for more studies to show if revising a text leads to improved accuracy in a new piece of writing. Thus, this is one of the issues the present study investigates.

### **1.3 Are some types of written CF more effective in treating error types than others?**

As discussed in the previous section, theoretically, written CF has the potential to facilitate L2 development. However, its effectiveness may differ according to the different types of written CF provided (Bitchener, 2012). In the literature, there are three main types of written CF: direct written CF, indirect written CF, and metalinguistic written CF. *Direct written CF* provides students with the corrected forms of their linguistic errors through crossing out, rewriting, or an addition near or above the linguistic error. In *indirect CF*, students' attention is drawn to the position of their errors by the circling and underlining of an error, but it is left to the students to work out the correction needed. The third form of written CF, *metalinguistic CF*, uses error codes to indicate the type of error a learner has made or, alternatively, it numbers errors and then provides students with a brief explanation, with or without examples, for the errors at the bottom of the page or at the end of the full text. Metalinguistic CF can also be in the form of a

handout to students that includes an explanation of the target structure (Shintani, Ellis, & Suzuki, 2014).

These written CF types have different degrees of explicitness; therefore, their effectiveness on L2 development may vary. For instance, indirect written CF, as a less explicit type of written CF, may enable learners to notice the gap in their existing knowledge. Those who favour indirect written CF suggest that it is the most useful because it engages learners in problem solving and guided learning (Lalande, 1982). In other words, it requires learners to do the work. However, it is unlikely to enable learners, especially those with low limited linguistic memory store, to form a new hypothesis about acceptable targeted linguistic feature use. On the other hand, more explicit types of feedback (e.g. direct CF) are more likely to help learners to form a correct hypothesis and thus may better facilitate L2 development. Additionally, more explicit types of feedback (e.g. direct CF) may reduce the confusion that learners may experience if they do not understand a less explicit type of CF (e.g. metalinguistic explanation in the form of a handout) (Bitchener & Ferris, 2012).

This study has compared the efficacy of a more explicit and a less explicit type of written CF on L2 development. Because these types of written CF provide learners with differing degrees of explicitness, this is an area that requires more investigation (Bitchener & Storch, 2016). In this study, direct CF is considered as more explicit because it provides learners with the correct form of the erroneous form of the passive voice as the target structure. Metalinguistic explanation, on the other hand, is less explicit because it is provided in the form of a handout to students that includes an explanation and examples of the use of the passive voice without identifying the errors in their writing text.

Only two recent studies (Shintani & Ellis 2013; Shintani, Ellis & Suzuki, 2014) have investigated the efficacy of direct written CF and metalinguistic explanation in the form

of a handout to students. These studies reported mixed findings on the effectiveness of the written CF types. The study by Shintani and Ellis (2013) reported that metalinguistic explanation was effective in improving the accuracy of new texts over time, while Shintani, Ellis, and Suzuki (2014) found that direct written CF was effective. The varied findings show a need for further studies on the efficacy of direct CF and metalinguistic explanation in the form of a handout to learners. Thus, this study will investigate the efficacy of direct CF and metalinguistic explanation on L2 development.

#### **1.4 Is written CF more effective for some linguistic forms/structures than others?**

The discussion in Section 1.2.1 focused on whether written CF has the potential to facilitate L2 development and noted that the extent to which it can effectively improve different types of linguistic errors may differ. Theoretically, the reason for this is that morphological, syntactic, and lexical errors represent different domains of linguistic knowledge (Bitchener & Knoch, 2008; Ellis, 2008; Ferris, 1999; Ortega, 2009; Truscott, 2007) and learners may need to focus their attention on more than just one linguistic element each time they make a hypothesis on the use of the correct linguistic form for a certain linguistic error type. Additionally, it has been argued (Young, 1996) that one linguistic form or structure may be more difficult to learn than another and that different linguistic forms and structures may be learnt at different cognitive stages (Pienemann, 1998). Thus, some linguistic forms and structures may be more ‘treatable’ than others (Ferris, 1999). Further, it has been argued (e.g., Ferris, 2002, 2003) that rule-based error types may be more ‘treatable’ than item-based errors. Rule-based errors are those that occur in a rule-governed way and for example can be corrected by referring to grammar books (e.g., the regular simple past tense), while item-based errors are those where there is no set of rules students can refer to (e.g., the irregular simple past tense). In other words,

these errors are less rule-governed so grammatical rules are therefore less likely to be useful in resolving them. In these cases, there is a need for specific knowledge of the targeted language.

Although it is likely that rule-based errors are more treatable than item-based errors, it is important to consider the complexity of the rule-based forms/structures. For instance, the rules for employing the simple past tense are relatively straightforward; however, the rules for some structures, such as the English passive voice, can be difficult to acquire as these are often complex. Because they are both semantically and syntactically complex learners are required to use more attentional capacity when employing them. Even though there is a growing amount of research that has examined the effect of written CF on targeting more rule-based linguistic error categories (e.g, Bitchener, 2008; Bitchener & Knoch, 2008; Frear, 2012; Sheen, 2007a Shintani & Ellis 2013), only two written CF studies have examined the efficacy of written CF on complex structures (Rummel, 2014; Shintani et al., 2014), so there is a need for more investigation before any generalized conclusion can be drawn. Thus, to fill the gap, this thesis will investigate the effectiveness of written CF in the use of the English passive voice as the target structure.

### **1.5 Does working memory moderate the effectiveness of written CF?**

Although many studies have focused on the efficacy of different types of written CF (e.g., Shintani & Ellis, 2013; Shintani et al., 2014), the mixed findings from these studies have revealed that other factors, such as individual factors, may impact the efficacy of written CF. Thus, scholars have called for an investigation into the moderating effect individual factors have on how learners respond to and use the written CF they receive (Bitchener & Storch, 2016; Bitchener & Ferris, 2012; Ellis, 2008). Working memory, as a cognitive individual factor, plays an important role in consolidating explicit knowledge in both the

initial single episode and in subsequent episodes of cognitive processing (Bitchener, 2016). Learners store and process the new information by drawing on working memory; that is, they can control both the production of meaning and a suitable form/structure of the new target structure and retrieve the newly integrated knowledge from their long-term memory. Additionally, Engle (2002) argued that working memory includes the ability to control attention in order to keep information in an active, promptly retrievable state; it is not just about individual differences and the working memory's storage capacity. Thus, it has been supported that working memory can play a deep role in cognitive activities such as attention and noticing.

Several studies have focused on the moderating effect of individual differences in working memory for oral corrective feedback (e.g., Li, 2013; 2014; Mackey & Sachs, 2012; Mackey, Adams, Stafford & Winke, 2010; Re'v'e'sz, 2012). However, no study has investigated the moderating effect of working memory on written CF. Thus, there is a need to investigate if working memory, as a cognitive factor, moderates the efficacy of written CF. Kormos (2012) argued that working memory may moderate how learners learn from different written CF types. She argued that in contrast to the oral context, opportunities to learn through feedback in the writing context are less restricted by time pressure; however, because writing learners are dependent on their working memory capacity, they may respond differently to feedback. Therefore, to fill the gap, the present study will investigate the moderating effect of working memory on written CF.

## **1.6 Conclusion**

In sum, the previous sections have revealed the areas where questions remain and need further investigation. To fill the gaps, this study aims to address the following questions: RQ1 investigates the effect of written direct corrective feedback and metalinguistic



explanation with and without revision on learners' immediate and delayed output in relation to the passive voice as the target structure in new writing tasks over time. RQ2 explores whether the opportunity for revision influences the effectiveness of direct corrective feedback and metalinguistic explanation. In order to address RQ2, the two revision groups (i.e., direct corrective plus revision and metalinguistic explanation plus revision) and the two groups that did not make revisions (i.e., direct corrective feedback and metalinguistic explanation) were combined and compared. RQ3 examines the relative efficacy of direct corrective feedback and metalinguistic explanation regardless of whether there is an opportunity for revision. In order to address RQ3, the two direct CF groups (i.e., direct corrective feedback plus revision and direct corrective feedback) and the two ME groups (i.e., metalinguistic explanation plus revision and metalinguistic explanation) were combined and compared. RQ4 investigates the extent to which working memory and phonological short-term memory moderate the effectiveness of direct corrective feedback and metalinguistic explanation with and without revision in new writing texts over time.

### **1.7 The overall design and focus of the present thesis**

This study employed a pre-test, treatment, post-test, and delayed post-test design, using intact English as a Foreign Language (EFL) classes. Many experimental studies utilize a pretest-posttest design because a pre-test is administered prior to the experimental manipulation, and a post-test following the manipulation. Accordingly, a pretest-posttest design allows the researcher to assess the impact of the experimental manipulation by observing the difference between the pre-test and post-test.

One hundred Iranian EFL university students participated in this quantitative study and were assigned to four experimental groups (direct corrective feedback, direct corrective

feedback and revision, metalinguistic explanation, metalinguistic explanation and revision) and one control group. The efficacy of one session of providing written CF on the use of the passive voice was investigated immediately and over time (four weeks). One week prior to the start of the CF treatment, participants performed a writing task as the pre-test. The immediate post-tests (revision and new test) were conducted immediately after the CF treatment session had been completed in Week 2. The delayed post-test was completed in Week 4. In order to investigate if working memory may moderate the effectiveness of written CF, participants completed a working memory test (reading span test) and a phonological short-term memory test (non-word span test) in Week 3.

## **1.8 Thesis outline**

This thesis consists of seven chapters. As an introduction, Chapter 1 has provided an overview of the thesis. Chapter 2 reviews cognitive information processing theories and identifies what role written CF might play and what factors may affect the efficacy of written CF in terms of cognitive processing. Chapter 3 critically reviews existing written CF studies that focus on cognitive information processing in order to identify the contribution of these studies to L2 development. Gaps and limitations in these studies are also highlighted. The chapter also reviews studies on working memory and CF. Finally, the research questions are presented. Chapter 4 discusses the methodological approach used in this research. Additionally, it describes the methods of data collection and data analysis of the writing tasks in relation to working memory and phonological short-term memory. Chapter 5 presents the findings of the statistical analysis of the quantitative data. Chapter 6 provides a discussion and interpretation of the findings in relation to the previous relevant literature and research questions. Chapter 7 concludes the study by

summarizing the findings and discussing their contribution to theory, research, and pedagogy. The limitations of the study are also discussed and suggestions for further studies are provided.

## **CHAPTER 2**

### **THEORETICAL EXPLANATION OF THE EFFECTIVENESS OF WRITTEN CF FOR L2 DEVELOPMENT**

#### **2.1 Introduction**

This chapter has two main aims. The first aim is to provide a theoretical overview of why we might expect written CF (especially direct CF, direct plus revision, metalinguistic explanation, metalinguistic plus revision) to facilitate L2 development immediately and over time. As L2 development is a key term in the study, the chapter begins with an explanation of the construct and its difference from L2 learning and L2 acquisition. The chapter then refers to the information processing theories, such as those by Anderson and McLaughlin, to argue that through meaningful, contextual practice (e.g., in the form of written CF) over time, explicit knowledge can be converted to implicit knowledge. In other words, through written CF (especially direct CF, direct plus revision, metalinguistic explanation, metalinguistic plus revision) on the target language (for instance, the English passive voice) learners may convert their explicit/declarative knowledge to an automatic use of implicit/procedural knowledge. Following this, by drawing on Gass's (1997) framework for the stages in the cognitive processing of explicit input, it is argued how information arising from written CF (especially direct CF, direct plus revision, metalinguistic explanation, metalinguistic plus revision), as explicit input, can be processed, produced as modified output and retained long term. Additionally, in order to understand the nature of the input that learners are provided with during the conversion of explicit knowledge to implicit knowledge, two types of input are explained: positive and negative input. Because written CF is a form of negative input, a theoretical discussion is provided on written CF and its types. This is followed by a further discussion of the theoretical arguments relating to the efficacy of different types of written CF for

L2 development. Then, by drawing on Swain's (1985) Comprehensible Output Hypothesis, the chapter argues how revision following written CF leads to "*pushed output*" (Swain & Lapkin, 1995) and consequently may lead to L2 development.

Revision and written CF types (especially direct CF, direct plus revision, metalinguistic explanation, metalinguistic plus revision) may lead to variations in the effectiveness of focused and unfocused written CF (Sheen, 2009; Van Beuningen et al., 2008, 2012). Thus, from there, the chapter explains focused and unfocused feedback and puts forward an argument concerning their efficacy for L2 development. Then, by drawing on Yang and Lyster's (2010) categorization of linguistic error types into rule-based and item-based errors, the chapter discusses the differential efficacy of written CF types relating to different grammatical structures. It is expected that simple rule-based forms and structures may be easier to acquire than item-based forms/structures and more complex structures.

It is acknowledged that the stages in the cognitive processing of input (during a single episode) may either be facilitated or impeded by the moderating presence of individual difference factors. It has been revealed that cognitive factors such as language analytical ability can impact the efficacy of written CF (Shintani & Ellis, 2015; Sheen, 2007). Therefore, it is necessary to investigate the moderating effect of other cognitive factors on L2 development (Bitchener, 2012, 2016; Ellis, 2010). For instance, oral CF studies have shown that working memory impacts the process of L2 development (Goo 2012, Li, 2013). Thus, there is a need to investigate if working memory, as a cognitive factor, may moderate also the efficacy of written CF. It has been argued that working memory may moderate the efficacy of written CF (Kormos, 2012). Thus, the chapter will conclude by focusing on the second aim of the study; that is, examining the extent to which working memory, as an internal individual factor, may moderate the efficacy of L2 development of the English passive voice immediately and over time.

## 2.2 An overview of L2 development

As L2 development is a construct in this study, the aim of this section is to provide an overview of what it means. In this study, the term *L2 development* (as opposed to *L2 acquisition* and *L2 learning*) is used as it is believed to be the most precise term to refer to the stage or stages in the process of L2 development from the initial written CF input stage to the automatized, unconscious output stage. Additionally, L2 development, L2 learning and L2 acquisition are often used interchangeably. Thus, the definitions and differences between them are also discussed.

Krashen (1985) distinguished between learning and acquisition. According to him, learning an L2 is a conscious process whereas acquisition of an L2 is a natural process. He argued that they lead to different competencies, namely, learned competence and acquired competence. Learned competence is achieved by paying conscious attention to the L2 and its rules while acquired competence is largely achieved through meaningful interaction in the L2. Ellis (2009) further stated that learned competence constitutes *explicit knowledge* and acquired knowledge can be labelled as *implicit knowledge*. *Implicit knowledge* can be employed unconsciously and automatically. *Explicit knowledge*, however, is the type of knowledge that learners can access only through controlled and conscious processing and they typically achieve this knowledge in an instructional/educational context.

The distinction between these two types of knowledge has resulted in considerable debate among theorists and researchers. First, Krashen (1985, 1994, 2003) argued that these two types of knowledge are formed in different parts of the brain and that explicit knowledge cannot be converted to implicit knowledge. In the SLA literature this is called the non-interface position. Truscott (2004, 2007) also supported this argument and claimed that making use of explicit knowledge has a superficial impact but that will not contribute to

L2 development over time. However several interaction theories reject this position and, by taking a strong interface position argue that explicit knowledge can be converted to implicit knowledge through meaningful, contextual practice over time (DeKeyser, 1998). The weak interface position also supports the conversion of explicit knowledge to implicit knowledge; however, it posits there are limitations on when and how the conversion can take place (N. Ellis, 2005).

These positions claim that explicit knowledge can help learners produce more accurate output by making use of their explicit knowledge during production. Additionally, a stronger claim can be made for making use of the explicit knowledge contained in written CF when writing, because learners have additional time to process and use explicit knowledge. However, learners have more limited time to draw on explicit knowledge in oral communication contexts. Accordingly, it is reasonable to believe that written CF can contribute to the development of explicit knowledge and, by practice, lead to implicit knowledge (Bitchener & Storch, 2016).

The key question is how explicit knowledge can be converted to implicit knowledge. To answer this, an explanation of the conversion is discussed using the skill acquisition theories of Anderson and McLaughlin.

### **2.3 Skill-based theories of second language acquisition**

This section presents an account of the well-known skill-based theory of McLaughlin's (1987, 1990) skill acquisition model and Anderson's (1983) adaptive control of thought model (VanPatten & Benati, 2010). These theories present the stages (from control to automatic) that learners are required to pass through to acquire any skill, including a language skill. Skill acquisition theories maintain that intentional learning (e.g., by means of written CF) can play an important role in the controlled stage and ultimately result in

automatic processing. Thus, skill acquisition theories explain how learners can improve L2 development of a target structure (e.g., English passive voice) immediately and over time.

### **2.3.1 McLaughlin's skill acquisition model**

McLaughlin proposed the skill acquisition model of language learning in 1983, and it suggests that L2 development involves shifting from controlled to automatic processing via practice. McLaughlin (1987) claimed that declarative knowledge or knowledge about the skill (which is used when the learner is doing controlled, conscious processing and practice of a specific target structure) can be converted into procedural or automatic knowledge in the L2 context. He stated that input can be processed in two ways – either controlled (drawing on explicit knowledge) or automatic (drawing on implicit knowledge) and that learning occurs as a result of a move from controlled towards automatic processing. In other words, explicit knowledge and explicit learning from instruction and CF (including written CF) can be changed to implicit knowledge, which is essential for L2 development. Other theories and researchers (e.g., Dekeyser, 1997, 2001, 2007; Hulstijn, 1995; Schmidt, 1990, 1994, 1995) have also posited the view that controlled activities, including receiving written CF, can facilitate the conversion of controlled activities (declarative knowledge) into automatic activities (automatic knowledge).

McLaughlin (1987) pointed out that learners initially resort to conscious processing of explicit L2 information. At this stage, conscious – or ‘controlled’ – processing places a high demand on the learner’s cognitive skills and working memory. Indeed, the short-term memory limits what can be consciously learned. However, eventually, a simple sentence like “I am a student” through repeated activation (i.e. practice) can be said and



written automatically, without any conscious thought or effort, making room for new structures to be produced consciously. Then, these automatic processes (i.e. implicit/procedural information) are stored as units in the long-term memory, and they can be accessed very rapidly when needed with minimal attentional control by the learner. This implies that CF can play an important role in the controlled processing component of this model. In other words, intentional learning, for instance in the form of CF and explicit instruction, can play a facilitating role at the controlled stage and eventually be converted to automatized processing.

### **2.3.2 Anderson's adaptive control theory**

Mclaughlin's Model (1983) does not distinguish between the nature of knowledge and only focuses on changes in the nature of processing. Anderson's (1983, 2000) adaptive control theory (ACT) model is based on a distinction between types of knowledge. Firstly, *declarative knowledge* is knowledge about the skill while *proceduralization of knowledge* is using declarative knowledge in skill performance. The final stage is *automatization*, which is the ability to access existing knowledge implicitly. In other words, Anderson regarded declarative knowledge as explicit knowledge and procedural knowledge as implicit knowledge and argued that declarative knowledge can be converted to procedural knowledge via practice, which results in automatization.

In order to progress from one stage to the next, learners need to practice their use of explicit knowledge (e.g., the knowledge they received in the form of written CF) in ways that make it more intuitive to use. Ultimately, providing learners with input enables them to develop declarative knowledge. For example, when students are taught the rule to add 'ed' in the past tense, they are exposed to input that leads to their declarative knowledge of this rule. However, in order to be able to use it in speaking and writing, they need to

practise the rule. Practice helps them to convert their declarative knowledge to procedural knowledge in the form of output. Then, if they continue practising, they can convert the modified output to the correct form of output in their speaking and writing. Explicit knowledge has an essential role in the process because it allows the skill to be broken into smaller stages and practiced (Dekeyser, 2007). Corrective feedback can be utilized to provide explicit knowledge that assists the learners by drawing their attention to a problematic area and helps prevent errors from becoming proceduralized.

Since ‘practice’ has a central role in the conversion of explicit knowledge to implicit knowledge, it is important to have a clear understanding about it in this context. Traditionally, practice was seen in target language learning contexts as an activity that dealt with the process of deliberately and repeatedly trying to produce a specific feature of the L2. However, DeKeyser (1998) argued that mechanical drills, which were central to the behaviorist approach, were unlikely to impact on learners’ long-term memory and result in a change from declarative to automatic processing. Thus, meaningful practice needs to be in the form of meaningful communication to facilitate progress from one stage to the next.

In order to understand the written CF input learners can receive for cognitive processing, in the following section different types of written CF, as a form of input, are discussed.

## **2.4 Types of written CF as input**

In this section in order to understand the nature of the written CF types, their definitions and the way they differ from other types of written CF are presented. L2 learners, depending on the context, can be provided with two types of input: either *positive* or *negative* input. Positive input refers to linguistically well-formed forms and sentences in the L2. In other words, it deals with what is linguistically acceptable in the L2. On the

other hand, negative input, including written CF, is about providing learners with information about the incorrectness of an L2 form or structure (i.e., what is not linguistically acceptable in L2).

In the literature, there are three main types of written CF provided as negative input: direct written CF, indirect written CF and metalinguistic written CF (Ellis, 2009). *Direct written CF* provides students with the corrected forms of their linguistic errors through crossing out, rewriting or addition near or above the linguistic error. This enables learners to edit their writing and improve their performance in future tasks. In *indirect CF*, students' attention is drawn to the position of their errors by the circling and underlining of an error, but it is left to the students to work out the correction needed. The third form of written CF, *metalinguistic CF*, uses error codes to indicate the type of error a student has made or by means of numbering errors and then providing students with a brief metalinguistic explanation for the errors at the bottom of the page or at the end of the full text. Metalinguistic CF can also be in the form of a handout to students that includes an explanation of the target structure (Shintani, Ellis, & Suzuki, 2014).

These three types of CF are provided in written form, thus they are all an explicit form of input that can help develop a learner's explicit knowledge. However, the degree of explicitness differs among them. Metalinguistic CF, as defined above, is more explicit than the other two as it provides more linguistic data on errors.

In order to understand how the input is converted from explicit declarative to automatic knowledge, one needs to understand the cognitive processing required. Gass' (1997) model, in the following section, explains how the conversion takes place.

## **2.5 Cognitive processing of input for explicit declarative knowledge**

This section draws on Gass's (1997) cognitive framework in order to describe the stages that apply to a single episode of input processing.

Gass's (1997) cognitive framework includes five stages in the development process, namely apperceived (noticed) input, comprehended input, intake, integration and output. Each of these stages needs to be completed before the next stage can be activated.

*Apperceived input* means that learners have to notice the features of the input in relation to their existing knowledge. *Comprehended input* can be considered as one stage beyond noticing. The input may be analyzed and has the potential to be assimilated prior to it becoming intake (stage three). "Psycholinguistic processing occurs at this stage where new information may be matched against existing stored knowledge" (Gass, 1997, p.23). *Intake* is a process of mental activity which moderates between comprehended input and grammar, through which the process of assimilating linguistic material occurs. The next stage, *integration*, deals with "storage of new information for later use, hypothesis formulation, and confirmation or reformulation of existing hypotheses" (Gass, 1997, p. 25). In this stage, the amount of attention a learner pays to CF may have an influence on the degree to which he/she retains the features in his/her long-term memory. *Output*, the final stage, involves deeper processing of linguistic forms; in other words, output leads to learning.

Bitchener and Storch (2016) pointed out that even though the five stages were designed to discuss the cognitive processing of oral corrective feedback, the same stages can be broadly applied to the cognitive processing of written CF. The first stage explains that learners are required to notice or apperceive the explicit input. In other words, learners are required to notice the gap between the explicit corrective feedback input and their erroneous production in order to modify it accurately. Similarly, learners in the written context are required to notice the explicit input in the same way as in the oral context. The second stage of information processing explains that input needs to be 'comprehended' by learners as described by Gass. This stage is also required in the written context for explicit input to go through central processing. Written CF can also

facilitate intake. At this stage, learners match the information provided by the input/CF with their existing explicit knowledge. In the writing context, as opposed to the speaking context, learners may have a better opportunity to match the input/written CF with their existing knowledge because of (1) the writing process enabling them to refer back as often as desired to their written text and the feedback provided on the text, and (2) having sufficient time to analyze and re-analyze their text. Written CF can facilitate *integration*. In the context of writing, hypothesis-testing is facilitated by time, in that learners have sufficient time to make a cognitive comparison between the written CF and existing knowledge retrieved from their long-term memory. Output, the final stage, is important because written CF is provided in the hope that it will allow learners to notice their errors and then modify them.

The extent to which the learners progress through the above various stages of cognitive processing may be dependent upon a number factors and variables, one of which is the types of written CF provided. This is theoretically discussed in the following section.

## **2.6 Theoretical accounts of the efficacy of different types of written CF**

This section presents a theoretical discussion on the efficacy of written CF types. In the written CF literature, there have been theoretical arguments on the efficacy of different types of written CF on L2 development (Bitchener & Storch, 2016). Those who favour indirect written CF suggest that it is most useful because it engages learners in problem solving and guided learning (Lalande, 1982). Those supporting direct feedback suggest that it reduces learners' confusion and offers more explicit feedback on the hypotheses and it is immediate as opposed to indirect feedback. Additionally, direct CF is more useful to resolve learners' errors on more complex forms and structures. However, if learners have partially acquired a particular form and structure in their L2, there is a need for a

fuller explanation regarding the error in order to hypothesize and produce the correct form and structure over time. Because direct CF only provides the correct form of an erroneous form/structure, it may not be useful for some learners (Bitchener & Ferris, 2012).

Accounts of the value of metalinguistic CF have recently drawn researchers' attention. Those supporting metalinguistic CF (Guo, 2015) suggest that it may be useful for any proficiency level because it can provide an explanation of new knowledge about the L2 or it can draw learners' attention to a partially acquired form/structure. However, indirect CF may be more useful for the on-going L2 development of advanced learners because it helps them to identify an incorrect production and, by drawing on their long-term memory, it helps them to hypothesize the correct version.

Additionally, written CF types may play different roles in terms of getting the learners attention and helping them to comprehend written CF. Bitchener (2012) argued that different types of written CF differ in their degree of explicitness and as a results may draw learners' attention to different levels. For instance, at the stage of knowledge modification, learners may notice and then comprehend more explicit types of written CF resulting in improved accuracy; however, if less explicit types of written CF are provided, learners may not be able to form a correct hypothesis and, consequently, the result may not be immediate improved accuracy. However, although less explicit types of written CF may not lead learners to form a correct hypothesis, they may enable learners to store the information and wait for more input. At the stage of knowledge consolidation, learners may benefit from less explicit types of written CF because they are able to retrieve the relevant explicit knowledge from their long-term memory. Consequently, learners may undertake deeper controlled processing because the feedback is not clearly explained. However, explicit types of written CF (e.g., direct correction and metalinguistic explanation) may not lead to deeper controlled processing because the information in the explicit feedback is clearly identified and explained.

Additionally, Bitchener and Storch (2016) argued that written CF types may play different roles in terms of helping learners to comprehend written CF. For instance, metalinguistic explanation feedback explicitly explains learners' errors and may help learners to more clearly and more fully comprehend their errors than less explicit CF such as indirect CF that identifies learners' errors by, for example, underlining or circling them. This is especially the case if the learners have partially stored the information on the target structure in their long-term memory. Thus, learners' proficiency in the target language as well as the extent to which they have stored knowledge in their long-term memory may impact whether the CF is comprehended. For learners with a low level of proficiency, more explicit types of feedback, such as metalinguistic feedback, may help them to comprehend input in the form of written CF, while for advanced learners less explicit feedback, such as indirect CF, is sufficient because they have more information stored and possibly have had more retrieval experience (Bitchener & Storch, 2016).

In summary, the degree of explicitness of written CF may impact the efficacy of written CF on L2 development. Besides the degree of explicitness of written CF, the question has been raised whether the efficacy of written CF may change due to text revision following feedback. The next section theoretically discusses the extent to which revision following feedback may promote the effectiveness of written CF.

## **2.7 Theoretical accounts of the efficacy of revision following feedback on L2 development**

This section provides a theoretical overview on the efficacy of revision in L2 development. A reason for asking learners to revise their texts is firstly to determine whether they have learnt from the feedback provided. Secondly, it is important to determine whether learners can apply what they have learnt from the feedback when revising their texts. If learners are able to accurately revise their texts, it reveals that they

have likely understood the feedback (Bitchener & Storch, 2016). However, whether or not revising a text leads to L2 development is controversial. For example, it has been argued that regardless of whether there is a single opportunity to revise following written CF (Frear, 2012; Van Beuningen et al., 2008) or multiple opportunities (Chandler, 2003; Hartshorn et al., 2010), revision following feedback leads to greater accuracy in new writing. This is because revision following written CF leads to ‘pushed output’ (Shintani et al., 2014). Swain (1985, 1995) argued that when learners are *pushed* to produce a language, they likely “notice a gap between what they want to say and what they can say, leading them to recognize what they do not know, or know only partially” (Swain, 1995, pp. 125-126). Swain takes the position that pushed output helps learners to notice the grammatical forms that probably otherwise would go unattended. Thus, when required to undertake revision, learners need to give explicit attention to the error and its correction, which may promote the storage of the target structures in the memory.

Although revision may facilitate L2 development, Truscott (1996) argued that a learner’s ability to produce an accurate revision of a text does not mean that they are able to produce these target forms in a new piece of writing. Truscott’s claim may be true because there is a distinction between revision and writing a new text. In revision, learners may only be required to focus on and revise the errors in the same linguistic context while in writing a new text, learners may be required to undertake deep processing and focus both on form and meaning in a new linguistic environment. This is especially so when learners are provided with a type of written CF (e.g., direct CF) where they only need to copy the correct form in the text revision. Thus, the effectiveness of written CF – either with revision or without revision – needs to be investigated through the writing of a new text. Additionally, it might be important to investigate whether the effectiveness of written CF – either with revision or without revision – may vary if it targets a single linguistic



form/structure (focused) rather than three or more (unfocused) at a time. Therefore, theoretical accounts of focused and unfocused written CF are explained in the following section.

## **2.8 Theoretical explanations of the efficacy of focused and unfocused written CF on L2 development**

The aim of this section is to present a theoretical account of the effectiveness of focused and unfocused written CF on L2 development. It has been argued that focused and unfocused written CF may have different effects on L2 development (Sheen, 2009; Van Beuningen et al., 2008, 2012); therefore, the extent to which focused written CF may promote L2 development, as opposed to unfocused written CF, is theoretically discussed in this section.

Written CF can be either unfocused or focused. Unfocused CF, or comprehensive feedback, is what many language teachers normally use in writing classes; that is, they generally provide feedback on all writing errors or at least on a range of errors. In contrast, focused CF can be either ‘highly focused’, that is, feedback is provided on only one error, or ‘less focused’, that is, feedback is provided on a limited number of errors (Ellis et al., 2008).

There have been a number of arguments over the years in support of the efficacy of both unfocused and focused written CF on L2 development (Bitchener 2016, Bitchener & Ferris, 2012). Those who tend to support unfocused feedback (e.g., composition teachers and writing teachers at tertiary level) consider focused CF as a luxury option and argue that focused feedback is not helpful when learners are immediately in need of feedback on a wide range of errors. Those supporting the need for focused feedback (e.g., L2 classroom teachers of the target language) maintain that with focused feedback, learners

are more likely to attend to error corrections provided on a single error (or a limited range of errors) and are more likely to gain a better understanding of the nature of the error and the correction required. Thus, since attention and understanding play important roles in L2 development (e.g., Schmidt, 1994; Ellis, 2005), focused CF produces positive results because focused CF focuses learners' attention on one or two features rather than on too many, which may lead them to be unable to give their attention and understanding to the task (Ellis et al., 2008). Additionally, by drawing on the limited processing capacity model of L2 development (Robinson, 1995, 2003; Schmidt, 2001; Van Patten, 1996, 2004), focused CF is likely to be more effective for students at a low level of proficiency because they have a limited store to draw upon when they compare the feedback with the errors. Students at a high level of proficiency are more likely to benefit from unfocused CF than those at a lower proficiency level as they have a larger memory store and possibly have had more retrieval experience, allowing them to attend to a wide range of input/CF on a single occasion (Bitchener, 2012).

In addition to a learner's amount of stored knowledge and their attentional capacity to enable them to draw upon it, the effectiveness of focused and unfocused written CF may change when different linguistic error types are treated. Theoretical accounts of this issue are discussed in the following section.

## **2.9 Theoretical explanations of the differential efficacy of written CF types for treating different types of linguistic error**

This section theoretically discusses whether written CF is effective in targeting certain linguistic error domains/categories. Yang and Lyster (2010) categorized linguistic error types into rule-based and item-based errors. Rule-based errors are those that occur in a rule-governed way and can be corrected, for example, by referring to grammar books,

while item-based errors are those for which there is no set of rules students can refer to or the rules vary according to use in different linguistic environment. In other words, these errors are less rule-governed and grammatical rules are less likely to be helpful in resolving them. In these cases, there is a need for specific knowledge of the targeted language

Although it is likely that rule-based errors are more treatable than item-based errors, it is important to consider the complexity of the rule-based forms/structures. For instance, the rules for employing the simple past tense are relatively straightforward; however, the rules for some structures such as the passive voice could be difficult to acquire as they involve complex structures.

Structures such as the passive voice are categorized as both rule-based and item-based structures. In the passive voice the verb “to be” is categorized as rule-based while the “past participle” is item-based. For instance, in the sentences, *The house **is cleaned** by Tom. Pods **are grown** in South America. Pods **are spread** in the sunshine* “to be” is rule based as there is a rule that identifies when learners need to use “am”, “is” or “are”. However, the past participle is item-based as it takes three different forms: regular verbs that end in “ed” (e.g., “cleaned”); irregular verbs in which the past participle changes form (e.g., “grown”); and irregular verbs where the past participle retains the same form (e.g., “spread”). Thus, structures that are both semantically and syntactically complex, such as the passive voice, require learners to use more attentional capacity when employing them. For complex structures, less explicit feedback is unlikely to be beneficial for low proficiency learners; rather, they require written CF such as metalinguistic CF that explicitly explains and illustrates the complex structures. Thus, it is expected that there is some correlation between error types and written CF types. In other words, it may be that for some learners more explicit feedback may need to be provided on complex structure errors such as the English passive voice and hypothetical

conditional, and it is likely that for some learners less explicit feedback may be helpful for simple rule-based errors such as the regular past tense.

Additionally, it is important to consider that learners' proficiency level and existing knowledge play a paramount role during cognitive processing. In other words, less explicit feedback may be helpful for learners with a high proficiency level and related existing knowledge, while more explicit feedback is required for learners with a low proficiency level and without related existing knowledge.

In summary, the above discussion showed that the effectiveness of written CF may vary by targeting different types of linguistic forms/structures. In addition to linguistic factors, individual differences may also have a moderating effect on written CF (Ellis, 2008). Thus, individual factors may impact cognitive processing. As Kormos (2012) explained "individual differences may be hypothesized to exert influence on how students process feedback, the extent to which they notice gaps in their knowledge, the aspects of language they pay attention to, and, consequently, how they exploit the learning opportunities provided by writing" ( Kormos, 2012:400). Working memory, as a cognitive individual factor, may play an important role in consolidating explicit knowledge in both the initial single episode and in subsequent episodes of cognitive processing (Bitchener & Storch, 2016). Therefore, this study has investigated the moderating effect of working memory on learners' cognitive processing of new information in the form of written CF.

## **2.10 The potentially moderating influence of working memory on cognitive processing**

Individual factors have been recognized as important in the process of L2 development (Bitchener & Ferris, 2012; Bichener & Storch, 2016). Individual differences can be

classified into the individual learner-internal motivational/affective factors and learner-internal cognitive factors (e.g. working memory) (Dörnyei, 2005, Ellis, 2010).

It has been argued that the stages of the cognitive processing of input (single episode) may either be facilitated or impeded by the impact of individual cognitive factors such as working memory (Bitchener, 2012, 2016; Ellis, 2010). Thus, the aim of this section is to theoretically discuss the extent to which working memory, as an internal individual factor, may moderate the efficacy of L2 development of the English passive voice immediately and over time.

Oral CF studies have shown that working memory moderates the process of L2 development (Li, 2013). Thus, there is a need to investigate whether working memory, as a cognitive factor, moderates the efficacy of written CF (Bichener & Storch, 2016). The extent to which working memory may act as a moderator is theoretically discussed in the following.

The concept of working memory has become a topic of interest and controversy over the past decades as a consequence of Baddeley and Hitch's (1974) introduction of their working memory model. Working memory was defined by Baddeley (2003) as "the temporary storage and manipulation of information that is assumed to be necessary for a wide range of complex cognitive activities" (p. 189). This is in contrast to the earlier theories of memory that focused only on its storage function. Baddeley and Hitch's (1974) model presents a dynamic approach to memory; namely, working memory integrates storage with the processing and manipulation of information. Additionally, Engle (2002), a cognitive psychologist, argued that working memory includes the ability to control attention in order to keep information in an active, promptly retrievable state; it is not just about individual differences and the working memory's storage capacity. Thus, working memory, in addition to storage, can play a much deeper role in cognitive activities such as, attention, noticing, hypothesizing, restructuring and practising.

Baddeley and Hitch's (1974) proposed model (see Figure 2.2) of working memory consists of a *central executive* and two subcomponents – the *phonological loop* and the *visuo-spatial sketchpad*. The *central executive* is thought to be responsible for the control and regulation of working memory.

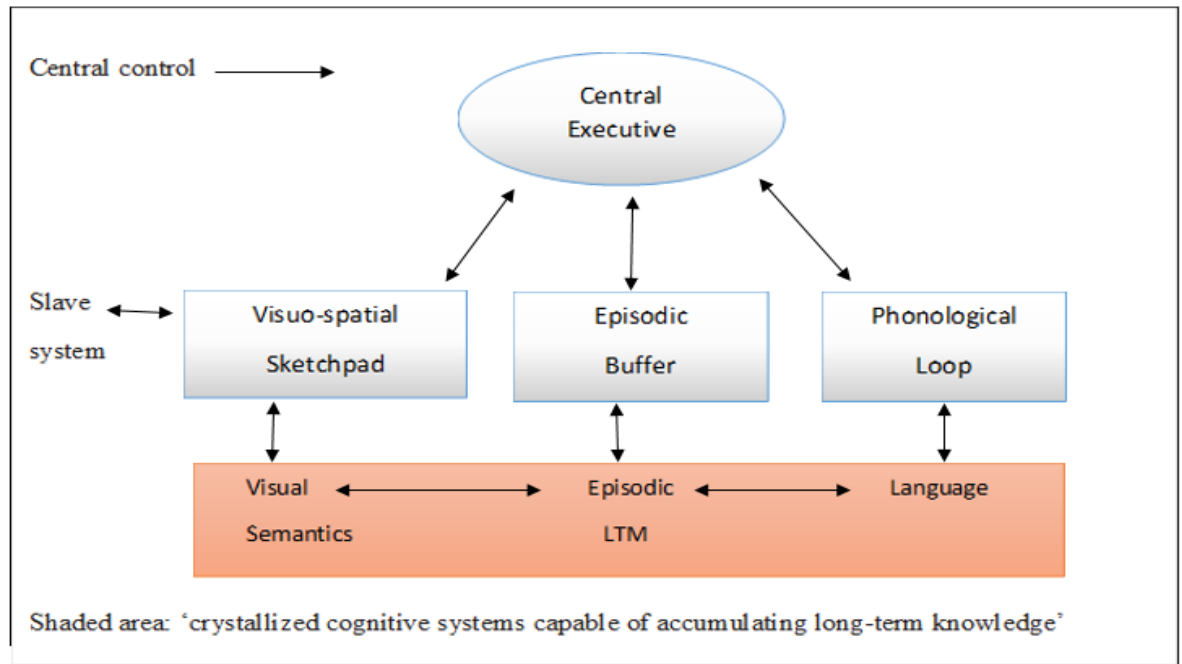


Figure 2.1: Development of the working memory model (Baddeley, 2000)

The *phonological loop* stores acoustic and verbal information. The *visuospatial sketchpad* is responsible for storing and processing spatial and visual information. Later, Baddeley (2000) added a fourth subcomponent to the working model called the *episodic buffer*. This subsystem integrates spatial, visual and verbal information from a) the phonological loop, b) the visuospatial sketchpad and c) long-term memory into single multimodal units or episodes (e.g., a story or a scene in a film). However, among the aforementioned components of working memory, only the phonological short-term memory and the central executive have been revealed to be mostly related to first and second language learning and processing (Gathercole & Baddely, 1993; Juffs & Harrington, 2011; Linck, Osthus, Koeth & Bunting, 2014; Wen, 2012, 2014; Williams, 2012).

Based on the capacity-limited model (e.g., Skehan, 1998), working memory has limited capacity and can process a limited amount of information at one time; therefore, learners with greater working memory capacity can better attend to and process input/CF (Ortega, 2009). Additionally, the model explains that learners with a lower level of proficiency might find it difficult to simultaneously attend to more than one aspect of language (e.g., meaning and form). Because learners with a lower level of proficiency need to process new knowledge consciously, they need greater attention and effort in processing the new information in their working memory. Although written CF allows more time to process information than oral CF, a learner's working memory still needs to coordinate the learner's attention to cognitive processing. On each new occasion during the consolidation stage, after written CF has been given and accurate output has been produced, learners draw upon their working memory to retrieve the new information from their long-term memory and produce it. Over time when new knowledge is proceduralized, the learner may put less effort into the working memory to process new information. Additionally, it is expected that learners with different working memory spans process and produce new information differently (Bitchener & Storch, 2016).

Kormos (2012) pointed out that working memory might moderate how learners learn from different types of written feedback. She argued that in contrast to the oral context, learning opportunities through feedback in the writing context are less constrained by time pressure; however, she stated that because writing learners are dependent on their working memory capacity, they may respond differently to feedback. Thus, this study will be the first to determine the extent to which working memory may moderate learners' use of different types of written CF (direct CF and metalinguistic explanation).

## **2.11 Summary**

This chapter firstly provided an overview of SLA theories and hypotheses regarding the cognitive processes of L2 development and discussed the extent to which written CF may

play a role in these processes. Then, the chapter examined the extent to which working memory, as an internal individual factor, may moderate the processes of L2 development. Thus, the chapter presented theoretical arguments in support of the potential role of written CF in facilitating learners' L2 development. In doing so, some important counter-arguments have been offered.

The chapter argued that from a cognitive information processing perspective, written CF may facilitate L2 development. In other words, written CF may facilitate knowledge modification when employed as a form of input or knowledge consolidation and utilized as a noticing-trigger. Additionally, it was discussed that because learners' existing knowledge, processing capacity and attention play a key role in cognitive information processing, the degree of explicitness of written CF, the complexity and number of linguistic forms/structures, and individual differences in working memory may moderate the effectiveness of written CF on L2 development.

The extent to which the theories proposed in this chapter are valid accounts of the contribution of written CF to L2 development is an issue that only empirical research can determine, as discussed in the following chapter.



## **CHAPTER 3**

### **A REVIEW OF WRITTEN CORRECTIVE FEEDBACK STUDIES**

#### **3.1 Introduction**

The theoretical notions and hypotheses presented in Chapter 2 describe how information arising from written CF can be processed, produced as modified output, retained and be used with accuracy over time. The discussion in the previous chapter revealed that the efficacy of written CF has been a topic of considerable debate among theorists and researchers (Krashen, 1985, 2003; Truscott, 1996, 2007). This chapter, by drawing on the empirical research, attempts to critically evaluate four groups of cognitively informed research that have examined the efficacy of written CF on L2 development. These groups are: (1) studies that investigated the effectiveness of written CF; (2) studies that examined whether certain types of written CF were more effective on L2 development than others; (3) studies that investigated whether written CF has a different level of efficacy when different types of linguistic errors are targeted; and (4) studies that explored the extent to which individual differences in working memory may moderate the effectiveness of different types of CF.

The first group of research has investigated the efficacy of written CF on learners' improved accuracy. These studies have been conducted on written CF in order to address Truscott's (1996, 1998, 2004, 2007, 2010) claim that written CF is ineffective in promoting L2 acquisition, and even can be harmful to it. These studies have examined whether or not written CF, as an explicit form of input, impacts L2 development immediately and over time. Additionally, a number of these studies have investigated the effect of text revision on accuracy improvement when writing new texts over time.

Providing learners with an opportunity to revise a text may play an important role in the development process (Chandler, 2003). This is because it *pushes* learners to attend to the feedback they have been provided with (Shintani et al., 2014), to process it across cognitive processing stages (Gass, 1997) and to make a hypothesis regarding an accurate modification.

The second group of research has investigated whether certain types of written CF are more effective at improving accuracy than others. These studies are both theoretically and pedagogically motivated. Theoretically, they have investigated the role that the explicitness of different types of written CF can play in L2 development. Pedagogically, teachers would like to know whether certain written CF practices (such as providing direct written CF and metalinguistic explanation) lead to higher levels of improved accuracy than others.

The third group has examined whether written CF has a different level of efficacy when different types of linguistic errors are treated. It has been argued (Bitchener, 2016) that written CF targeting simple, rule-based forms/structures may be more effective than written CF targeting more complex, ruled-based forms/structures (because learners have to process more than one linguistic element in complex forms/structures) and item-based forms/structures that are not rule-governed. Most of these studies have targeted *forms*, and only two studies have targeted *structures* (present perfect tense, hypothetical conditional) (Rummel, 2014 & Shintani et al., 2014). Thus, because of the limited number of studies, there is a need to further investigate the effectiveness of written CF when treating different structures, for example the English passive voice, as targeted in the present study.

The last group of research has explored the extent to which individual differences in working memory may moderate the efficacy of different types of oral CF. However, we

do not know whether working memory moderates the cognitive processing of corrective feedback in the written context (i.e. written CF) in the same way that it does in the oral context (i.e., oral CF). This section, therefore, focuses only on those studies that have investigated the moderating effect of working memory on L2 development in oral CF. Kormos (2012) pointed out that working memory may moderate the effectiveness that different types of written CF may have on improved accuracy or modified output. She argued that in contrast to the oral context, learning opportunities through feedback in the written context are less constrained by time pressure; however, because L2 writers are dependent on their working memory capacity, they may respond differently to some types of feedback. For example, learners at a lower proficiency level may respond better to metalinguistic explanation, while advanced learners may respond better to direct CF as it acts as a reminder of what they already know. Thus, this study investigates the moderating effect of working memory on written CF.

These four groups of studies are discussed in the following sections (3.1-3.4).

### **3.2 Research on whether or not written CF can facilitate L2 development**

A growing number of studies have been conducted on written CF in order to address Truscott's (1996) claim that written CF is ineffective and even harmful in promoting L2 acquisition. In these studies (e.g., Bitchener, 2008; Shintani et al., 2014; Van Beuningen et al., 2012) three types of output have been investigated as an indication that the L2 development process has begun: (1) output that is produced when a learner revises his/her original output; (2) output that is produced when a learner writes a new text; (3) output that is produced when a learner writes a new text over time. These studies have measured output by employing a pre-test/written CF treatment/post-test design and compared a learner's accuracy performance prior to providing written CF on linguistic errors in the

pre-test with accuracy performance immediately after receiving feedback (immediate post-test) or at various periods of time after receiving feedback (delayed post-test). The immediate post-test writing task can come in the form of pre-test revision or a new text or both. If learners in the experimental groups, that is, the groups that receive feedback, reveal a significant increase in accuracy between the pre-test and immediate post-test scores and this improvement is significantly higher than the group of learners who do not receive feedback (control group), then it is understood that written CF has facilitated the learning process. If a learner maintains accuracy improvement in delayed post-tests, it is concluded that improvement is being consolidated. In order to evaluate the research findings on the efficacy of written CF on L2 development, each of the studies are reviewed. To do so, in the following section a review of revision studies is presented followed by an investigation of the effectiveness of written CF on new texts immediately and over time.

### **3.2.1 The efficacy of written CF on a revised version of a text**

Providing learners with opportunities to revise their texts may play an important role in the development process because it invites them to notice the feedback they have been provided with (Shintani et al, 2014). A number of the initial studies that investigated the efficacy of written CF asked learners to revise their text after providing them with written CF. These studies attempted to find out if learners had learnt anything from feedback and also if they could accurately use their learning when revising their texts. The early studies on the effectiveness of revision are summarized in Table 3.1.

*Table 3.1 Studies of written CF on the effectiveness of text revision*

<b>Studies</b>	<b>Participants</b>	<b>Treatment</b>	<b>Error types</b>	<b>Effectiveness</b>
<b>Fathman and Whalley (1990)</b>	72 ESL learners at US colleges	1. Grammar feedback (underlining) 2. Content feedback 3. Grammar & content feedback 4. Control	Comprehensive	Yes
<b>Ashwell (2000)</b>	50 EFL learners at a university college	1. Content then form 2. Form then content 3. Form and content 4. Control	Lexical Grammatical Mechanical	Yes
<b>Ferris and Roberts (2001)</b>	72 ESL learners at US colleges	1. Error code 2. Underline 3. Control	Noun, verb, article errors, wrong word, sentence structure	Yes

In their studies, Fathman and Whalley (1990) and Ferris and Roberts (2001) reported that ESL learners who received written CF employed more accurate linguistic forms in their revision text than the learners who did not receive feedback. Ashwell (2000) also reported similar findings with EFL learners. Although the above findings showed positive evidence for improved accuracy in text revision, Truscott (1996) argued that learners' accuracy when revising their texts does not constitute evidence of L2 learning and that they need to show their learning in new pieces of writing. Thus, several studies (see Table 3.2) have investigated the extent to which revision results in increased accuracy in new pieces of writing.

*Table 3.2 Studies of written CF on the effectiveness of text revision and new text writing*

<b>Studies</b>	<b>Revision</b>	<b>New text writing</b>
<b>Truscott and Hsu (2008)</b>	Improved accuracy	No improved accuracy
<b>Van Beuningen et al. (2008)</b>	Improved accuracy	Improved accuracy
<b>Van Beuningen et al. (2012)</b>	Improved accuracy	Improved accuracy

Truscott and Hsu (2008) reported that the increase in accuracy revealed by their experimental group when revising their texts was not shown in their writing of a new text. Thus, they concluded that “the successful error reduction during revision is not a predictor

... of learning” (p. 299). However, Bruton (2009) challenged the validity of the claim as he found that learners made only a few errors in their pre-test writing, and that they, therefore, had little room for improvement. Furthermore, the findings by Van Beuningen et al. in both their studies (2008, 2012) contradicted the findings of Truscott and Hsu (2008). In their pilot study of 62 learners, Van Beuningen et al. (2008) found that both experimental groups (error code and direct error correction) increased their accuracy in the text revision; however, the first experimental group (i.e., direct error correction) was able to write a new piece of writing with improved accuracy a week later. In their main study (2012), the authors reported that after four weeks all 268 learners retained the same level of accuracy in the delayed post-test as was recorded in the text revision.

The main difference between Truscott and Hsu’s (2008) and Van Beuningen et al.’s (2008, 2012) studies is the degree of explicitness of written CF, and this may have led to the contradictory findings. In contrast to Truscott and Hsu’s (2008) study in which one type of written CF with a low level of explicitness (underlining) was employed, in Van Beuningen et al.’s (2008, 2012) study the researchers used two more explicit written CF types (direct correction and error codes). This may indicate that the degree of explicitness of written CF can impact on the efficacy of written CF.

The conflicting findings in the above studies reveal a need for more studies to show if revising a text leads to improved accuracy in a new writing text. Thus, this is one of the issues the present study investigates. It is important and logical to investigate the effectiveness of written CF in a new piece of writing because writing a new text is very different to text revision. In writing a new text, learners are likely to perform deep processing and focus on both meaning and form in a new context; however, in text revision, learners may only focus on form and revise errors in the same context. This is especially so when learners are provided with a type of written CF (e.g., direct CF) where they only need to copy the correct form in the text revision. Thus, the effectiveness of

written CF – either with revision or without revision – needs to be investigated in new writing.

### **3.2.2 The efficacy of written CF on new pieces of writing**

The studies that have examined the effectiveness of written CF on new texts immediately and over time can be placed into two categories. First, the studies that cover a wide range of different types of linguistic errors (including punctuation and spelling errors, and lexical and structural errors) are called *unfocused* written CF studies. All the revision studies discussed above are unfocused studies. Second, the studies that focus on only a limited number of language domains (less focused) or a few linguistic error types (highly focused) are called *focused* written CF studies. In order to justify the first research question, in the following three sections unfocused studies will firstly be reviewed as these early studies were all unfocused. The design/execution shortcomings of a number of them will also be discussed. Secondly, the focused studies that have tended to adopt a more considered methodology will be reviewed. Those studies that have compared focused and unfocused written CF will also be discussed.

#### **3.2.2.1 Studies on unfocused new texts**

There are only a few research studies in the literature that examined the effectiveness of unfocused written CF for L2 development. These studies can be categorized into early and recent studies. The four early new text studies that investigated the efficacy of written CF on new texts were unfocused studies (Semke, 1984; Robb *et al.*, 1986; Kepner, 1991; Sheppard, 1992). These four studies claimed that written CF does not lead to accuracy in the new text. These studies are summarized in Table 3.3 below.

*Table 3.3 Early unfocused studies on the effectiveness of written CF on new texts*

<b>Studies</b>	<b>Participants</b>	<b>Treatments</b>	<b>Methodological flaws</b>	<b>Effectiveness</b>
<b>Semke (1984)</b>	141 German FL learners	1.Direct error correction 2.Direct error correction plus content comments 3. Error codes	Not a real control group (content comments)	No
<b>Robb et al. (1986)</b>	134 Japanese EFL freshman	1.Direct error correction 2.Error codes 3.Highlighting 4.Marginal error total	Not a real control group (marginal error total)	No
<b>Kepner (1991)</b>	60 intermediate Spanish FL learners	1.Direct error correction 2.Content comments	No pre-test Not a real control group (content comments)	No
<b>Sheppard (1992)</b>	26 upper-intermediate EFL learners	1.Direct error correction plus conference 2.Error codes	Not a real control group (content comments plus conferences)	No

Due to methodological shortcomings, there are a number of controversies related to the validity of the conclusions drawn from the findings of these four studies. In her study of 141 freshman German EFL learners in an American university, Semke (1984) divided the participants into three experimental groups and a control group. The experimental groups were provided with direct written CF, error code and correction plus positive comments respectively. The control group was given written comments. The study found no significant improvement in accuracy in new texts between the treatment groups and the control group, and Semke concluded that written CF was not effective in improving accuracy in new texts. However, there are some concerns with regard to the control group; that is, the questions asked of them likely raised the students' attention to certain linguistic errors. Additionally, the study had several shortcomings in terms of measurements; that is, the treatment groups and the control group were measured differently. The control group was graded on the basis of the number of words that were written but the treatment groups were marked according to the ratio of the number of errors to the number of words



written. Different methods of measurement can lead to conflicting findings. Thus, these issues mean that caution is necessary with regard to the validity of the findings – in this case, the ineffectiveness of written CF on new texts.

It is also debatable whether Robb et al.'s (1986) study of 134 EFL learners in Japan had a real control group. The four groups were either provided with direct correction, error coded feedback, highlighted feedback or the number of errors per line was pointed out. The study had no traditional control group (the control group refers to the group that is not provided with treatment and is then employed as a scale to measure how the other groups perform). However, Truscott (2007) argued that the fourth group, which was provided with the total number of errors, could be considered a control group 'because the information provided to (this) group was so limited that it could not have been helpful' (p.261). The result of the study revealed that all four groups improved and there was no significant difference among the four groups. In other words, there was no difference between the groups which received written CF and the group which did not receive it. Thus, Truscott (2007) posited the view that the findings of this study showed the ineffectiveness of written CF.

The study by Kepner (1991) examined the efficacy of written CF by comparing the improved accuracy of intermediate Spanish FL learners when dealing with new texts. The first problem with the study was that it did not have a control group and it only involved two treatment groups – that is, one group that was given direct error correction and the other that was provided with content comments on their texts. The second shortcoming, as Ferris (2003) argued, was that the study did not have a pre-test writing task in order to identify the pre-treatment level of accuracy of the learners. In other words, it was not clear (1) whether the participants had the same initial level of accuracy or (2) how the participants' improvements were measured. Thirdly, the study gave no details on the controls and conditions placed on participants who did their writing tasks out-of-class.

For these reasons, the study did not provide solid evidence of the ineffectiveness of written CF on new texts.

Sheppard's (1992) study of 26 upper-intermediate ESL learners also had no real control group. The study investigated two groups, one of which received error code treatment and the other which received comments relating to content. Because the comments on the content were not about linguistic accuracy, the second group can be considered as the control group. However, the findings of a study are only valid if its design is unambiguous and fully defined. The second problem with Sheppard's study is that the design was based on one-on-one conferences between teacher and students in the comments on content group. It is unclear whether the discussion between this group and the teacher avoided any discussion of problems in understanding the meaning that may have arisen from linguistic errors. Additionally, Ferris (2004) commented that there was a lack of inter-rater reliability in the coding of data. Again, based on the design flaws, it can be concluded that the validity of the results is questionable.

Thus, the above early *unfocused* studies which investigated the effectiveness of written CF on new texts reported that written CF was ineffective. However, these studies had a number of methodological flaws. Recently, three studies have attempted to avoid the methodological shortcomings and have investigated the effectiveness of unfocused CF on new texts (see Table 3.4).

*Table 3.4 Recent studies on unfocused written CF*

<b>Studies</b>	<b>Participants</b>	<b>Error categories</b>	<b>Effectiveness</b>
<b>Truscott and Hsu (2008)</b>	47 high intermediate EFL learners	All grammar, punctuation, spelling and errors	No
<b>Van Beuningen et al. (2008)</b>	62 secondary school Dutch EFL learners	All grammatical forms/structure, incomplete sentences, word omission or inclusion, lexical choice, punctuation, spelling, capitalization	Yes

<b>Van Beuningen et al. (2012)</b>	268 secondary school Dutch EFL learners	All grammatical forms/structure, incomplete sentences, word omission or inclusion, lexical choice, punctuation, spelling, capitalization	Yes
------------------------------------	---	--	-----

Truscott and Hsu (2008) examined the efficacy of unfocused CF on L2 development by investigating the improved accuracy of 47 high-intermediate EFL learners at a university in Taiwan. The study had one experimental and one control group. The experimental and control groups were required to write a guided narrative story based on eight pictures over 30 minutes. One week later, the experimental group was given their texts with all the grammatical errors underlined. However, the control group was not provided with marked errors. Both groups were then asked to revise their written texts. A week later, both of the groups were required to write another guided narrative story based on a new series of eight pictures. The findings revealed that the experimental group outperformed the control group in revision; however, neither group showed a significant difference in improved accuracy in new texts. Even though the study included a real control group and avoided methodological flaws, one study is an insufficient basis upon which to draw any conclusions regarding the ineffectiveness of written CF on L2 development. Furthermore, the researchers applied only one type of written CF, the indirect underlining of errors, which has a very low level of explicitness. The results may have been different with more explicit types of written CF.

Van Beuningen et al. (2008) conducted research on 62 Dutch EFL learners in a secondary school. The study included a wide range of error categories: spelling, pronunciation, word choice, word forms, capitalization, omission or addition of a word in incomplete sentences, and punctuation. In contrast to Truscott and Hsu's (2008) study, in which one type of written CF with a low level of explicitness (underlining) was employed, in this study the researchers used two more explicit written CF types (error codes and direct

correction) and compared them with two control groups (writing practice and self-correction without written CF). Both the direct correction group and the error code group showed a significant improved accuracy in revisions, but only the direct correction group revealed a significant improvement in accuracy in new texts. On the other hand, the two control groups did not show improved accuracy in either revisions or new texts. Thus, the main difference between Truscott and Hsu's (2008) and Van Beuningen et al.'s (2008) studies is the degree of explicitness of the written CF, which may have resulted in the contradictory findings. This may suggest that the degree of explicitness of written CF can impact the effectiveness of written CF.

Van Beuningen et al. (2012) conducted a further study with 268 participants and examined the efficacy of written CF on two different language domains (grammatical and non-grammatical errors). As each domain included a wide range of linguistic types, the study is still categorized as an *unfocused* written CF study. This study adopted a similar study design to Van Beuningen et al.'s (2008) study by utilizing four groups; however, it added another delayed post-test after four weeks to investigate the long-term efficacy of written CF. The results showed that when non-grammatical errors were considered alone, the error code group revealed more improvement than the direct correction group. However, when grammatical errors were considered alone, the direct correction group revealed significant improvement in accuracy in new texts after four weeks. Thus, the study revealed both the efficacy of direct correction in new texts and its retention over time (four weeks).

To summarize, the early *unfocused* studies (Semke, 1984; Robert et al., 1986; Kepner, 1991; Sheppard, 1992) investigated the efficacy of written CF on new texts. However, due to the methodological flaws and contradictory results, no solid conclusions could be made. On the other hand, more recent unfocused studies (See Table 3.4) have avoided the methodological flaws and both negative and positive findings have been reported. From

the findings, it would seem that the degree of the explicitness of written CF may have impacted the results. For instance, in Truscott and Hsu's (2008) study, underlining as an implicit type of written CF may not have contributed to L2 development; however, in Van Beuningen et al.'s (2008, 2012) studies, L2 development may have occurred through direct correction as it is a more explicit type of written CF.

There is another factor that may impact on the effectiveness of written CF. In the above studies, the feedback was *unfocused*; that is, feedback was given on a wide range of linguistic error categories. Thus, it can be argued whether it is reasonable to expect written CF to be effective when a heavy cognitive load is placed on learners who are not at advanced level proficiency. The participants in the above studies are at beginner or intermediate level proficiency. Thus, written CF may have better efficacy when it only focuses on one or a limited number of error types. A growing number of studies over the last 15 years have investigated the efficacy of *focused* written CF on L2 development. These studies are reviewed in the following section.

#### **3.2.2.2 Studies on focused new text**

Ellis (2005), Gass (1997) and Schmidt (1994) argued the importance of *attention* and *understanding* in cognitive information processing. Thus, it would seem that if written CF is given on only one or a few linguistic errors (i.e. focused written CF), learners may be more likely to *attend* to the focused written CF feedback and *understand* the reason for the error and how to correct it.

Focused written corrective feedback has been employed to investigate the efficacy of written CF over time. The majority of these studies have explored the efficacy of written CF on only one or a limited number of error types. A large number of these studies have been *highly focused* (written CF was provided on one error type only) while others have

been *less focused* (written CF was provided on a limited range of errors) (Ellis et al., 2008).

Bitchener and Knoch's studies adopted a very sound methodology and reported that written CF was effective on the functional use of articles over two months (Bitchener & Knoch, 2008), 10 months (Bitchener & Knoch 2010a) and 10 weeks (Bitchener & Knoch, 2010b). Written CF has also been reported to be effective on the use of English articles over nine weeks (Sheen 2007) and 10 weeks (Ellis et al., 2008). It has also been found that the effectiveness of written CF is retained for other targeted structures. A summary of these studies is presented in Table 3.5. This section has only introduced the focused studies with very sound methodology in new writing text. A detailed analysis of them was explained in section 3.2 of this chapter.

*Table 3.5 Recent focused studies investigating the efficacy of written CF in new texts*

<b>Studies</b>	<b>Participants</b>	<b>Error categories</b>	<b>Duration</b>
<b>Bitchener (2008)</b>	75 low-intermediate ESL learners	Two functional uses of the English article: definite article (anaphoric mention), indefinite article (first mention)	2 months
<b>Bitchener and Knoch (2008)</b>	144 low-intermediate ESL learners	Two functional uses of the English article: definite article (anaphoric mention), indefinite article (first mention)	2 months
<b>Bitchener and Knoch (2010a)</b>	52 low-intermediate ESL learners	Two functional uses of the English article: definite article (anaphoric mention), indefinite article (first mention)	10 months
<b>Bitchener and Knoch (2010b)</b>	63 advanced ESL learners	Two functional uses of the English article: definite article (anaphoric mention), indefinite article (first mention)	10 weeks (not circling group)
<b>Ellis et al. (2008)</b>	49 intermediate EFL learners	Two functional uses of the English article: definite article (anaphoric mention), indefinite article (first mention)	10 weeks
<b>Guo (2015)</b>	147 EFL learners at a Chinese university	Regular and irregular past simple tense; prepositions indicating space	19 weeks

<b>Mawlawi Diab (2015)</b>	57 ESL Lebanese learners	Pronoun agreement	9 weeks
<b>Rummel (2014)</b>	72 advanced EFL learners at Kuwait & Laos universities	Simple past tense; present perfect tense	7 weeks
<b>Sheen (2007)</b>	91 intermediate ESL learners	Two functional uses of the English article: definite article (anaphoric mention), indefinite article (first mention)	4 weeks
<b>Sheen et al. (2009)</b>	80 intermediate ESL learners	Two functional uses of the English article: definite article (anaphoric mention), indefinite article (first mention)	9 weeks
<b>Shintani and Ellis (2013)</b>	49 low intermediate ESL learners	Indefinite article	3 weeks
<b>Shintani et al. (2014)</b>	214 ESL learners at a Japanese university	Indefinite article, hypothetical conditional	4 weeks
<b>Shintani, Aubrey and Donnellan (2016)</b>	61 EFL Japanese at a Japanese university	hypothetical conditional	4 weeks
<b>Stefanou and Revesz (2015)</b>	89 Greek EFL learners	Article with generic & specific plural referents	4 weeks

Recent focused and unfocused studies have both shown that written CF can result in improved accuracy in new texts over time. However, two studies have compared focused and unfocused written CF with a single research design to find out which one of these approaches is more effective in improving learners' linguistic accuracy in new texts. These studies are discussed in the next section.

### 3.2.2.3 Studies comparing focused and unfocused written CF

A controversial debate in support of either unfocused or focused written CF for L2 development has been taking place over the last two decades. Those who support

unfocused argue that focused feedback is not helpful when learners are immediately in need of feedback on a wide range of errors. On the contrary, those supporting the need for focused written CF argue that students are more likely to attend to error corrections provided on a single error (or a limited range of errors) and are more likely to achieve a better understanding of the nature of the error and the correction required (Bitchener, 2016).

Two studies (Ellis et al., 2008; Sheen et al., 2009) have compared the efficacy of focused and unfocused written CF on L2 development. Ellis et al. (2008) conducted a study to compare the efficacy of unfocused and focused written CF in a Japanese university. The unfocused group was provided with correction on article errors and other types of errors while the focused group was only provided with direct error correction on article errors. The accuracy of focused and unfocused feedback was compared by examining the overall accuracy for unfocused feedback (article errors and other types of errors) versus the accuracy for focused feedback (i.e. articles). The study found that both unfocused and focused feedback were equally effective. Additionally, there was no significant difference between unfocused and focused groups in the delayed post-test. However, the researchers admitted that because article errors were highly represented in the two types of feedback, both types of feedback were not sufficiently distinguished from one another.

In the second study, Sheen et al. (2009) also compared the effectiveness of unfocused and focused written CF. Eighty ESL intermediate participants were divided into four groups: the focused written CF group received direct correction on a single grammatical target (the English article system); the unfocused written CF group was provided with direct correction on a broader range of grammatical structures (regular past tense, irregular past tense, articles, copula 'be', prepositions); the writing practice group was assigned to complete two writing tasks; and the fourth group was a control group. The researchers reported that focused written CF was more effective than unfocused written CF after nine



weeks. However, it is necessary to be cautious when interpreting the findings because, as the researchers acknowledged, the written CF in the unfocused groups was not systematic; that is, only some errors were corrected.

In summary, both focused and unfocused studies with sound methodological design have found written CF to be effective in L2 development; however, it is important to determine whether some types of written CF may facilitate L2 development more than others. Each type of written CF provides a different degree of explicitness and it is possible that more explicit feedback may be more effective than less explicit feedback. Because direct CF and ME feedback differ in their degree of explicitness, they may draw learners' attention to different extents. For instance, less explicit feedback is unlikely to be beneficial for low proficiency learners; instead, they require written CF such as metalinguistic CF that explicitly explains and illustrates the target structure. Thus, direct CF and metalinguistic explanation feedback may impact L2 development differently. Their comparative effectiveness with regards to other studies on L2 development is discussed in the following section.

### **3.3 The relative effectiveness of different types of written CF**

This section investigates the relative effectiveness of written CF types (i.e. direct CF and metalinguistic explanation). Research conducted on the relative merits of direct CF feedback and metalinguistic explanation feedback can be categorized according to whether a comparison has been made among (i) direct CF and less explicit types of written CF; (ii) direct CF and direct CF combined with more explicit types of written CF; and (iii) metalinguistic and other types of written CF. These are discussed in the following section.

### 3.3.1 Studies comparing direct CF and less explicit types of written CF

The studies that have compared direct error correction and less explicit types of feedback are summarized in Table 3.6.

*Table 3.6. Studies comparing direct error correction and less explicit types of written CF*

Studies	Feedback types	Effectiveness
<b>Lalande (1982)</b>	1. Direct error correction 2. Indirect coding	Indirect coding more effective than direct error correction but not statistically significant
<b>Semke (1984)</b>	1. Direct error correction 2. Content comments 3. Direct error correction and content comments 4. Indirect coding	No difference
<b>Chandler (2003)</b>	1. Direct error correction 2. Underlining 3. Error codes	Direct correction and underlining more effective than error codes but no difference between underlining and error codes
<b>Van Beuningen et al. (2008)</b>	1. Direct error correction 2. Indirect feedback 3. Writing practice 4. Self-correction revision	Direct error correction more effective long term; both direct and indirect feedback effective short term
<b>Van Beuningen et al. (2012)</b>	1. Direct error correction 2. Indirect feedback 3. Writing practice 4. Self-correction revision	Direct feedback more effective for grammar but indirect feedback more effective for non-grammar items

Lalande (1982) conducted research on 60 intermediate German EFL learners. The researcher found an advantage for indirect CF compared to direct CF, but the between group difference in accuracy improvement was not significant. However, the error codes employed in the study represent metalinguistic information rather than indirect feedback. The error code can be classified as less explicit metalinguistic feedback. In their study, Van Beuningen et al. (2012) argued that the indirect group performed more form-focused activities than the direct CF group. Semke's (1984) study of 141 German EFL learners found no significant difference among direct CF, comments and direct CF plus comments.

The study by Chandler (2003) involved 20 intermediate ESL learners and investigated the effectiveness of consecutively providing participants with direct CF, underlining and error codes. The researcher found that both direct CF and underlining were significantly more effective than error codes in improving accuracy in new texts over time. Additionally, the study found no significant difference between direct CF and underlining. Because learners in this study were provided with consecutive treatment rather than one treatment, the study cannot be compared with the two studies previously discussed in this set. Thus, none of the above studies provided evidence in support of direct CF being less or more effective than indirect, less explicit types of written CF.

Van Beuningen et al. (2008, 2012) managed to avoid methodological flaws and they reported that even though both direct and indirect groups showed significant improvement in accuracy in the short term, direct CF resulted in a more significant long-term effect than indirect written CF. These findings are in line with the theoretical explanations in Chapter 2 section 2.6 where it was argued that more explicit types of written CF (direct CF) may be more effective than less explicit types of written CF (indirect CF) for learners at a low level of proficiency. Even though Van Beuningen et al.'s (2008, 2012) studies were reliable and valid, caution is needed in drawing on only two studies to conclude the superiority of direct CF over indirect CF.

### **3.3.2 Studies comparing direct CF and direct CF with more explicit types of written CF**

A number of studies have compared the efficacy of direct CF with direct CF that is accompanied by more explicit forms of written CF. These studies are summarized in Table 3.7. As a large number of ESL and EFL language learning classrooms have combined direct CF with more explicit types of written CF, from a pedagogical view it is important to investigate this approach.

*Table 3.7 Studies comparing direct error correction and direct error correction plus more explicit forms of written CF*

<i>Studies</i>	<i>Participants</i>	<i>Target form</i>	<i>Types of direct written CF</i>	<i>Effectiveness</i>
<b>Bitchener et al. (2005)</b>	52 advanced ESL learners	Definite and indefinite articles; prepositions; past simple tense	1.Direct error correction 2.Direct error correction plus written ME 3.Direct error correction plus written and oral meta-linguistic explanation	Group 2 more effective than direct error correction for articles and past simple tense only
<b>Bitchener (2008)</b>	73 low intermediate ESL learners	First mention of indefinite article; anaphoric mention of definite article	1.Direct error correction 2.Direct error correction plus written ME 3.Direct error correction plus written and oral meta-linguistic explanation	Groups 1 and 3 more effective than group 2
<b>Bitchener and Knoch (2008)</b>	144 low intermediate ESL learners	First mention of indefinite article; anaphoric mention of definite article	1.Direct error correction 2.Direct error correction plus written ME 3.Direct error correction plus written and oral meta-linguistic explanation	No difference in types of written CF
<b>Bitchener and Knoch (2010a)</b>	52 low intermediate ESL learners	First mention of indefinite article; anaphoric mention of definite article	1.Direct error correction 2.Direct error correction plus written ME 3.Direct error correction plus written and oral meta-linguistic explanation	No difference in types of written CF
<b>Sheen (2007)</b>	91 intermediate ESL learners	First mention of indefinite article; anaphoric mention of definite article	1.Direct error correction 2.Direct error correction with meta-linguistic explanation	No difference between two treatment groups in immediate post-test but group 2 more effective than group 1 over two months
<b>Stefanou and Revesz(2015)</b>	89 EFL learners in Greece	Articles with generic & specific plural referents	1. Direct error correction 2. Direct error correction with meta-linguistic explanation	No difference in types of written CF

The study by Bitchener et al. (2005) involved 52 advanced ESL migrant learners over 12 weeks. The researchers compared direct written CF and direct CF plus oral metalinguistic explanation (in the form of five-minute one-to-one conferences). They found that the

group who received both direct CF and oral metalinguistic explanation CF showed an improvement in using the definite article and the simple past tense in new pieces of writing over time, in contrast to those who received direct written CF alone. Bitchener (2008) also conducted a two-month investigation on the efficacy of different types of written CF: (1) direct CF and oral and written meta-linguistic explanation; (2) direct CF and written metalinguistic explanation; (3) direct corrective feedback only; and (4) the control group that received no corrective feedback. The focus was only the two functional uses of the English article system (referential indefinite “a” and referential definite “the”) with 75 low intermediate ESL learners. The study found that Groups 1 and 3 were more effective than Groups 2 and 4. However, other studies with lower intermediate ESL learners on the use of “the” and “a” showed no advantage in adding metalinguistic explanation to direct CF in order for the CF to be more effective than direct CF alone (Bitchener & Knoch, 2008, 2010a).

Bitchener and Knoch’s (2008) study on 144 low intermediate ESL learners found no differences among the treatment groups. Similarly, in another longitudinal study (10 months) Bitchener and Knoch (2010a) examined the relative effectiveness of the different types of the same feedback approaches on two functional uses of the English article system. The participants were 52 low-intermediate ESL students in New Zealand. The findings revealed that all three groups who received treatments outperformed the control group on all post-tests. They also found that there was no difference among the three treatment groups. Stefanou and Revesz (2015) also found no advantage in adding written metalinguistic explanation to direct CF.

In her study, Sheen (2007) carried out an investigation into the effectiveness of two types of CF (i.e., direct CF and direct CF plus metalinguistic CF) on 91 intermediate ESL learners’ acquisition of articles. She reported no difference between the two treatment groups in the immediate post-test; however, in the delayed post-test (two months later),

it was found that direct CF plus written metalinguistic explanation was more effective than direct CF alone. Sheen suggested that the passage of time may have been a major factor in terms of the effectiveness of feedback types on L2 development.

The mixed findings of these studies show a need for further research on employing combined types of feedback prior to drawing any conclusions. There is a possibility that providing learners with separate feedback (e.g., separating the feedback so learners either receive direct CF or metalinguistic explanation) rather than a combination of two or more feedback types may result in more consistent findings. Thus, in the following section, a number of studies are discussed that have adopted this approach and compared the efficacy of metalinguistic explanation and other types of written CF.

### 3.3.3 Studies comparing metalinguistic explanation and other types of written CF

Several recent studies that have investigated the relative efficacy of metalinguistic CF and other types of written CF are summarized in Table 3.8.

*Table 3.8 Studies comparing metalinguistic feedback and other types of written CF*

Studies	Participants	Target form	Feedback types	Effectiveness
<b>Bitchener and Knoch (2010b)</b>	63 advanced ESL learners	First mention of indefinite article; anaphoric mention of definite article	1. Written metalinguistic explanation 2. Underlining/circling 3. Written & oral metalinguistic explanation 4. Control	No difference between 3 treatment groups in the short term. Groups 1 and 3 outperformed group 2 over time (10 weeks)
<b>Guo (2015)</b>	147 Chinese EFL learners	Regular and irregular simple past tense; prepositions indicating space	1. Direct error correction 2. Direct error correction plus written metalinguistic explanation 3. Written metalinguistic explanation 4. Underlining 5. Error code 6. Control	Groups 1-3 outperformed groups 4, 5 and 6. There was no significant difference between groups 1-3

<b>Mawlawi Diab (2015)</b>	57 ESL Lebanese learners	Pronoun agreement	1.Direct error correction plus metalinguistic explanation 2. Metalinguistic explanation	No difference between 2 treatment groups overtime
<b>Rummel (2014)</b>	72 advanced EFL learners at Kuwaiti and Laotian universities	Simple past tense; present perfect tense	1.Direct error correction 2.Indirect error correction 3.Written metalinguistic explanation 4. Control	No difference in types of written CF for Laotian learners; however, the Kuwaiti direct error correction group outperformed the other groups
<b>Shintani and Ellis (2013)</b>	49 low intermediate ESL learners	Indefinite article	1.Direct error correction 2.Metalinguistic explanation 3.Control	Group 2 were more effective than group 1 in the immediate post-test but not in the delayed post-test
<b>Shintani et al. (2014)</b>	214 EFL learners at a Japanese university	Indefinite article; hypothetical conditional	1. Direct error correction 2. Direct error correction plus revision 3.Metalinguistic explanation 4.Metalinguistic explanation plus revision 5. Control	Direct error correction more effective than metalinguistic explanation over time
<b>Shintani, Aubrey and Donnellan (2016)</b>	61 EFL Japanese at a Japanese university	hypothetical conditional	1.Pre-task metalinguistic explanation 2.Post-task metalinguistic explanation 3.Control	Pre-task metalinguistic explanation more effective than Post-task metalinguistic explanation overtime

In the first set of these studies, Bitchener and Knoch (2010b) investigated the effectiveness of different types of CF on students' learning and acquisition of articles. The participants were 63 advanced L2 students at a university in the USA. They were divided into three treatment groups which received written metalinguistic explanation, indirect underlining/circling ME and written and oral metalinguistic explanation and one control group. The targeted forms were the functional uses of articles ("a" and "the"). The findings revealed that there were no significant differences in efficacy among the three treatment groups in the immediate post-test. However, they found that there were significant differences between the metalinguistic groups and the indirect

underlining/circling group over time (after 10 weeks). As Sheen (2007) suggested, the passage of time may have been a critical factor in terms of the effectiveness of delayed post-tests. In her study on 147 low intermediate EFL learners, Guo (2015) also reported that learners who were provided with more explicit types of feedback (metalinguistic explanation, direct error correction, direct error correction plus metalinguistic explanation) outperformed those who were provided with less explicit types of written CF (error code and underlining). The researcher also found no significant difference among the three most explicit types of written CF. These findings were corroborated by Shintani and Ellis (2013) who investigated the comparative effect of different types of written CF (metalinguistic explanation and direct CF) on 49 low-intermediate ESL learners' accurate use of the English indefinite article. They reported that even though the metalinguistic group outperformed the direct correction group in the immediate post-test, there was no difference between the efficacy of metalinguistic explanations and direct correction groups in improving accuracy over time (after two weeks).

Similarly, in her study of Laotian EFL learners, Rummel (2014) found no difference between the direct error correction group, the indirect error correction group and the metalinguistic explanation group in treating target structures (simple past tense and present perfect). However, in the same study, Rummel found that for Kuwaiti learners, the group that received direct error correction outperformed the other two treatment groups. The researcher recommended that the differences in teaching and learning approaches in Kuwait and Laos as well as the learners' beliefs towards different types of written corrective feedback may have produced the different findings in Kuwait and Laos.

Shintani et al. (2014) reported similar findings to those of Rummel (2014) in the context of Kuwait. Shintani et al. compared the effects of direct CF and ME on 214 Japanese learners' accurate use of two grammatical structures: the hypothetical condition and the indefinite article. The participants were first and second-year university students majoring



in a variety of subjects. The students were provided with both types of CF and were not given an opportunity to rewrite. Accurate use of the targeted structures was then measured in a new writing task. The researchers found that direct CF was more effective than ME in the delayed post-test (after two weeks). They also compared the effectiveness of these two types of feedback when revision was added. They found that direct error correction plus revision was more effective than metalinguistic explanation plus revision. The authors suggested that the reason the findings differed from those of Shintani and Ellis (2013) may have been that, in the earlier study, there was a single structure while, in the second study, the focus was on two structures which may have led to an overload of information for learners at a lower proficiency level.

The study by Shintani, Aubrey and Donnellan (2016) resulted in similar findings to those of Shintani et al. (2014). In their recently published study of 61 Japanese EFL learners, Shintani et al. (2016) investigated the comparative efficacy of pre- and post-task metalinguistic explanation on learners' improved accuracy of the hypothetical condition. Similar to Shintani et al.'s (2014) study, learners received metalinguistic explanation in the form of a handout. The pre-task group was provided with metalinguistic explanation before the writing task and the post-task metalinguistic explanation group was provided with ME following completion of the writing task. The results showed that both pre- and post-task groups had improved accuracy in the short term but only the pre-task group maintained accuracy over time (after 3 weeks). Because of the similarity of findings to those of Shintani et al. (2014), the researchers suggested that metalinguistic explanation mainly led to developing explicit knowledge, which is less durable than automatized or implicit linguistic knowledge. However, as Bitchener and Storch (2016) argued, before generalizing and drawing any conclusion regarding the efficacy of written CF types, the effectiveness of written CF types such as ME needs to be investigated taking into account learners' individual and contextual factors and different linguistic structures.

In a recent study, Mawlawi Diab (2015) also investigated the efficacy of direct plus metalinguistic explanation and metalinguistic explanation alone on 57 ESL learners' accurate use of pronoun agreement at an American university in Lebanon. The findings showed improved accuracy for both experimental groups in new writing texts in the immediate context and over time. The researcher argued that the rule-based nature of pronoun agreement led to improved accuracy in both experimental groups. Additionally, the findings showed that direct plus metalinguistic correction was more effective than metalinguistic correction alone in the immediate post-test. However, there was no significant difference between groups in delayed post-tests. The researcher suggested that learners succeeded in retaining the procedural knowledge of the pronoun agreement rule over time (i.e., 9 weeks) facilitated by the rule-based nature of pronoun agreement.

In summary, the mixed findings of the above studies reveal that while there has been growing interest in testing the efficacy of different types of written CF in different ways, there are as yet no firm conclusions. The studies also reveal that there is currently a dearth of research to support the view that metalinguistic explanation can be used in place of other types of written CF, and that more research needs to be conducted in this regard. However, as a whole, the findings reveal that written CF is effective in comparison to no CF, and some form(s) of written CF (direct and metacognitive explanation) may help learners to improve their accuracy further. However, there is again a lack of research to support this view. This is the issue that this study will endeavour to investigate, that is, the comparative effectiveness of direct error correction and metalinguistic explanation on L2 development. Additionally, the recent studies by Shintani et al. (2013) and Shintani et al. (2014) showed that the efficacy of different types of written CF vary when different error types are investigated. It seems error types may act as a moderating factor on the efficacy of different types of written CF. In order to investigate the extent to which error types impact the efficacy of different types of written CF, there is a need to design more

studies (Bitchener, 2016). Thus, in the following section includes a discussion of the effect of error types on the efficacy of different types of written CF.

### **3.4 The effect of feedback on different grammatical structures**

This section investigates the moderating effect of linguistic error type on the effectiveness of different types of written CF. It was highlighted in the previous section that a number of variables (including linguistic error type and written CF type) may interact with one another and thus impact the extent to which written CF impedes or facilitates the accuracy of learners' output. The reason for this is that the development of morphological, syntactic and lexical items needs an understanding of meaning, form and use in relation to other parts of the language system and to other words (Ortega, 2009) and students may need to focus their attention on more than one linguistic element each time they use a form or structure. Some forms and structures may develop more easily than others as a result of the written CF provided. Thus, investigating the extent to which error type may moderate the efficacy of written CF has led to a growing number of research studies over the decades.

The review of the studies in the previous sections showed that English indefinite and definite articles have been extensively investigated as error types in the recent written CF studies, and the findings have shown that written CF on the English article system can facilitate L2 development. However, a number of studies have shown mixed findings when targeting other linguistic error types, for instance, prepositions, simple past tense and other structural errors (e.g., Bitchener et al., 2005; Shintani et al., 2014). Thus, the research has so far been too limited to conclude whether or not written CF is effective in targeting certain linguistic error domains and categories. This area, therefore, requires further exploration (Bitchener & Ferris, 2012). Specifically, it is important to investigate

whether there is any interactional relationship between the complexity of error type and the degree of explicitness of written CF. Thus, the present study investigates the effectiveness of different types of written CF on the passive voice as a complex structure. Table 3.9 shows a number of recent studies, along with the error type that was targeted, and the findings regarding the extent to which written CF facilitated learners' accuracy in using the targeted forms/structures.

*Table 3.9 The efficacy of written CF for developing accuracy in the use of linguistic form and structure*

<b>Studies</b>	<b>Linguistic focus</b>	<b>Findings</b>
<b>Bitchener et al. ( 2005)</b>	Articles, simple past tense, prepositions	Effective for articles and simple past tense
<b>Bitchener (2008); Bitchener &amp; Knoch (2008, 2009b, 2010a, 2010b)</b>	Indefinite article “a” for first mention & definite article “the” for subsequent or anaphoric mentions	Effective for both
<b>Sheen (2007)</b>	Indefinite article “a” for first mention & definite article “the” for subsequent or anaphoric mentions	Effective for both
<b>Ellis et al. (2008)</b>	Indefinite article “a” for first mention & definite article “the” for subsequent or anaphoric mentions	Effective for both
<b>Frear (2012)</b>	Regular and irregular past tense	Effective for regular but not for irregular form
<b>Shintani &amp; Ellis (2013)</b>	Indefinite article	Effective in immediate post-test but not in delayed post-test
<b>Shentani et al. (2014)</b>	Indefinite article & hypothetical conditional	Effective for hypothetical conditional but not over time. Not effective for indefinite article.
<b>Rummel (2014)</b>	Past simple tense & present perfect tense	Effective for both
<b>Stefanou &amp; Revesz (2015)</b>	Article with generic & specific plural referents	Effective for both
<b>Guo (2015)</b>	Regular and irregular past tense, prepositions indicating space	Effective for irregular simple past tense but not over time

Ferris (1999) distinguished between errors as treatable and untreatable. She suggested that treatable errors are those that occur “in a patterned, rule-governed way” and can be

corrected by referring to grammar books (e.g. regular simple past tense), while untreatable errors are those where “there is no handbook or set of rules students can consult” (e.g. irregular simple past tense) (p. 6). This distinction has been examined by other studies. For example, Frear (2012) provided evidence to support the hypothesis that errors in the regular past tense can be treated while those in the irregular past tense are possibly less treatable. Frear found that students who received focused direct written CF on the regular past tense, but not on the irregular past tense, showed improvement in new pieces of writing. Additionally, Bitchener et al. (2005) suggested that written CF is possibly more effective with treatable errors. They found that the students who received written CF had more accuracy in their use of the past tense and definite article, both regarded as patterned and rule-governed, but not in the use of prepositions, which can be considered as a more idiosyncratic item-based feature. However, Bitchener et al. showed the efficacy of written CF in targeting the use of articles in general. Follow-up studies (Bitchener, 2008; Bitchener & Knoch 2008, 2009, 2010a, 2010b; Ellis et al., 2008; Sheen, 2007) found the effectiveness of written CF in targeting both the indefinite article for first mention and the definite article for anaphoric or subsequent mentions. Two recent studies (Shintani & Ellis, 2013; Shintani et al., 2014) argued that because learners tend to overgeneralize the use of the definite article, knowing whether or not they have learnt the definite article for a specific grammatical function is difficult. The researchers explained that “restricting the analysis to ‘a’ for first mention allows for a more reliable scoring of the effect of instruction on acquisition” (Shintani & Ellis, 2013, p. 292). These two studies reported that written CF was ineffective for improving accuracy in the use of articles over time (two weeks), but in the first study, Shintani and Ellis (2013) found that providing learners with metalinguistic explanation resulted in improved accuracy in the immediate post-test. By drawing on the results of interviews with the participants, the researchers argued that

learners had given less attention to the indefinite article than to the hypothetical conditional (the other targeted structure in their study).

In addition, a number of studies have examined the efficacy of written CF on the improvement of accuracy by targeting the simple past tense and prepositions. Bitchener et al. (2005) and Rummel (2014) found that written CF helped learners improve accuracy in the use of the simple past tense; however, they did not make a distinction between the irregular and regular simple past tense in their studies. Frear (2012) compared the regular and irregular past tense and found that written CF improved the accuracy of regular past tense use over seven weeks, reporting that because the regular past tense is rule-based, learners had less difficulty in acquiring this form. Thus, it is expected that a rule-based form (e.g. regular past tense) may be easier to learn than an item-based form (e.g. irregular form).

Two studies examined the effectiveness of written CF on improved accuracy in the use of prepositions. In their study, Bitchener et al. (2005) found that written CF was not effective and argued that prepositions have many subcategories (for instance, prepositions indicating time, direction, space and so on) and those that have specific rules for specific functions may be more easily developed through written CF than those that, depending on the individual's stylistic preference and linguistic environment, are used for various functions. Additionally, Bitchener et al. suggested that it is possible that the accuracy of some of these sub-categories, especially those that are frequently used, may be improved if they are targeted by written CF. As stated, the prepositions targeted in the study by Bitchener et al. (2005) were from different categories. In a follow-up study, Guo (2015) found that written CF was not effective on improving the use of prepositions of space. However, there is a need for further research on the other sub-categories and a comparison

of these different types of prepositions must be made before any conclusions are drawn (Bitchener, 2016).

The studies discussed in this section have shown that the effectiveness of written CF on only a few linguistic error categories; thus, it is necessary to investigate the effectiveness of written CF on different error types. Additionally, most of these studies have targeted *forms*, while the effectiveness of written CF on improved accuracy in the use of *structure* is still in need of investigation. There are only two studies that have targeted structures: Shintani et al. (2014) and Rummel (2014) targeted the hypothetical conditional and present perfect tense respectively. While Rummel (2014) reported accuracy gain in the use of the present perfect tense over time, Shintani et al. (2014) found that learners did not sustain improved accuracy in the use of the hypothetical conditional over time. Thus, since structures have only been investigated in two studies with different findings regarding the improved use of these structures, the present study investigates the effectiveness of written CF by targeting the structure of the passive voice. The passive voice is a complex structure because it is categorized as both rule-based and item-based. In the passive voice the verb “to be” is categorized as rule-based as there is a rule that identifies when learners need to use “*am*”, “*is*” or “*are*”. However, the past participle is item-based as it takes three different forms: regular verbs that end in “*ed*” (e.g., “*cleaned*”); irregular verbs in which the past participle changes form (e.g., “*grown*”); and irregular verbs where the past participle retains the same form (e.g., “*spread*”).

The above discussion has shown that the effectiveness of written CF may vary when different types of linguistic forms/structures are targeted. In addition to linguistic factors, individual differences may also have a moderating effect on how effective written CF can be. In other words, the effect of individual factors on L2 development may differ from individual to individual (Bitchener, 2012; Ellis, 2008). Understanding the extent to which

individual factors may facilitate or impede L2 development may help to understand why some learners learn from the feedback provided while other learners seem to fail to learn from the feedback they receive (Bitchener, 2016). There has been a great number of studies on the impact of individual learner differences on L2 acquisition (e.g., Li, 2013; Mackey *et al.*, 2002; Re'v'e'sz, 2012; Shintani & Ellis, 2015); however, few studies have investigated the effect of these differences on learners' responses to written CF. Studies that have investigated the extent to which individual factors may moderate the effectiveness of written CF are discussed in the following subsection.

### **3.5 Individual Factors and Written Corrective Feedback**

This section firstly provides a general overview of a number of empirical studies on the moderating effect of individual factors on written CF. Then in the following section (3.5.1) the studies focusing specifically on the potentially moderating influence of another individual difference (working memory) on the effectiveness of written CF on L2 development are discussed.

Individual learner factors include motivation, working memory, language aptitude, age, personality, learning style, language anxiety, and learners' attitudes and beliefs. Several studies have investigated the moderating effect of language aptitude on the efficacy of written CF. Sheen (2007, 2011), for example, studied the potential influence of three learner factors (analytic ability, anxiety, and learners' attitudes towards error correction) on learners' uptake and retention after receiving different types of corrective feedback. Sheen (2007) found that students with high analytic ability benefited more from direct written CF and direct metalinguistic CF than learners with lower analytic ability. She also reported that learners with high aptitude benefited more from metalinguistic explanation than direct CF. Moreover, Sheen



(2011) found that while learners' analytic ability and their attitudes impacted the effects of written CF, anxiety had no impact on written CF.

In a more recent study, Stefanou and Revesz (2015) investigated the moderating effect of learners' grammatical sensitivity and knowledge of metalanguage on the effectiveness of direct written CF and direct written CF plus metalinguistic explanation. They found that learners who were provided with direct written CF alone and had greater grammatical sensitivity and more knowledge of metalanguage were more able to improve the accuracy of their use of article.

Learners' belief and goals have also been investigated in written CF studies. For instance, Storch and Wigglesworth (2010) used a case study approach to investigate students' beliefs about two types of CF. They found that when CF contradicted advanced learners' beliefs, they did not engage with it and did not draw on it when revising or rewriting their texts. However, the researchers pointed out that the finding may not represent the case of lower proficiency level students. Storch and Wigglesworth (2010) also found that the goals of individual students impacted their uptake; for example, learners who sought to enhance their writing accuracy achieved higher levels of uptake. Similarly, Hyland (1998, 2000, 2003) found that students who attached more importance to grammatical accuracy benefited more from CF; however, they resisted feedback if they felt the teacher had too much control over the feedback process or their individual goals were not taken into consideration.

Even though the findings from these few studies do not lead to any firm conclusions about the moderating effects of individual factors on the effectiveness of written CF, they suggested that the output arising from written CF may be moderated by individual factors. No study to my knowledge has examined the moderating effect of working memory on the effectiveness of written CF types, even though this individual difference factor has been investigated with respect to oral CF.

### **3.5.1 Studies on working memory, phonological short-term memory and oral corrective feedback**

Working memory is a cognitive device with the dual function of storing and processing information. Working memory is operationalized as either complex working memory, which refers to both storage and processing components, or phonological short-term memory (PSTM), which only consists of the storage component.

Some empirical literature suggests that phonological short-term memory plays an important role in L2 development. For instance, phonological short-term memory has been shown to be related to L2 aspects of speech production (Kormos & S´af´ar, 2008; O’Brien, Segalowitz, Freed, & Collentine, 2007), L2 oral performance (Ellis & Sinclair 1996; O’Brien, Pollett, Gallinger, & Dick, 2006), L2 grammar learning (Speciale, Ellis, & Bywater 2004) and reading comprehension (Masoura & Gathercole, 1999, 2005; Papagno, Valentine, & Baddeley, 1991; Service & Craik, 1993; Service & Kohonen, 1995).

Several studies have been conducted on the relationship between phonological short-term memory and the efficacy of corrective feedback, and these studies are related to recasts. Trofimovich et al. (2007) reported that learners with superior phonological loop ability benefit more from recasts in the longer term; however in their study the testing was conducted 2 to 12 minutes following the immediate post-test. Revesz (2012) also found a significant positive correlation between phonological short-term memory and recasts in an immediate oral post-test. Thus, the literature shows unclear findings regarding the role of phonological short-term memory in processing recasts.

Several studies have also investigated the relationship between complex working memory and the efficacy of recasts (Li, 2013; Goo, 2012; Mackey et al., 2002; Revesz, 2012; Sagarra, 2007; Trofimovich et al., 2007). Mackey et al. (2002) conducted a study on Japanese EFL learners to investigate the relationship among noticing recasts, working

memory and the efficacy of recasts in the learning of the target structure (English question formation). The researchers found that learners with low working memory achieved more gain from recasts in the short term, but learners with high working memory capacity noticed the recasts better in the long term. Goo (2012) confirmed the noticing function of working memory and the relationship with recasts. He found working memory moderates recasts, but there was no correlation between working memory and metalinguistic feedback and he argued that the explicit nature of metalinguistic feedback neutralized the participants' individual differences in their ability to notice the target language. Li (2013) also reported that complex working memory did not predict the effect of recasts on the learning of the target structure (*-le*) in grammatical judgment and elicited imitation tests. However, in the metalinguistic feedback group, working memory capacity negatively moderated performance on the delayed grammar judgement test.

Both Sagarra (2007) and Trofimovich et al., (2007) investigated the moderating effect of working memory in the effectiveness of computerized recasts. The former study found that the effectiveness of recasts was related only to working memory; however, this was not the case in the latter study. Goo (2012) argued that the mixed findings were due to the letter-string test that Trofimovich et al. (2007) employed, which did not measure the processing component of working memory.

One of the main limitations of the above studies (except Li , 2013) is that the tests they employed only measured the storage component and did not measure reaction time and veracity judgment (indicators of the processing component of working memory). Thus, this is an area that needs consideration in working memory studies. Further, the studies discussed above only investigated the moderating effect of working memory on the effectiveness of oral corrective feedback. So far, however, no research has examined how complex working memory and phonological short-term memory might mediate learners use of different types of written CF. This study intends to investigate this issue. Kormos

(2012) pointed out that working memory may moderate the effectiveness of written CF types. She argued that in contrast to the oral context, learning opportunities through feedback in the writing context are less constrained by time pressure; however, because writing learners are dependent on their working memory capacity, they may respond differently to feedback. Additionally, in contrary to oral feedback which is online, written CF is more offline and so there is more time to process feedback, so working memory may play a better role in written CF because it deals with processing over a longer period of time. While in oral CF it has to be instantaneous. Thus, this study has investigated the moderating effect of working memory on written CF.

In summary, this chapter has shown that even though a growing amount of research has contributed to written CF knowledge, there are further questions that need to be addressed. For instance, does focused written CF facilitate L2 development immediately and over time? Does revision have an impact on improved accuracy in the short and long term? Is one type of written CF (especially direct CF, metalinguistic explanation, direct CF plus revision, metalinguistic explanation, and metalinguistic explanation plus revision) more effective in improving accuracy than other types? Does written CF work better or worse in treating complex structures (such as the English passive voice)? Do learners' individual differences in cognitive processing (i.e., working memory and phonological short-term memory) moderate the effectiveness of different types of written CF? These questions are addressed in this study.

### **3. 6 Rationale of the present study and research questions**

The study has two main purposes: the first aim is to examine the effect of different types of written CF (direct and metalinguistic) on learners' output (immediate and delayed) in relation to the targeted structure (passive voice). The motivation for exploring this aim is

drawn from the mixed findings from several recent studies (e.g., Shintani & Ellis, 2013; Shintani, et al., 2014) on the effectiveness of written CF. Additionally, it is still not clear how explicit the feedback needs to be and whether certain types of CF are more effective than others in facilitating L2 development (Bitchener, 2016). For this reason, the study focuses on two different types of written CF (direct and metalinguistic) in order to show to what extent the type of corrective feedback and its explicitness can impact on learners' subsequent output (immediate and delayed).

The second purpose of this study is to determine the extent to which individual differences in working memory and phonological short-term memory may moderate different types of written CF (direct and metalinguistic), and whether these differences have an effect on learners' subsequent output (immediate and delayed). The second aim is in response to Bitchener and Storch's (2016), Bitchener and Ferris' (2012) and Ellis' (2010) calls for an investigation into the moderating effect of individual factors on how learners respond to and use written CF. Understanding whether certain types of individual factors have the potential to moderate progress (that is, the extent to which they may have an impeding or facilitating effect on L2 development) helps researchers understand why some learners are able to develop their L2 knowledge more easily than others and why some learners are successful in learning from the feedback provided and the others fail to learn from the feedback. To date, very few studies have examined the moderating effect of individual factors, from a cognitive perspective, on the effectiveness of written CF on L2 development. This study, therefore, has investigated the moderating effect of working memory and phonological short term memory on the efficacy of written CF types (namely, direct CF, direct plus revision, metalinguistic explanation and metalinguistic plus revision). In order to address the above purposes of the study, the following research questions were designed.

RQ1: What effect do focused direct corrective feedback (DCF) and metalinguistic explanation (ME) with and without revision have on learners' use of the English passive voice in an immediate text revision and in new texts over time?

RQ2: Does the opportunity for the revision influence in the efficacy of DC and ME?

RQ3: Is there any difference in the effect of DCF and ME regardless of whether there is an opportunity for revision?

RQ4: To what extent do working memory and phonological short-term memory moderate the effects of the different types of feedback?

The next chapter, Methodology, will include an outline of how these RQs were operationalized.

## **CHAPTER 4**

### **METHODOLOGY**

#### **4.1. Introduction:**

This chapter presents the methodological approach, research design, and the methods employed in this thesis to collect the data needed to answer the research questions. In section 4.2 the aims of the research are contextualized by discussing past studies that have inspired this research. Following this, section 4.3 explains the post-positivist methodological approach that underlies this study. Since this research draws on cause and effect thinking, detailed empirical observations, and measurement (Creswell & Plano Clark, 2011), a post-positivist approach is viewed as the most appropriate method to answer the research questions.

Section 4.4 explains how the study has been designed to answer the research questions. Section 4.5 provides the context for the study as well as detailed information about the participants and the research site in which the data were collected.

Section 4.6 describes the target structure of the study and provides the reasons for utilizing this structure. Following this, section 4.7 provides a detailed illustration and examples of the types of instruments that were employed to collect the data. The section also provides reasons for why these instruments are regarded as appropriate to answer the research questions (RQ1, RQ2, RQ3 and RQ4).

Section 4.8 discusses the data collection and accounts for the procedures and different stages employed when utilizing the instruments. Subsequently, section 4.9 illustrates the way the data were analysed. An account of the validity and reliability of the study is provided in section 4.10. This is followed by an overview of ethical considerations in section 4.11.

## **4.2 Scope of the Research**

The study has two main purposes. The first aim of the study is to examine the effect of different types of written Corrective Feedback (CF) (direct CF, metalinguistic explanation, direct CF plus revision, metalinguistic plus revision) on learners' output (immediate and delayed) in relation to the English passive voice. This aim was motivated by the mixed findings of recent studies (e.g., Shintani & Ellis, 2013; Shintani, Ellis, & Suzuki, 2014).

The second purpose of this study is to determine the extent to which individual differences in working memory and phonological short-term memory may moderate how learners respond to and use different types of written CF (direct CF, metalinguistic explanation, direct CF plus revision, metalinguistic plus revision), and if these differences have an effect on learners' subsequent output (both immediate and delayed). The second aim was in response to Bitchener's (2012) and Ellis' (2010) calls for more research into the moderating effect of individual factors on how learners respond to the written CF they receive.

## **4.3 Philosophical Approach**

According to Creswell (2013), "Philosophical assumptions are typically the first ideas in developing a study" (p. 17). As this study utilizes a quantitative approach, "the first ideas" that underlie this study are post-positivist. Researchers who draw upon post-positivism, including the researcher in this study, collect data objectively, using agreed-on definitions of variables and checks to eliminate bias (Creswell & Plano Clark, 2011).

Post-positivist researchers utilize knowledge based on "(1) determinism or cause and effect thinking; (2) reductionism, by narrowing and focusing on select variables to interrelate; (3) detailed observations and measures of variables; and (4) the testing of



theories that are continually refined (Creswell & Plano Clark, 2011, p. 40)”. In doing so, researchers collect data objectively, using agreed-on definitions of variables and checks to eliminate bias (Creswell & Plano Clark, 2011).

Hence, by adopting a post-positivist framework, this study endeavours to address its research questions by employing empirical observation and measurement. To do so, the independent variables (different types of written CF) have been manipulated with the purpose of determining whether they have any impact on dependent variables (working memory and phonological short term memory). Thus, a scientific method using a pre-test, treatment, and post-test has been conducted to address RQs 1-3; and an investigation was undertaken to determine the moderating effect of individual differences (RQ 4).

#### **4.4 Design**

This study employed a pre-test, treatment, post-test, and delayed post-test design, using intact English as a Foreign Language (EFL) classes. Many experimental studies utilize a pre-test-post-test design whereby a pre-test is administered prior to the experimental manipulation, and a post-test is administered following the manipulation. Accordingly, a pre-test-post-test design allows the researcher to assess the impact of the experimental manipulation by observing the difference between the pre-test and post-test (See Table 4.1).

In this study, four experimental groups (direct corrective feedback, direct corrective feedback and revision, metalinguistic explanation, metalinguistic explanation and revision) and one control group were included. The treatments were operationalized as four different written CF strategies: direct written CF, direct CF and revision, written metalinguistic CF and written metalinguistic explanation and revision. The control group did not receive any feedback on their writing texts to contrast the impact of treatment on

the experimental groups; instead, they were given very brief feedback on the quality and organization of their content. One week prior to the start of the CF treatment, participants completed a writing task as the pre-test. The immediate post-tests (revision and new tests) were conducted immediately after the CF treatment session has been completed in Week 2. Participants completed a working memory test (reading span test) and a phonological short term memory test (non-word span test) in Week 3. The delayed post-test was completed in Week 4.

*Table 4.1 Study Design*

	DC Group (N=20)	DC+R Group (N= 20)	ME Group (N= 20 )	ME+R Group (N= 19 )	Control Group (N= 21 )
Week 1	Written task, pre-test (Time 1)				
Week 2	DC (10 min)	DC (10 min) + Revision (30 min)	ME (10 min)	ME (10 min)+ Revision (30 min)	No Treatment
	Written task, immediate post-test (Time 2)				
Week 3	Working memory test (reading span test) (Time 3) Phonological short term memory test(non-word span test)				
Week 4	Written Task, delayed post-test (Time 4)				

**Note:** DCF= Direct corrective feedback

DCF+R= Direct corrective feedback and revision

ME= Metalinguistic explanation

ME+R= Metalinguistic explanation and revision

## **4.5 Context and Participants**

### **4.5.1 Context**

The Departments of Foreign Languages and Literature at two universities in Iran was chosen as the research site for the study. The departments offer courses in English Translation, English Literature, and Teaching English as a Foreign Language for Bachelor's degrees, and in Teaching English as a Foreign Language for Master's degrees.

As Duff (2008) points out, familiarity with the context and the participants works as an advantage for the researcher, as it enables him/her to gain "insider" status to facilitate the difficult process of "gaining entry to the research context and access to the case ... for a longitudinal study" (p. 117). Because I used to be a student at the context of data collection, I am quite familiar with the lecturers and site of the research, and thereby had "insider" status.

The rationale for using the context was that it provided the appropriate number of participants required for the study. Furthermore, the context provided the English proficiency level (intermediate level) required for this study. The reason for choosing the intermediate level was that although these students were familiar with the target structure (passive voice), they did not have the ability to employ the target structure accurately in a writing text. This conclusion was reached after sending samples of writing tests to several teachers in the context and gaining their opinion regarding the test.

### **4.5.2 Participants**

The participants were L1 Persian EFL learners at an intermediate level. They included both males and females who were 18-25 years old. The English proficiency of participants was at the intermediate level, which suited the study. Their proficiency level was

measured through an internal (Iranian) proficiency test. The test included sections on grammar, vocabulary and reading comprehension.

They were randomly divided into five groups – one class in which direct written CF was provided, one class in which direct written CF with revision was provided, one class in which metalinguistic explanation was provided, one class in which metalinguistic explanation with revision was provided, and one class that acted as a control group.

#### **4.6 Target Structure**

The target structure of the study was the English passive voice. Much of the research conducted on focused written CF (e.g., Sheen 2007; Frear, 2012) has concentrated on the same targeted structures such as English articles and the past tense. Thus, in order to ensure findings are more reliable, further studies are required on the use of different types of grammatical features such as syntactic structures and on the use of more than one targeted grammatical feature to measure the level of learners' retention over a long period of time. This study therefore used a complex structure that has not previously been the focus of written CF studies – the present simple English passive voice.

In a passive construction the object of an action takes the place of the subject of a sentence. The subject of the sentence (the agent or source of the action) may be omitted or added in a prepositional phrase.

Both passive and active constructions exist in English and Persian. In the English language, the passive voice is formed by a combination of an auxiliary and the past participle of a transitive verb. For example:

Active

Passive

*He opens the door.    The door is opened by him.*

In the Persian language, the passive voice is formed by a past participle and a derivation of the auxiliary verb "šodan".

For example:

سیبها چیده و شسته شده اند.

Birjandi, Maftoon and Rahemi (2011) argued that the passive voice is particularly problematic for Iranian EFL learners because they have to apply a default processing strategy that ascribes the position of subject/agent to the first noun or phrase they hear or see in the input. Thus, the reason that Iranian learners have difficulty in processing the English passive voice is that they need to process the first noun as the patient /object and not the agent. Pedagogically, the reason for using this structure as the focus of the study is that the English passive voice does not frequently appear in Iranian EFL textbooks, therefore adding to the difficulty faced by Iranian students in coming to terms with this structure (Birjandi, Maftoon, & Rahemi, 2011). Although many EFL textbooks include a chapter about the passive voice, most EFL learners, even at advanced levels, incorrectly form the passive voice in their speaking and writing (Hinkel, 2002). Additionally, the participants of the main study were Iranian students who had academic writing and scientific reporting as two components of their study. The ability to use the passive voice could be helpful for the students when writing their assignments for the course. Most academic writing texts do not focus on who is the “doer” in an action, but on who/what is experiencing or receiving the action. Furthermore, it is irrelevant or repetitive in academic writing to state who is doing an action; therefore, the passive voice can be employed to construct these types of sentences. For these reasons, the study focused on the simple present form of the passive voice. For example: *Diamonds are mined in South Africa.*

## 4.7 Instruments

Three different types of instrument were employed during the process of data collection. That is, writing tasks, a reading span test, and a non-word span test.

1. Students were provided with three writing prompts, each requiring an explanation of a process – for example, how chocolate is made. For each session (pre-test, immediate post-test, delayed post-test) one writing prompt was utilized. The writing tasks were used to address RQ1, RQ2 and RQ3.
2. A reading span test was administered to measure working memory capacity. This type of test has been employed in several studies as a measure of working memory capacity (e.g., Daneman & Carpenter, 1980; Harrington & Sawyer, 1992; Lesser, 2007; Light & Anderson, 1985; Osaka & Osaka, 1992; Salthouse & Frisk & Milner, 1990; Shahnazari, 2012; Swanson, 1993). Thus, it can be regarded as a valid measure of working memory. Accordingly, participants read a set of unrelated sentences and judged whether they make sense or not (processing assessment). Their reaction time during the judgment was measured. They were then asked to recall the final word of each sentence at the end of the set (storage assessment). For this task, DMDX software was utilized. DMDX is a Win 32-based display system employed to measure reaction times to auditory and visual stimuli. The reading span test was used to address RQ 4.
3. Generally, phonological short-term memory capacity is measured by digit or word span tests. The tests involve the recall of unrelated numbers and words presented in either aural or written form. However, since both digit or word span tests require some knowledge of the language, they confound phonological memory capacity.

One way to reduce this confounding is to utilize a non-word repetition task (NWR), in which nonsense words are presented aurally and participants are required to recall them (Baddeley et al., 1998). Thus, a non-word repetition task was conducted to answer RQ4.

The research instruments are discussed as follows.

#### **4.7.1 Student Writing Task**

The target structure in the study was the English simple present passive tense. Thus, process tasks were designed to motivate participants to use this tense. Process tasks deal with a number of stages that are in time order. Accordingly, it is necessary to start at the beginning and describe each stage through to the last one. Process tasks are generally targeted at the description of processes (how something happens) or procedures (how something is done). Process writing often uses the passive structure (Swales & Feak, 2001).

Participants were administered three process tasks in Weeks 1, 2 and 4. They were required to write one process tasks in each session. The rubric for all the tasks was the same: the participants were required to write approximately 200 words within 40 minutes for each task. All three tasks were similar in focus in that they were process writing tasks. The three task topics dealt with how chocolate is produced, how apples are canned, and how coffee is produced.

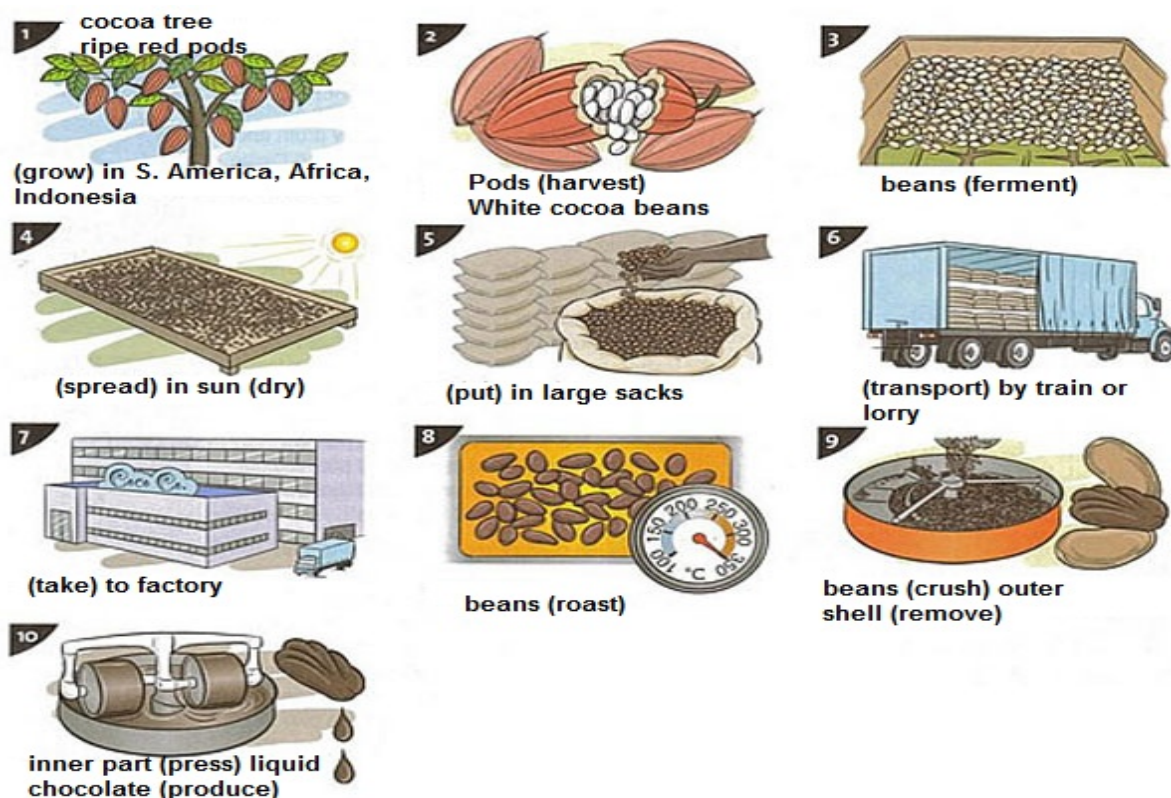
The following is an example of one of the tasks:

##### **Writing task 1:**

You should spend about 40 minutes on this task.

*The illustrations show how chocolate is produced.*

Summarize the information by selecting and reporting the main features and make comparisons where relevant.



In order to motivate (prompt) participants to use the English passive voice in their writing, three text-reconstruction tasks were administered. Accordingly, participants were provided a sample writing answer on each task and asked to read it for 5 minutes along with the task. Then, the sample answer was collected and they wrote their own version of this text according to the process diagram.

The following is an example of a sample answer for the above task:

### Sample Answer

*"The diagram illustrates the process for the manufacturing of chocolate which can be divided into 10 main stages, starting from the growth of cacao tree and ending in the production of chocolate.*

*Firstly, ripe red pods are collected from cacao trees. The pods are mainly grown in South America, Africa, and Indonesia. These pods are then harvested into outer shells and white cocoa beans. Then, the pods are fermented, spread, and dried in the sunshine. Following that, the pods are placed into large bags and transported to a factory by train or lorry.*



*Inside the factory, the beans are roasted in an oven at a temperature of about 350 degrees Celsius. Next, the roasted beans are transferred to a mill. In the mill, the beans are first crushed and the outer shells are removed. Then, the inner parts are pressed and liquid chocolate is produced. At this point, the process is completed and chocolate is available to be sold in the market”*

#### **4.7.2 Reading Span Test**

Daneman and Carpenter introduced the first reading span test in 1980. Such tests are utilized to measure working memory capacity and to give an index relating to the processing and storage of the components of working memory (WM). The reading span test includes two steps. First, learners read a set of unrelated sentences and judge whether they make sense or not (processing assessment). Second, they attempt to recall the final word of each sentence at the end of the set (storage assessment) (Daneman & Carpenter, 1980).

In this research, a Persian reading span test was conducted to measure WM. The Persian WM test included 64 sentences, that is 10 practice sentences, and 54 test sentences. The test was made up of 12 sets, and the sets were arranged with 3, 4, 5, and 6 sentences. The sentences were in an affirmative and active form within a range of 13-16 words. The test included both “sense” and “nonsense” sentences. In other words, half of the sentences were semantically plausible and the remaining half were implausible. This type of instrument is constructed by rearranging some content words so that sentences were syntactically acceptable, but semantically anomalous (Harrington & Sawyer, 2011; Lesser, 2007). This is to ensure that the students process sentences for meaning rather than focusing only on the retention of recall items. The test was administered individually using DMDX software. Accordingly, each sentence appeared on a computer screen for 7 seconds. The pilot tests revealed that 7 seconds is an appropriate time to read sentences within a range of 13-16 words. Following each set, a page appeared asking participants to recall the last word of each sentence in the set and write it on a piece of paper that was

provided at the beginning of the test. The test measured working memory components, namely processing and storage (e.g., Chun & Payne, 2004; Daneman & Carpenter, 1980; Lesser, 2007; Waters & Caplan, 1996). In order to assess working memory processing, firstly the participants were required to read each sentence. Then, they judged whether each sentence made sense or was nonsense. Participants gave their judgments by pressing on the letters M or Z on the keyboard of the computer and they stood for yes (sense) or no (nonsense) respectively. In order to avoid any confusion while using the keys, the letters Y and N, representing “yes” or “no”, were attached on the M and Z keys respectively. As a measure of processing efficiency, DMDX also recorded their reaction times when judging sentences. To measure storage, the participants were required to recall the final word of each sentence up to the end of each set until a visual prompt came up on the computer screen. At this time, the participants were required to write them on a piece of paper. In order to control for recency, they were needed to recall the final words in the order in which they appeared (Baddeley & Hitch, 1993; Waters & Caplan, 1996).

#### **4.7.3 Phonological Short-term Memory**

A non-word repetition task was administered to participants to measure their phonological short-term memory (PSTM) (Baddeley et al., 1998; Kormos & S´af´ar, 2008; Service, 1992; Service & Kohonen, 1995). The participants were required to listen to a sequence of non-words and then repeat each sequence. Non-words were employed in the test since they reduced the impact of vocabulary knowledge on phonological short-term memory, thereby leading to a relatively accurate measure of phonological short-term memory (Gathercole et al., 2001). Phonological short-term memory is important for the temporary storage of information without understanding. Additionally, phonological short-term

memory is important in terms of accuracy; that is, when learners accurately apply stored information (e.g., in the form of discrete corrected passive sentences) when rewriting.

This task was adapted from Gathercole et al. (2001). It was composed of 22 sequenced pairs of English non-words. The length of each sequence was gradually increased across the pairs within the range of 4 to 7 non-word syllables. Since the non-words were made from English letters, an English native speaker with six years of English teaching experience in the UK and New Zealand was asked to record the test. The task was conducted in Week 3. Each participant was tested individually, seated in a quiet room. The test was recorded on an HP laptop. Participants listened to each recorded sequence and then repeated each sequence. The instructions in the audiotape were: “You will hear some made-up words. Repeat them after hearing in the same order. Items A, B and C are for your practice. The test starts from number 1 and proceeds to 22.”

Participants’ repetition was recorded by the researcher for analysis. Their answer for each item was rated as correct or incorrect using a binary criterion of right or wrong.

For instance:

*1- peb kib bon deet*

*2- peeb kol goob mab*

*3- pib kom gook tam*

#### **4.8 Treatment Procedure**

All experimental groups received focused CF rather than unfocused CF. This decision was motivated by a number of studies (Bitchener & Knoch, 2008, 2010a, 2010b; Ellis, Sheen, Murakami, & Takashima, 2008; Frear, 2012; Shintani & Ellis, 2013; Sheen, 2007; Sheen, Wright, & Moldawa, 2009) that have found that focused CF leads to improvement in accuracy in new pieces of writing. To demonstrate the effectiveness of CF on output or retention, researchers need to show that the correction of a specific error results in the

reduction or elimination of the error in a new piece of writing. Statistically, focused CF has been shown to result in significant gains in grammatical accuracy in new writing tasks (Bitchener & Ferris, 2012; Shintani & Ellis, 2013). Furthermore, one of the advantages of focused CF is providing multiple corrections of the same error in a piece of writing. Thus, it is more likely learners will notice and attend to the error. Robinson (1995b) defines *noticing* as “detection with awareness and rehearsal in short term memory ... necessary to learning and subsequent encoding in long-term memory” (p.34). This follows Schmidt’s noticing hypothesis (1990, 1993, 1995) which claims learning cannot occur without noticing. In Schmidt’s model learners are required to notice linguistic output consciously in order for it to become intake. Thus, noticing can be crucial for the learning of the target structure in feedback studies as well. Multiple instances of feedback on the same error in one piece of writing can facilitate learners’ noticing of the target structure, especially when a complex structure such as the passive voice is targeted. Thus, focused CF was used in the study.

The study included four experimental groups and one control group.

**Group 1 DCF:** The participants completed task (1) in Week 1. To do so, they were provided with writing task 1 along with a sample answer, which they were asked to read for 5 minutes. The sample answer was then collected, and learners started writing task 1. In Week 2, the direct CF group received written CF on the writing they had produced in the pre-test. The corrective feedback focused only on the passive voice structure and the group was provided with 10 minutes to look over the corrections to their writings. Then, these first texts with feedback were collected and the group undertook their next task (2) (immediate post-test) using the same procedures as in the pre-test. Since the study also investigated the effect of feedback on the immediate post-test, it was necessary to collect the first texts with feedback. In Week 4, participants completed the delayed post-test

(writing task 3) following the same procedures as they had followed for the immediate post-test.

**Group 2 DCF+ R:** The direct CF plus revision group followed the same procedures as the DCF group, but in Week 2, participants were given 30 minutes to rewrite their first text. One of the aims of this study was to investigate how participants with different levels of phonological short-term memory can store feedback and use it when rewriting. Therefore, they were not allowed to look over their corrected text while rewriting. Phonological short-term memory involves storing information. Applying stored information (in this case in the form of discrete corrected passive sentences) assists learners' accuracy when rewriting. Furthermore, to encourage students to focus only on the target structure (the passive voice), and to determine if they could use the target structure accurately in the rewritten texts and the post-test, participants were asked to only copy the corrections while rewriting. Then, the rewritten texts were collected by teachers prior to completing the immediate post-tests. After that, students immediately wrote their second writing text (immediate post-test) following the same procedures as in the pre-tests. The learners wrote their fourth writing tasks in Week 4, following the same procedures as in the immediate post-test.

**Group 3 ME:** The metalinguistic explanation group completed their first task in Week 1. In task 1, participants were provided with a sample answer that they were asked to read for 10 minutes. The sample was then collected and participants began writing task 1. Participants did not receive any feedback on their writing in pre-test. In Week 2, the researcher gave participants their first written text and a handout with an explicit explanation about the targeted structure (the passive voice). The handout was in Farsi to

prevent any lack of English vocabulary affecting their understanding. It included the definition of the English passive voice, how to make the English passive voice and had some examples of the English passive voice in both Farsi and English. The group was given 10 minutes to read the handout and check their text to see if/where there were passive voice errors. Then, the teachers collected the handouts and the written texts. Students then immediately wrote their second writing text (immediate post-test) using the same procedures as the pre-test. In Week 4, the learners wrote their third writing task following the same procedure as the immediate post-test.

**Group 4 ME+R:** The metalinguistic explanation plus revision group completed their first writing task using the same procedures as the ME group in the pre-test. In Week 2, participants were provided with the first test they had completed in week 1 and the same handout that had been provided to the ME group (explicit explanation about the passive voice as the targeted structure). They were given 10 minutes to look over the handout. Then, the first text and the handout were collected and participants were given 30 minutes to rewrite their first texts. After that, the rewritten text was collected. Participants then immediately wrote their second writing text (immediate post-test) following the same procedures as in the pre-tests. The participants wrote their third writing task in Week 4, following the same procedures as in the immediate post-test.

**Group 5:** The control group did not receive any feedback on their writing texts. Participants performed their writing tasks using the same procedures as the experimental groups in Week 1. Then, in Week 2, participants were given their written texts back without any corrective feedback; instead, for 5 minutes they were given feedback on the quality and organization of their content. Then, their texts were collected and the group

was asked to undertake their second writing tasks. In week 4, participants undertook the delayed post-test, using the same procedures as in the immediate post-test.

#### **4.9 Data Analysis**

This section covers in detail the analyses of the data for RQ1, RQ2, RQ3 and RQ4. RQ1 investigated the effect of written direct CF and metalinguistic explanation with and without revision on learners' immediate and delayed output in relation to the passive voice as the target structure in new writing tasks over time; RQ2 investigated whether the opportunity for revision influences the effectiveness of direct CF and metalinguistic explanation. In order to address RQ2, the two revision groups (i.e., metalinguistic explanation plus revision and direct CF plus revision) and the two groups that did not make revisions (i.e., metalinguistic explanation and direct CF) were combined and compared. The DMPR group stands for the combined direct CF plus revision and metalinguistic explanation plus revision groups and the DMWR group stands for the combined direct CF and metalinguistic explanation groups. RQ3 explored whether the opportunity for revision influenced the effectiveness of direct CF and metalinguistic explanation. In order to address RQ3, that is, to investigate the relative efficacy of direct CF and metalinguistic explanation regardless of whether there is an opportunity for revision, the two direct CF groups (i.e., direct CF plus revision and direct CF) and the two metalinguistic explanation groups (i.e., metalinguistic explanation and metalinguistic explanation plus revision) were combined and compared. The DCO group stands for the combined direct CF groups and the MEO group stands for the combined metalinguistic explanation groups. RQ4 investigated the extent to which working memory and phonological short term memory mediated efficacy in the use of direct CF and metalinguistic explanation with and without revision in new writing texts over time.

This study is a quantitative method research, thus data analysis consists of analysing the quantitative data (Creswell & Plano Clark, 2011). Quantitative data was collected from writing tasks, reading-span tests, and non-word-span tests. The section begins with an illustration of how participants' writing tasks were scored followed by explanations of the statistical analyses used to calculate the accuracy scores in the written tasks. The section continues by outlining how statistical analyses were employed to draw comparisons between the different treatments of writing task scores. This is followed by a description of the method used to score working memory and phonological short-term memory as well as the statistical analyses used to investigate if working memory and phonological short-term memory mediate different types of written CF. The method of data analysis will be discussed further in the following.

#### **4.9.1 Writing Tasks:**

Research questions 1, 2 and 3 focus on the efficacy of each of the different types of written CF overtime and the comparative effects of the four types of treatment of learners' use of the English passive voice.

The data analyses for the three research questions are reported as follows:

##### **4.9.1.1 Scoring**

The simple present passive voice is the target structure in this study. In the English language, the simple present passive voice is formed with a combination of "to be" and the past participle. The conjugation of the verb "to be" depends on the subject (i.e., the object of the active statement) which can be either "am", "is" or "are". The past participle can also take three different forms: regular verbs that end in "ed" (e.g., "collected");



irregular verbs in which the past participle changes form (e.g., “grown”); and irregular verbs where the past participle retains the same form (e.g., “spread”).

The passive voice can be scored in two ways. In the first scoring method (Type 1), participants are scored on the passive voice only if they use both of its components accurately. In other words, if they use only one out of two components accurately they do not receive a score. Obligatory occasion analysis is then applied to calculate the total percentage score for each student. In the second method (Type 2), a point system is defined in order to score the use of the passive voice based on the accurate use of each its components (‘to be’ and the past participle), and then a partial obligatory occasion analysis is applied to calculate a total percentage score for each student. It can be argued that the Type 2 scoring method can provide a more accurate analysis of data than the Type 1 scoring method because the passive voice is made up of two components and students may acquire only one of its components instead of two as a result of feedback. Thus, they need to get credit for each correct component. Written corrective feedback may help them to learn both of the components. The Type 2 method also provides a score for participants who attempt to apply the passive voice but use it inaccurately, because it shows they are in the process of learning the target structure (see below). In contrast the Type 1 method fails to provide a score even if learners show development in one component of the passive voice. In other words, a score is only granted if complete learning is shown. The literature shows that previous studies that have targeted structural problems (e.g., the hypothetical conditional) have used a Type 2 point system to analyse data (e.g., Shintani & Ellis, 2014). Therefore, since the focus of this study is the passive voice with a recognition of the importance of both components of the structure, the study designed a scoring system based on Type 2 as described below.

To do so, if participants used the correct form of the passive voice, they received 2 points. That is, 1 point for the correct usage of “to be” and 1 point for the correct usage of the

past participle. If only one of these was correct, they received only 1 point. If both “to be” and the past participle were applied incorrectly, they did not receive any points, that is, zero (0) points. If participants showed they were attempting to employ the passive voice, they were also awarded; that is, if participants applied the incorrect form of the verb “to be” (e.g., “are” was used instead of “is”), they were awarded .5 point as they had attempted to use the verb “to be”. If they used the wrong form of the past participle, (e.g., “spreaded” instead of “spread”), they were awarded .5 point (See Tables 4.2, 4.3 and 4.4).

Correct sentences relating to Tables 4.2, 4.3 and 4.4:

Pods ***are collected*** from cacao tree.

Pods ***are grown*** in South America.

Pods ***are spread*** in sunshine.

*Table 4.2 Regular Past Participle*

Error	Student Answers	Mark
-----	Pods <b>are collected</b>	1+1
No <i>be</i>	Pods <b>collected</b>	0+1
Wrong form of <i>be</i>	Pods <b>is collected</b>	.5+1
No past participle	Pods <b>are collect</b>	1+0
No <i>be</i> / no past participle	Pods <b>collect</b>	0+0

*Table 4.3 Irregular Past Participle*

Type of Error	Student Answers	Mark
-----	Pods <b>are grown</b>	1+1
No <i>be</i>	Pods <b>grown</b>	0+1
Wrong form of <i>be</i>	Pods <b>is grown</b>	.5+1
Wrong form of past participle	Pods <b>are growed</b>	1+.5
No past participle	Pods <b>are grow</b>	1+0
No <i>be</i> / no past participle	Pods <b>grow</b>	0+0

*Table 4.4 Irregular Past Participle – No Change in Form*

Type of error	Student Answer	Mark
-----	Pods <b>are spread</b>	1+1
Wrong form of past participle	Pods <b>are spreaded</b>	1+.5
No <i>be</i>	Pods <b>spread</b>	0+0
Wrong form of <i>be</i>	Pods <b>is spread</b>	.5+1

After scoring all the writing tasks, the following formula was utilized to calculate a total percentage score for each student:

$$\frac{\text{Number of points scored}}{\text{Number of points possible (i.e., number of passive uses x 2)}} \times 100$$

## **4.9.2 Statistical analyses**

### **4.9.2.1 Statistical analyses for writing tasks**

SPSS version 22 was employed to analyse the statistical data for the writing tasks. To avoid misleading findings resulting from the effect of extreme values, outliers in each test were identified and excluded prior to conducting data analysis. Outliers are data points or scores that are distinct from other data scores in a set of data. Thus, to identify outliers, raw scores were converted to Z scores, and any score 3 standard deviation units below or above the mean was considered an outlier. Z scores identified the distance of a particular score from the mean (Pallant, 2013).

The participants in the study totalled 135; however, only 100 participants completed all the testing sections. Thirty-five participants were excluded due to not attending one or more of the writing tasks. The data collected from the writing tasks of 100 participants were subsequently calculated in SPSS.

In order to address RQ1, 2 and 3, that is, to determine to what extent written CF contributes to learners employing the passive voice more accurately across different types of written CF (between groups) and over time (within a group), firstly descriptive statistics were calculated. Then, a series of One-way ANOVAs was conducted for the groups pre-test scores. One-way ANOVAs are utilized to compare the variance between different groups and are utilized when there is one independent variable (e.g., feedback) with three or more

levels (e.g., direct, metalinguistic, control) (Pallant, 2001). One-way ANOVAs were also run to locate the source of significance.

The initial one-way ANOVAs detected significant effects in pre-test for RQ1 and RQ2. The significant effect of the pre-test could be a concern because if there were significant differences between groups in the pre-test, group differences in the immediate and delayed post-tests could be partially due to differences in the pre-test and not due to treatment. In order to control this concern, one-way ANOVA with gain scores was conducted. In experimental studies, as covariate adjustment for a pre-test can cause biased results, it is recommended to employ gain scores because they are an unbiased estimate of true change (Rogosa, 1988). Thus, gain scores were performed to investigate learners' improvement from the pre-test to the post-tests for whole groups.

A post hoc Bonferroni adjustment test was also utilized to investigate differences between groups. The reason for using Bonferroni post-hoc tests is that they are considered the most conservative, in other words, if they show a significant difference, the difference is certain (Pallant, 2013).

Within-group ANOVAs were chosen to see whether a group had improved over time. Within-group ANOVA is used when the same participants are measured over time (Pallant, 2013). Therefore, repeated measure ANOVAs with pre- and post-test scores were conducted. Accordingly, firstly, *F* tests were conducted followed by the use of Wilks' lambda in multivariate analysis of variance to explore whether there were differences between the mean of each group over time.

In pairwise comparisons, both *p* value and effect size were considered. The *p* value was employed to evaluate the significance of difference in mean scores. To ensure similarity with other research in the area, a confidence level of .05 ( $p < .05$ ) was used during analyses. However, *p* value is sensitive to sample size, and does not show how strongly

the two variables are associated. It is also susceptible (prone) to Type II error. To overcome the limitation, effect size in the form of Cohen's  $d$  and partial eta squared ( $\eta^2$ ) were considered. Cohen's  $d$  is based on mean differences and was used to show the size of difference between two groups. Partial eta squared ( $\eta^2$ ) was also utilized to show how strong the effect found in the ANOVAs was. Thus, Cohen's  $d$  was used for pairwise comparisons and partial eta squared ( $\eta^2$ ) was employed for ANOVAs. Table 3. 4 shows Cohen's benchmarks that were employed to indicate the value of partial eta squared ( $\eta^2$ ) and Cohen's  $d$  in the study (Pallant, 2013).

*Table 4.5 Cohen's Guideline to Interpret Partial Eta Squared and Cohen's  $d$*

<b>Size</b>	<b>Partial Eta Squared</b>	<b>Cohen's <math>d</math></b>
<b>Small</b>	.01 or 1%	.2
<b>Medium</b>	.06 or 6%	.5
<b>Large</b>	.138 or 13.8%	.8

In order to answer RQ3, the same procedure explained for RQ1 and RQ2 was employed. Although for RQ3 there was no significant difference in the pre-test, to be consistent with RQ1 and 2, one- way ANOVAs with gain scores were used.

#### **4.9.2.2 Statistical analyses for reading-span test**

To address RQ4 (the extent to which working memory and phonological short term memory mediate the effectiveness of focused direct CF and metalinguistic explanation with and without revision in new texts over time), a reading-span test was used to measure working memory. The participants' reading-span raw scores were calculated. The scores for the reading-span test included the following: (a) the number of sentence-final words correctly recalled; (b) the number of correctly judged sentences; and (c) mean reaction times for the correctly judged sentences. In order to weight tests equally and identify

outliers, all the raw scores calculated in (a), (b) and (c) were transformed into z-scores. Because higher reaction times represent slower responses, the reaction time z-scores were multiplied by (-1). Accordingly, a higher score reflected better performance for all three components (reaction times, sentence judgments, word recall). Then, an average score of (a), (b) and (c) was calculated as a composite working memory score for each participant. This included the correlations and regressions for the working memory and each feedback type. Pearson correlation coefficients ( $r$ ) – simple bivariate correlation – was conducted to measure the correlation between working memory and the post-tests for each type of feedback. Pearson correlation coefficients( $r$ ) give a value between -1 to +1 where 1 is total positive correlation, 0 is no correlation, and -1 is total negative correlation. The size of the value (ignoring the sign) indicates the strength of the relationship.

In order to reveal the unique contribution of working memory and phonological short term memory to the efficacy of different types of written CF, the scores of working memory, phonological short term memory and feedback types were subjected to standard multiple regression analyses. Multiple regression is used to explain the interrelationship among a set of variables, that is, a dependent variable and one or more independent variables (or 'predictors'). Multiple regression analysis was conducted so that feedback type was considered as a dependent variable and working memory as an independent variable. The regression was computed to identify to what extent working memory may predict the efficacy of different types of feedback. Accordingly, standardized coefficients, the model summary ( $R$  squared), and the ANOVA were assessed. The standardized coefficient represents the number of standard deviation changes in the outcome variable as a result of one standard deviation change in the predictor variable.  $R^2$  refers to the percentage of variance in the dependent variable which is explained by the independent variable (e.g.,  $R^2 = .43$  means 43% of the variance in the dependent variable was accounted for).

#### **4.9.2.3 Statistical analyses for non-word-span test**

A non-word-span test in the form of repetition was used to measure phonological short-term memory (RQ4). Accordingly, one point was given for each correctly recalled item in the correct sequence, and then the total of all correctly recalled items was computed. To remove outliers, the total raw scores were transformed into z scores for each participant. Simple bivariate correlation was conducted to measure the correlation between phonological short-term memory and the post-tests for each types of feedback. Standard multiple regression was computed so that feedback type was considered as a dependent variable and phonological short-term memory as an independent variable. Multiple regression analysis was conducted to identify to what extent phonological short-term memory may predict the efficacy of different types of feedback.

In order to investigate RQ4, one analysis in the form of regression was conducted. Additionally, post-test scores rather than gain scores were used with correlation and regression. The reason was that it has been shown that cognitive variables, such as phonological short-term memory, are traits not achievements (Carroll, 1981). Furthermore, the correlation of working memory and phonological short-term memory was investigated in terms of each group's scores so differences between groups were not relevant.

#### **4.10 Validity and reliability**

The validity and reliability of the main study were established in several ways. Validity means how well the study measures what it is supposed to measure. Internal validity is an important area of concern (Mackey & Gass, 2005). Mackey and Gass (2005) define internal validity as “the extent to which the results of a study are a function of the factor that the researcher intends” (p.109). The writing tasks, in this thesis, were designed to



elicit the target structure, that is, English passive voice. Even though these participants had substantial metalinguistic knowledge of English grammar (including the English passive voice), they had a limited ability to produce English in writing tasks. Thus, the students' improved accuracy on the English passive voice likely indicates the effectiveness of written CF, which establishes internal validity.

Furthermore, the validity of the reading span test and non-word span test as a measure of working memory and phonological short-term memory respectively were examined and confirmed by previous research (e.g., Danman & Carpenter, 1996; Deneman & Merikle, 1996; Juffs & Harrington, 2011).

The reliability of the main study was measured as well. Reliability refers to consistency; that is, it is the degree to which a test produces consistent and stable results. Reliability includes rater and instrument reliability (Mackey & Gass, 2005). The reliability of the scoring procedure was determined by calculating an *inter-rater reliability*. Inter-rater reliability is a measure of reliability employed to examine the degree to which different raters come to agreement about their assessment decisions (Bachman, 1995). Accordingly, all the texts were checked by a lecturer who has been teaching English for eight years in a university in Iran. This revealed an overall agreement of 96% for the writing tasks, 98% for the reading span test and 97% for the non-word test. Additionally, the instrument reliability was measured by test-retest in the pilot study.

#### **4.11 Ethical Approval**

Participants for this study followed a protocol that considers the rights of participants. They were invited to voluntarily participate in the study. To motivate students to participate, they were provided with general information about the study and how their participation could improve the field of SLA theoretically, empirically, and pedagogically. Participants were given consent forms that included a tick box to check if

they would like to see an analysis of the data. Participants were told that pseudonyms were used when analyzing and reporting the data (Mackey & Gass, 2005). Because formal approval from the University Ethics Committee was required prior to conducting the research, ethical approval was obtained from the AUT Ethics Committee (15/184).

## CHAPTER 5

### RESULTS

#### 5.1 Introduction

The results for RQ1, RQ2, RQ3 and RQ4 are presented in this chapter. RQ1 investigated the effect of written direct CF and metalinguistic explanation with and without revision on learners' immediate and delayed output in relation to the passive voice as the target structure in new writing tasks over time; RQ2 explored whether the opportunity for revision influenced the effectiveness of direct CF and metalinguistic explanation; RQ3 examined the relative efficacy of metalinguistic explanation and direct CF regardless of whether there was an opportunity for revision; and RQ4 investigated the extent to which working memory and phonological short term memory mediated efficacy in the use of direct CF and metalinguistic explanation with and without revision in new writing texts over time.

As explained in the Methodology chapter, to address RQ1, RQ2 and RQ3 a series of one-way ANOVAs was employed to test the comparative effects of the treatments between groups on the writing task. Repeated measure ANOVAs with pre- and post-test scores were conducted to measure within-group improvement of the experimental feedback groups and control group over time. In pairwise comparisons, both *p* value and effect size in the form of Cohen's *d* were used to evaluate the significance of the difference in mean scores. To address RQ4, firstly, Pearson correlation coefficients (*r*) were conducted to measure the correlation between working memory and phonological short term memory and the post-tests for each type of feedback. This was followed by regression to identify to what extent working memory and phonological short term memory can predict the

efficacy of different types of feedback. In the following sections, firstly, the findings for RQ1 will be presented followed by the findings for RQ2, 3 and 4.

**5.2. RQ1: What effect does focused DCF (direct CF) and ME (metalinguistic explanation) with and without revision have on learners' use of the English passive voice in an immediate text revision and in new texts over time?**

In order to address RQ1, that is, the efficacy of direct CF and metalinguistic explanation with and without revision on learners' use of the English passive voice in an immediate text revision and in new texts over time, firstly, descriptive statistics with raw scores for the accuracy scores of the English passive voice in the written tasks were calculated. Table 5.1 shows the descriptive statistics for the four treatment groups at the three different testing periods – Time 1 (pre-test), Time 2 (immediate post-test) and Time 3 (delayed post-test). The mean scores refer to the mean percentage accuracy in partial obligatory occasions. Figure 5.1 presents a visual illustration of the group mean of feedback types over time.

*Table 5.1 Descriptive statistics for the accuracy scores in the written tasks.*

Group	N	Time 1		Time 2		Time3	
		M	SD	M	SD	M	SD
<b>DCF</b>	20	29.63	23.44	78.80	19.62	70.69	25.66
<b>ME</b>	20	45.83	28.01	79.86	16.93	80.21	18.70
<b>DC+R</b>	20	51.45	24.79	85.50	11.62	84.04	13.05
<b>ME+R</b>	19	55.20	26.65	85.55	12.07	84.50	11.31
<b>CN</b>	21	54.40	31.92	58.70	29.11	41.07	33.21

*Note: DCF = direct corrective feedback with no revision, ME = metalinguistic explanation with no revision, DCF+R = direct corrective feedback with revision, ME+R = metalinguistic explanation with revision, CN = control group.*

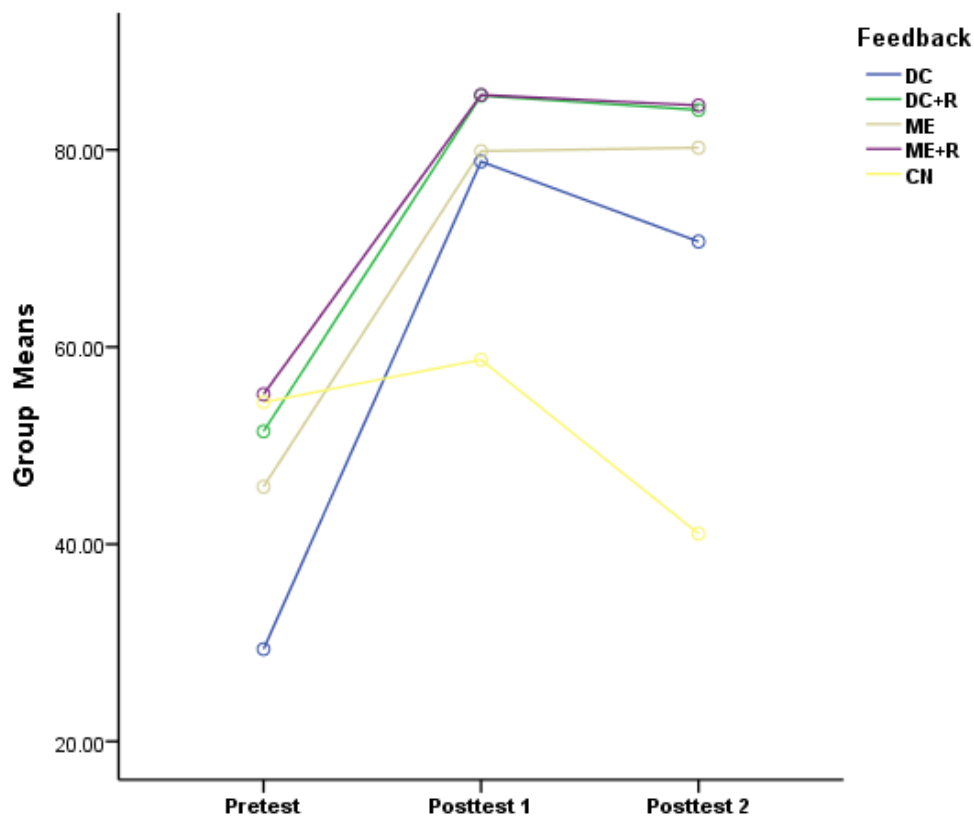


Figure 5.1. Group means of different types of feedback over time

Note: DCF = direct corrective feedback with no revision, ME = metalinguistic explanation with no revision, DCF+R = direct corrective feedback with revision, ME+R = metalinguistic explanation with revision, CN = control group.

Both Table 5.1 and Figure 5.1 show that from the pre-test to the immediate post-test, the accuracy of all four experimental groups increased considerably; however, from the immediate post-test to the delayed post-test the changes in the experimental groups were slight, that is, accuracy in the metalinguistic explanation group slightly increased while accuracy in the other groups slightly decreased. Accuracy in the control group also increased slightly from the pre-test to immediate post-test but it then decreased sharply in the delayed post-test.

Overall, the findings showed that (a) experimental groups performed better in the post-tests than in the pre-test; (b) the experimental groups showed higher scores than the control group on both post-tests; (c) in the pre-test, the mean scores for direct CF ( $M=29.63$ ,  $SD=23.44$ ) were low in comparison to other groups –direct CF plus revision ( $M=51.44$ ,  $SD=24.79$ ), metalinguistic explanation ( $M=45.83$ ,  $SD=28.01$ ), metalinguistic explanation plus revision ( $M=55.20$ ,  $SD=26.65$ ) and control group ( $M=54.40$ ,  $SD=31.92$ ) (See Table 5.1). This could be a concern because if there were significant differences between groups in the pre-test, group differences in Times 2 and 3 could be partially due to differences in Time 1 and not due to treatment. Thus, a one-way between groups ANOVA was employed to compare groups at Time 1.

In the pre-test (Time 1), the results of a one-way ANOVA showed significant differences among the pre-test scores of the five groups:  $F(4, 95) = 3.067$ ,  $P = .02$ . The learners' pre-test scores were subjected to post hoc pairwise comparisons to investigate the comparative effects of feedback types on the passive voice structure. Post-hoc comparison using Bonferroni adjustments indicated that direct CF was significantly lower than metalinguistic plus revision and the control group. Thus, the finding at Time 1 indicated that the possibility that group differences in Times 2 and 3 could be partially due to differences in Time 1 instead of the instructional treatments cannot be precluded. As explained in section 4.9.2.1 of the methodology chapter, in order to control this concern, raw scores were changed to gain scores. Table 5.2 shows the descriptive statistics for the four treatment groups for the immediate gain (gain 1) and the delayed gain (gain 2). Gain 1 was calculated by subtracting immediate post-test scores from pre-test scores, and gain 2 was measured by subtracting delayed post-test scores from pre-test scores. Figure 5.2 presents a visual illustration of gain scores 1 and 2 for different types of feedback.

Table 5.2 Descriptive Statistics of Gain Scores

Groups	N	Gain 1		Gain 2	
		M	SD	M	SD
DC	20	49.44	22.41	41.33	29.70
ME	20	34.03	28.63	34.38	23.94
DC+R	20	34.05	25.91	32.59	24.97
ME+R	19	30.35	23.19	29.30	22.25
Control	21	4.30	30.38	-13.32	32.77

Note: DCF = direct corrective feedback with no revision, ME = metalinguistic explanation with no revision, DCF+R = direct corrective feedback with revision, ME+R = metalinguistic explanation with revision, CN = control group.

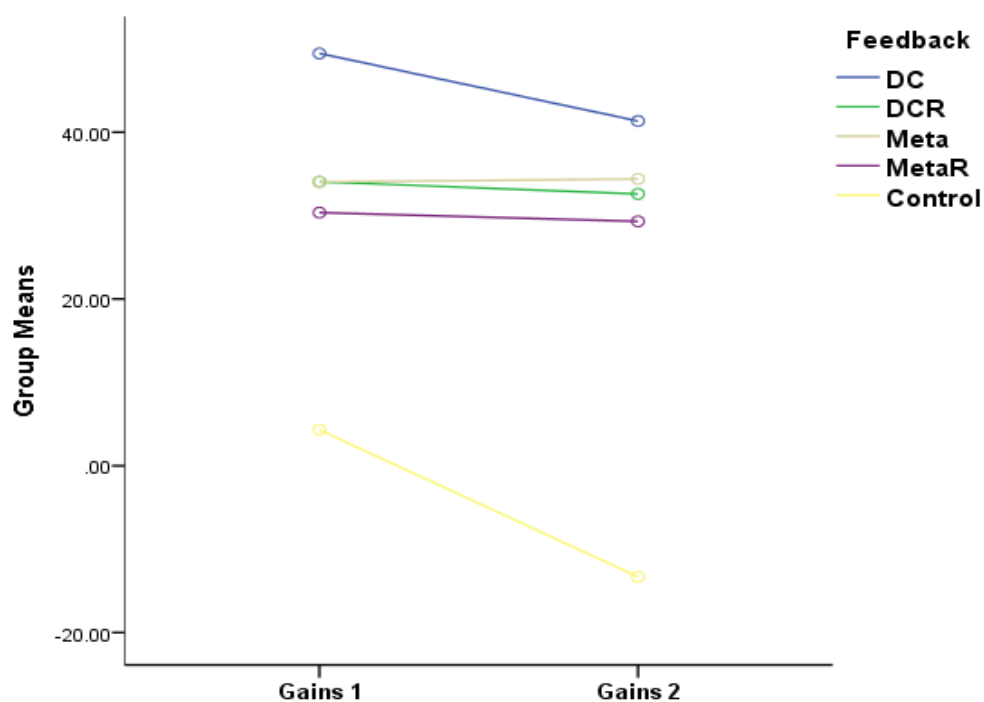


Figure 5.2. Group means of different types of feedback with gain scores

In order to investigate the comparative effects of the treatments (between groups) on writing task scores, one-way ANOVAs and pairwise comparisons were conducted with gain scores. The results revealed that for gain 1, the result of a one-way ANOVA showed significant differences between groups:  $F(4, 95) = 7.94, p < .001, \eta^2 = .25$ .

The findings reported in Table 5.3 show the resulting pairwise comparisons and Cohen's  $d$  values for the experimental and control groups. For gain 1, the findings indicated that the experimental groups significantly outperformed the CN group. Cohen's  $d$  values for direct CF, direct CF plus revision, metalinguistic explanation and metalinguistic explanation plus revision versus CN were 1.69, 0.97, 1.00, and 0.96 respectively which in Cohen's terms would be considered large effect sizes.

*Table 5.3. Effect Sizes in the Form of Cohen's  $d$  for the Accuracy Scores between Groups*

Group Contrast	Gains 1		Gains 2	
	$d^1$	$p^2$	$d$	$p$
<b>DC vs. ME</b>	0.59	0.67	0.25	1.00
<b>DC vs. DC+R</b>	0.63	0.68	0.31	1.00
<b>DC vs. ME+R</b>	0.83	0.26	0.45	1.00
<b>ME vs. ME+R</b>	0.14	1.00	0.21	1.00
<b>ME vs. DC+R</b>	0.00	1.00	0.07	1.00
<b>ME+R vs. DC+R</b>	0.15	1.00	0.13	1.00
<b>DC vs. CN</b>	1.69	.00	1.74	.00
<b>ME vs. CN</b>	1.00	.005	1.66	.00
<b>DC+R vs. CN</b>	0.97	.005	1.57	.00
<b>ME+R vs. CN</b>	0.96	.02	1.52	.00

*Note 1: Effect size (Cohen's  $d$ ). 2: Results of null hypothesis significance testing. ME = metalinguistic explanation with no revision, DC= direct corrective feedback with no revision, ME+R = metalinguistic explanation with revision, DC+R = direct corrective feedback with revision, CN = control group.*



Although all the experimental groups had large effect sizes, the value of direct CF was higher than other groups. The analyses failed to detect significant differences between the experimental groups. Cohen's  $d$  values for all experimental groups versus each other ranged from small to large, that is, 0.00 to 0.83.

For gain 2, the finding of a one-way ANOVA showed significant differences between groups:  $F(4, 95) = 13.39, p < .001, \eta^2 = 0.36$ . The findings of pairwise comparisons and Cohen's  $d$  values for gain 2 are shown in Table 5.3. The results indicated that the experimental groups performed significantly better than the control group. Cohen's  $d$  values for direct CF, direct CF plus revision, metalinguistic explanation and metalinguistic plus revision versus the control group were 1.74, 1.57, 1.66, and 1.52 respectively which in Cohen's terms would be considered large effect sizes. The value of direct CF was the largest. There were no significant differences between experimental groups with small effect sizes. The values ranged from 0.07 to 0.45.

Within-groups ANOVA is used when the same participants are measured under different conditions (or measured at different points in time) (Pallant, 2013). Thus, within-groups ANOVA was chosen to compare scores on writing of the experimental feedback groups at Time 1 (pre-test), Time 2 (immediate post-test) and Time 3 (delayed post-test). To do so, repeated measure ANOVAs with pre- and post-test scores were conducted. The means and standard deviations for groups are presented in Table 5.1. The findings of repeated measures ANOVA showed that experimental and control groups had a significant effect for time:

$F(2, 18) = 65.90, p < .001, \eta^2 = .58$ . Additionally, the results of repeated measures ANOVAs for treatment and control groups revealed significant differences for each group, that is, direct CF group:  $F(2, 18) = 46.25, p < .001, \eta^2 = .83$ ; direct CF plus revision group:  $F(2, 18) = 17.33, p < .001, \eta^2 = .65$ ; metalinguistic explanation group:  $F(2, 18) =$

20.55,  $p < .001$ ,  $\eta^2 = .69$ ; metalinguistic explanation plus revision group:  $F(2, 17) = 15.73$ ,  $p < .001$ ,  $\eta^2 = .64$  and control group:  $F(2, 19) = 5.62$ ,  $p = .01$ ,  $\eta^2 = .37$ .

The findings of pairwise comparisons (paired t-test) and the within-group effect sizes for experimental and control groups are reported in Table 5. 4.

*Table 5.4. Effect Sizes in the Form of Cohen's d for the Accuracy Scores Overtime*

Group	Time 1–Time 2		Time 2–Time 3		Time 1–Time3	
	<i>N</i>	<i>d</i> <sup>1</sup>	<i>p</i> <sup>2</sup>	<i>d</i>	<i>p</i>	<i>d</i>
<b>DCF</b>	20	2.28	.00	.35	.35	1.68
<b>ME</b>	20	1.47	.00	0.01	1.00	1.44
<b>DC+R</b>	20	1.75	.00	0.11	1.00	1.64
<b>ME+R</b>	19	1.46	.00	0.08	1.00	1.43
<b>CN</b>	21	0.14	1.00	0.56	.008	0.40

*Note. 1 Effect size (Cohen's d). 2 Results of null hypothesis significance testing. ME = metalinguistic explanation with no revision, DC = direct corrective feedback with no revision, ME+R = metalinguistic explanation with revision, DC+R = direct corrective feedback with revision, CN = control group*

The findings revealed that all the experimental groups significantly improved from Time 1 to Time 2. Cohen's d values for direct CF, direct CF plus revision, metalinguistic explanation and metalinguistic explanation plus revision from Time 1 to Time 2 were 2.28, 1.75, 1.47, and 1.46 respectively, which in Cohen's terms would be considered large effect sizes. The value of direct CF was the largest. There were no significant differences for all groups from Time 2 to Time 3 as small effect sizes were found, with values ranging from 0.01 to 0.35. That is, the experimental groups were able to sustain this improvement. However, the mean scores at Time 3 were still significantly greater than those of Time 1 with large effect sizes. Cohen's d values for direct CF, direct CF plus revision, metalinguistic explanation and metalinguistic explanation plus revision from Time 1 to

Time 3 were 1.68, 1.64, 1.44, and 1.43 respectively, which in Cohen's terms would be considered large effect sizes. The direct CF group had the biggest value. The control group showed no significant differences from Time 1 to Time 2 with small effect size (i.e., 0.14), but the scores significantly decreased from Time 2 to Time 3 with medium effect size (i.e., 0.56). There was no significant difference from Time 1 to Time 3 with small effect size (i.e., 0.40). In summary, all experimental groups were effective over time. However, the efficacy of the direct CF group proved to be relatively more effective in both short and long terms than other groups.

### **5.3. RQ2: *Does the opportunity for the revision influence the efficacy of DC and ME?***

In order to address RQ2, that is, investigating whether requiring the participants to do revision had any efficacy on accuracy in subsequent pieces of writing, the two revision groups (i.e., metalinguistic explanation plus revision and direct CF plus revision) and the two groups that did not make revisions (i.e., metalinguistic explanation and direct CF) were combined and compared. DMPR group stands for the combined direct CF plus revision and metalinguistic explanation plus revision groups and DMWR group stands for the combined direct CF and metalinguistic explanation groups. Table 5. 5 shows the descriptive statistics for the treatment groups at the three different testing periods, while Figure 5.3. displays graphically the group means of feedback types over time.

Both Table 5.5 and Figure 5.3 show that from the pre-test (Time 1) to the immediate post-test (Time 2), the accuracy of two experimental groups increased sharply; however, from the immediate post-test to the delayed post-test (Time 3) the changes in the experimental groups were slight, that is, accuracy in both DMWR (direct CF and metalinguistic explanation) and DMPR (direct CF plus revision and metalinguistic explanation plus revision) groups slightly decreased. Accuracy in the control group also

increased slightly from the pre-test to immediate post-test but it then decreased sharply in the delayed post-test.

Table 5.5 Descriptive statistics for the accuracy scores in the written tasks

Groups		Time 1		Time 2		Time 3	
	N	M	SD	M	SD	M	SD
DMWR	40	37.59	26.82	79.33	18.09	75.45	22.68
DMPR	39	53.27	25.44	85.52	11.68	84.26	12.07
CN	21	54.40	31.92	58.70	29.11	41.07	33.21

Note: DMWR group stands for the combined DC and ME groups, DMPR group stands for the combined direct CF plus revision and metalinguistic explanation plus revision groups and CN stands for the control group.

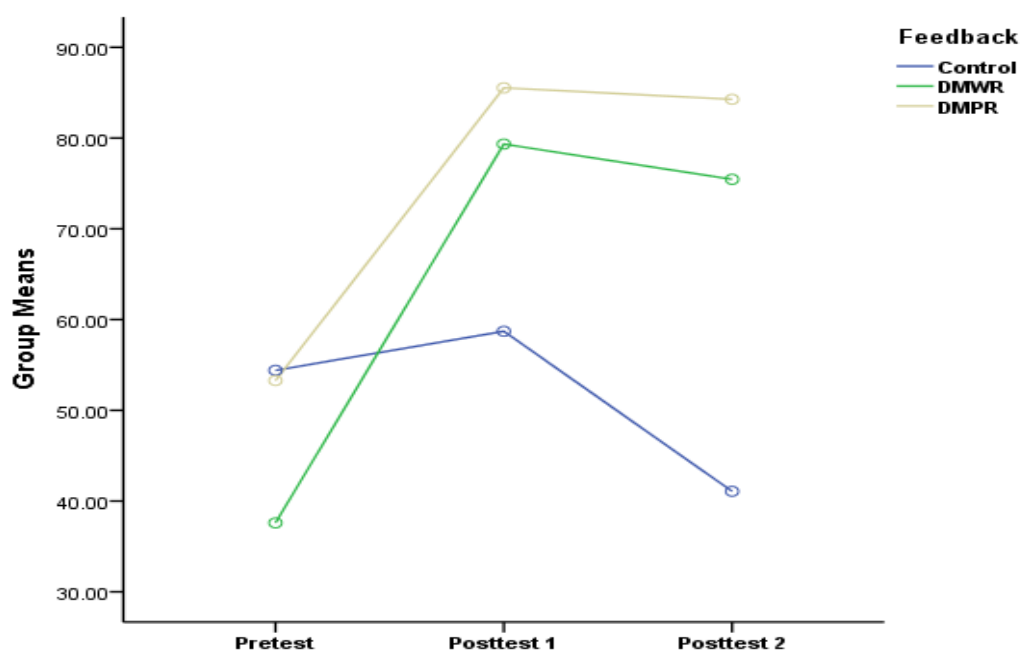


Figure 5.3. Group means of DMWR (direct CF and metalinguistic explanation) and DMPR (direct CF plus revision and metalinguistic explanation plus revision) feedback over time

Overall, the findings showed that (a) the experimental groups performed better in the post-tests than in the pre-test; (b) the experimental groups showed higher scores than the

control group on both post-tests; (c) in the pre-test, the mean scores for DMWR (direct CF and metalinguistic explanation) ( $M=37.59$ ,  $SD=26.82$ ) were low in comparison to DMWR group ( direct CF plus revision and metalinguistic explanation plus revision) ( $M=53.27$ ,  $SD=25.44$ ) and control group ( $M=54.40$ ,  $SD=31.92$ ) (See Table 5.5). This could be a concern because if there were significant differences between groups in the pre-test, group differences in Times 2 and 3 could be partially due to differences in Time 1 and not due to treatment. Thus, a one-way between groups ANOVA was employed to compare groups at Time 1.

The results of one-way between groups ANOVA revealed significant between group differences at Time 1 (pre-test):  $F(2, 97) = 4.128$ ,  $P=.019$ ,  $\eta^2 = 0.07$ . 4. As explained in section 3.10.1.2 of the methodology chapter, to control this concern, raw scores were changed to gain scores. Table 5.6 shows the descriptive statistics for the treatment and control groups for the immediate gain (gain 1) and the delayed gain (gain 2). Figure 5.4 presents a visual illustration of gain scores 1 and 2 for the treatment and control groups. Then one-way ANOVAs with gain scores 1 and 2 were conducted.

Table 5.6 *Descriptive statistics of the gain scores for the accuracy scores in the written tasks*

Groups		Gain 1		Gain 2	
	N	M	SD	M	SD
DMWR	40	41.74	26.55	37.85	26.86
DMWR	39	32.25	24.37	30.99	23.43
CN	21	4.30	30.38	-13.32	32.77

*Note: DMWR group stands for the combined direct CF and metalinguistic explanation groups, DMWR group stands for the combined direct CF plus revision and metalinguistic explanation plus revision groups and CN stands for the control group.*

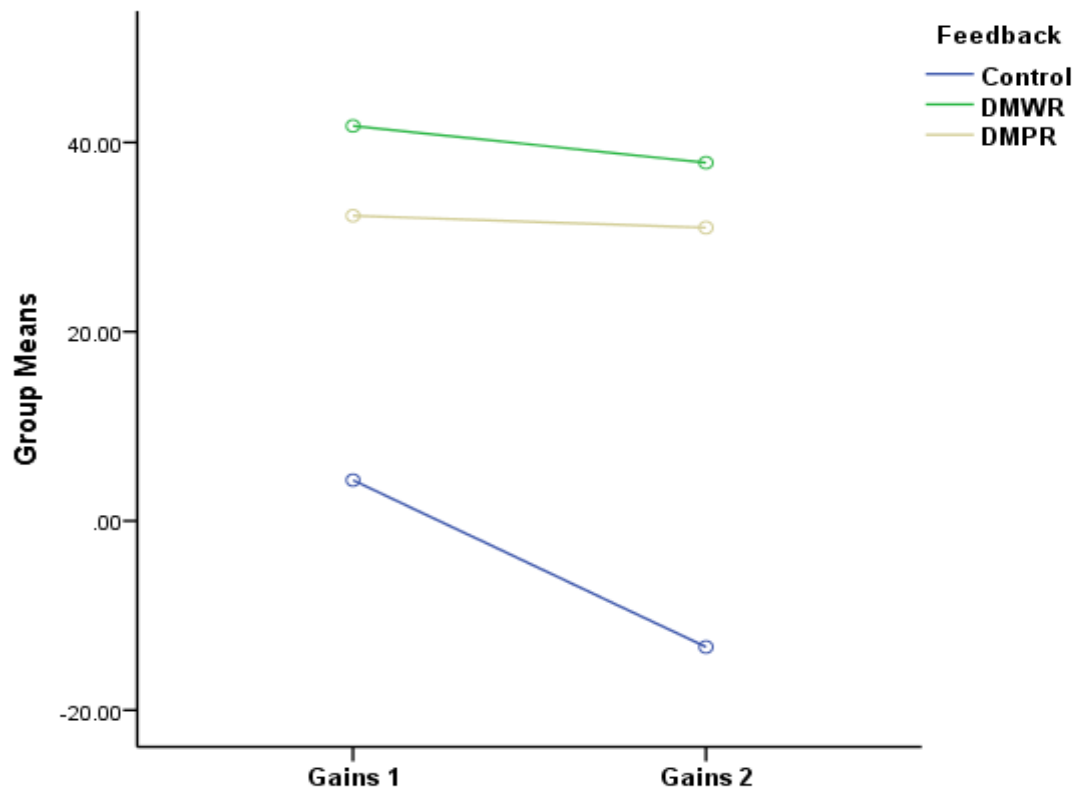


Figure 5.4. Group means of DMWR (direct CF and metalinguistic explanation) and DMPR (direct CF plus revision and metalinguistic explanation plus revision) feedback with gain scores

For gain 1, the results of a one-way ANOVA showed significant differences between groups:  $F(2, 97) = 13.84, p < .001, \eta^2 = 0.22$ . Table 5.7 shows the findings of pairwise comparisons and Cohen's  $d$  values for gain 1. The results revealed that the DMWR (direct CF and metalinguistic explanation) and DMPR (direct CF plus revision and metalinguistic explanation plus revision) groups performed significantly better than the control group. Cohen's  $d$  values for DMWR (direct CF and metalinguistic explanation) and DMPR (direct CF plus revision and metalinguistic explanation plus revision) versus control group were 1.31 and 1.01 respectively which in Cohen's terms would be considered large effect sizes. However, DMWR (direct CF and metalinguistic explanation) had a higher value than DMPR (direct CF plus revision and metalinguistic explanation plus revision). The analyses failed to detect any significant differences between DMWR (direct CF and metalinguistic explanation) and DMPR (direct CF plus revision and metalinguistic

explanation plus revision). Cohen's  $d$  value for DMWR (direct CF and metalinguistic explanation) versus DMPR (direct CF plus revision and metalinguistic explanation plus revision) detected a small effect size, that is, 0.37.

Table 5.7: *Effect Sizes in the Form of Cohen's  $d$  for the Accuracy Scores Between Groups*

Group Contrast	Gain 1		Gain 2	
	$d^1$	$p^2$	$d$	$p$
<b>DMWR vs. DMPR</b>	0.37	.34	0.27	.78
<b>DMWR vs. CN</b>	1.31	.00	1.70	.00
<b>DMPR vs. CN</b>	1.01	.001	1.55	.00

*Note. 1 Effect size (Cohen's  $d$ ). 2 Results of null hypothesis significance testing. DMWR group stands for the combined direct CF and metalinguistic explanation groups, DMPR group stands for the combined direct CF plus revision and metalinguistic plus revision groups and CN stands for the control group.*

For gain 2, the findings of a one-way ANOVA showed significant differences between groups:  $F(2, 97) = 26.71, p < .001, \eta^2 = 0.35$ . The findings of pairwise comparisons and Cohen's  $d$  values for gain 2 are reported in Table 5.7. The findings revealed that the mean scores for DMWR (direct CF and metalinguistic explanation) and DMPR (direct CF plus revision and metalinguistic explanation plus revision) were significantly higher than control group. The resulting Cohen's  $d$  values for DMWR (direct CF and metalinguistic explanation) and DMPR (direct CF plus revision and metalinguistic explanation plus revision) versus control group were 1.70 and 1.55 respectively, which in Cohen's terms would be considered large effect sizes. However, the value of DMWR (direct CF and metalinguistic explanation) was higher than DMPR (direct CF plus revision and metalinguistic explanation plus revision) as opposed to control group. The analyses failed to detect any significant differences between DMWR (direct CF and metalinguistic explanation) and DMPR (direct CF plus revision and metalinguistic explanation plus revision).

revision). Cohen's  $d$  value for DMWR (direct CF and metalinguistic explanation) versus DMPR (direct CF plus revision and metalinguistic explanation plus revision) was 0.27 which was a small effect size.

In sum, for gains 1 and 2, DMWR (direct CF and metalinguistic explanation) and DMPR (direct CF plus revision and metalinguistic explanation plus revision) groups were significantly higher than control group with large effect sizes. However, the value of DMWR (direct CF and metalinguistic explanation) was higher than DMPR (direct CF plus revision and metalinguistic explanation plus revision) for both gain scores 1 and 2. In other words, DMWR was relatively more effective than DMPR.

In order to measure within-group effects, repeated measure ANOVAs with pre- and post-test scores were conducted to compare scores on writing of the DMWR (direct CF and metalinguistic explanation) versus DMPR (direct CF plus revision and metalinguistic explanation plus revision) groups at Time 1 (pre-test), Time 2 (immediate post-test) and Time 3 (delayed post-test). The means and standard deviations are presented for both groups in Table 5.5.

The findings of repeated measure ANOVAs showed that all groups had a significant effect for time:  $F(2, 96) = 45.01, p < .001, \eta^2 = .48$

The results of repeated measure ANOVAs for each group showed significant effect for time for the DMWR (direct CF and metalinguistic explanation) group:  $F(2, 38) = 49.94, p < .001, \eta^2 = .72$  and the DMPR (direct CF plus revision and metalinguistic explanation plus revision) group:  $F(2, 37) = 34.70, p < .001, \eta^2 = .65$ .

Table 5.8. *Effect Sizes in the Form of Cohen's  $d$  for the Accuracy Scores Overtime*

Group		Time 1–Time 2		Time 2–Time 3		Time 1–Time3	
	N	$d^1$	$p^2$	$d$	$p$	$d$	$p$
DMWR	40	1.82	.00	0.18	.53	1.52	.00



<b>DMPR</b>	39	1.62	.00	0.10	1.00	1.55	.00
<b>CN</b>	21	.14	1.00	0.56	.008	0.40	.23

*Note: DMWR group stands for the combined direct CF and metalinguistic explanation groups, DMPR group stands for the combined direct CF plus revision and metalinguistic explanation plus revision groups and CN stands for the control group.*

The results shown in Table 5.8 report the pairwise comparisons (paired t-test) and effect sizes for DMWR (direct CF and metalinguistic explanation), DMPR (direct CF plus revision and metalinguistic explanation plus revision) and control groups. The findings showed that the DMWR (direct CF and metalinguistic explanation) and DMPR (direct CF plus revision and metalinguistic explanation plus revision) groups significantly improved from Time 1 to Time 2 with large effect sizes, that is, 1.82 and 1.62 respectively. The value of the DMWR (direct CF and metalinguistic explanation) group was bigger than the other group from Time 1 to Time 2. For both the DMWR (direct CF and metalinguistic explanation) and DMPR (direct CF plus revision and metalinguistic explanation plus revision) groups, there were no significant differences from Time 2 to Time 3. The values of effect sizes for the DMWR (direct CF and metalinguistic explanation) and DMPR (direct CF plus revision and metalinguistic explanation plus revision) groups were small from Time 2 to Time 3, that is, 0.18 and 0.10 respectively. However, the scores at Time 3 were significantly greater than those of Time 1. The value of effect sizes for the DMWR (direct CF and metalinguistic explanation) and DMPR (direct CF plus revision and metalinguistic explanation plus revision) groups increased from Time 1 to Time 3 with large effect sizes, that is, 1.52 and 1.55 respectively. However, the DMPR (direct CF plus revision and metalinguistic explanation plus revision) group had the bigger value. The control group showed no significant differences from Time 1 to Time 2 with small effect sizes, but the scores significantly decreased from Time 2 to Time 3 with medium effect size. There was no significant differences from

Time 1 to Time 3 with small effect size. Overall, both the DMWR (direct CF and metalinguistic explanation) and the DMPR (direct CF plus revision and metalinguistic explanation plus revision) treatments were effective. However, DMWR was relatively more effective in the short term and DMPR proved relatively more effective in the long term.

**5.4. RQ3: *Is there any difference in the effect of DC and ME regardless of whether there is an opportunity for revision?***

In order to address RQ3, that is, to investigate the relative efficacy of direct CF and metalinguistic explanation regardless of whether there is an opportunity for revision, the two direct CF groups (i.e., direct CF and direct CF plus revision) and the two metalinguistic explanation groups (i.e., metalinguistic explanation and metalinguistic explanation plus revision) were combined and compared. DCO group stands for the combined direct CF groups (i.e., direct CF and direct CF plus revision) and MEO group stands for the combined metalinguistic explanation groups (i.e., metalinguistic explanation and metalinguistic explanation plus revision). Table 5.9 shows the descriptive statistics for the four treatment groups at the three different testing periods, that is, Time 1 (pre-test), Time 2 (immediate post-test) and Time 3 (delayed post-test). The mean scores refer to the mean percentage accuracy in partial obligatory occasions. The means of feedback types are plotted on the graph in Figure 5. 5.

*Table 5.9 Descriptive Statistics for the Accuracy Scores in the Written Tasks*

Groups		Time 1		Time 2		Time 3	
	N	M	SD	M	SD	M	SD
DCO	40	40.40	26.31	82.15	16.27	77.36	21.20
MEO	39	50.39	27.40	82.63	14.86	82.30	15.49
CN	21	54.40	31.92	58.70	29.11	41.07	33.21

*Note: DCO group stands for the combined direct CF groups (i.e., direct CF and direct CF plus revision), MEO group stands for the combined metalinguistic explanation groups (i.e., metalinguistic explanation and metalinguistic explanation plus revision) and CN stands for the control group.*

Both Table 5.9 and Figure 5.4 show that from Time 1 to Time 2, the accuracy of the two experimental groups increased sharply; however, from Time 2 to Time 3 the changes in the experimental groups were slight, that is, accuracy in both the DCO (direct CF and direct CF plus revision) and MEO (metalinguistic explanation and metalinguistic explanation plus revision) groups slightly decreased. Accuracy in the control group also increased slightly from Time 1 to Time 2 but it then decreased sharply in Time 3.

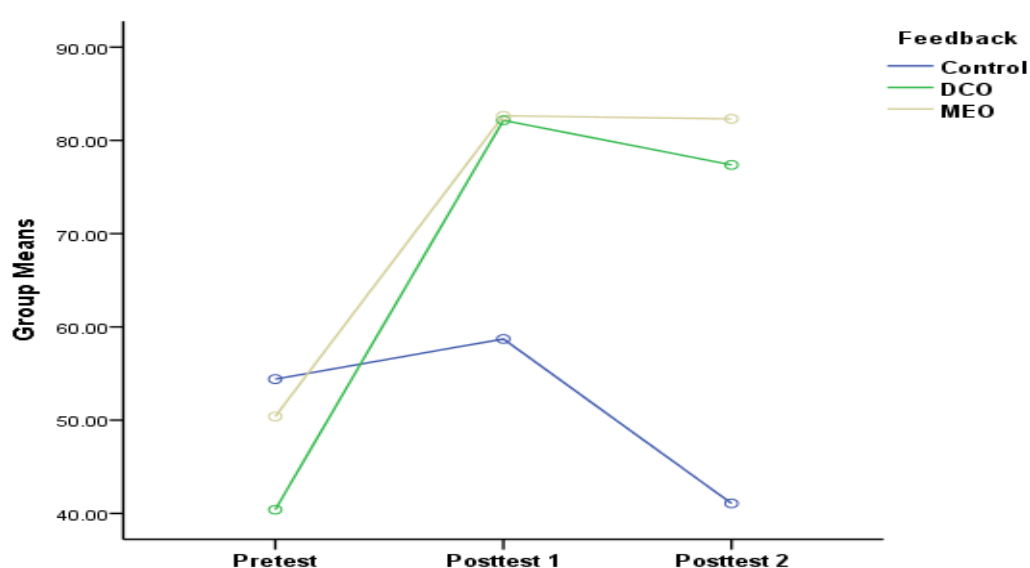


Figure 5.5. Group means of DCO (direct CF and direct CF plus revision) and MEO (i.e., metalinguistic explanation and metalinguistic explanation plus revision) feedback over time

One-way ANOVA was conducted to measure whether there was significant differences in the pre-test results. The result of one-way ANOVA showed that there were no significant differences between the three groups at Time 1 (see Table 5.9.):  $F(2, 97) = 2.13, P = .12, \eta^2 = 0.04$ . However, gain scores were employed in RQ1 and RQ2, and in order to be consistent, for RQ3, raw scores were also changed to gain scores to measure between groups. Table 5.10 shows the descriptive statistics for the treatment and control groups for the immediate gain (gain 1) and the delayed gain (gain 2). Figure 5.6 presents a visual illustration of gain scores 1 and 2 for the treatment and control groups.

Table 5.10 Descriptive Statistics for the Accuracy Scores in the Written Tasks

Groups	N	Gain 1		Gain 2	
		M	SD	M	SD
DCO	40	41.75	25.15	36.96	27.44
MEO	39	32.24	25.84	31.90	22.97
CN	21	4.30	30.38	-13.32	32.77

Note: DCO group stands for the combined direct CF groups (i.e., direct CF and direct CF plus revision), MEO group stands for the combined metalinguistic explanation groups (i.e., i.e., metalinguistic explanation and metalinguistic explanation plus revision) and CN stands for the control group.

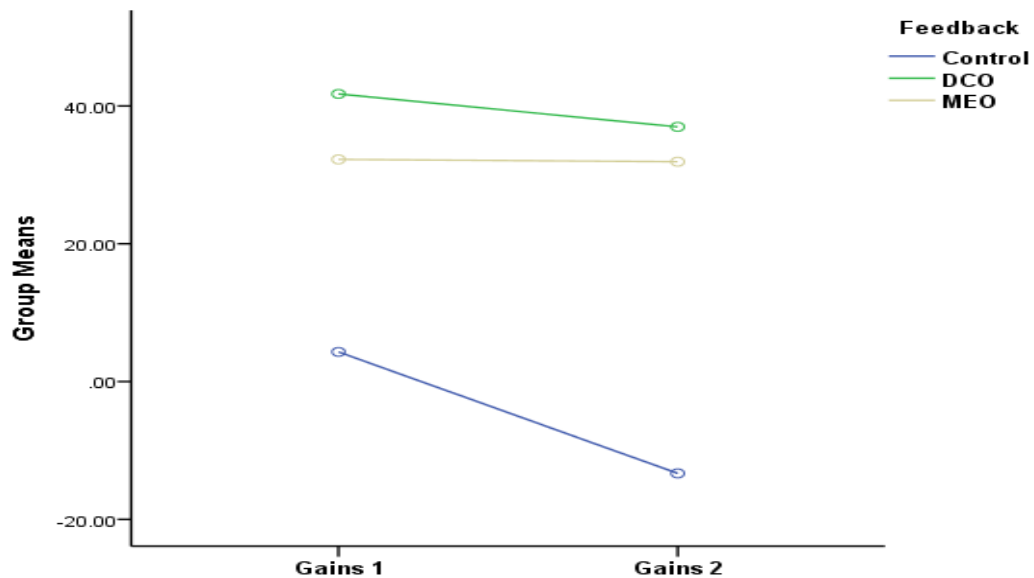


Figure 5.6. Group means of DCO (direct CF and direct CF plus revision) and MEO (i.e., metalinguistic explanation and metalinguistic explanation plus revision) feedback over time.

The findings of one-way ANOVAs for DCO (direct CF and direct CF plus revision), MEO (i.e., metalinguistic explanation and metalinguistic explanation plus revision) and control group for gains 1 and 2 are explained in the following: For gain 1, the result of a one-way ANOVA showed significant differences between three groups:  $F(2, 97) = 13.85$ ,  $p < .001$ ,  $\eta^2 = 0.22$ .

*Table: 5.11 Effect Sizes in the Form of Cohen's d for the Accuracy Scores Between Groups*

Group Contrast	Gain 1		Gain 2	
	$d^1$	$p^2$	$d$	$p$
DCO vs. MEO	0.37	.34	0.20	1.00
DCO vs. CN	1.34	.00	1.66	.00
MEO vs. CN	0.99	.001	1.59	.00

*Note. 1 Effect size (Cohen's d). 2 Results of null hypothesis significance testing. DCO group stands for the combined direct CF groups (i.e., direct CF and direct CF plus revision), MEO group stands for the combined metalinguistic explanation groups (i.e., metalinguistic explanation and metalinguistic explanation plus revision) and CN stands for the control group.*

Table 5.11 shows that for gain 1, both DCO (direct CF and direct CF plus revision) and MEO (metalinguistic explanation and metalinguistic explanation plus revision) scored significantly higher than control group. Cohen's d values for DCO (direct CF and direct CF plus revision) and MEO (metalinguistic explanation and metalinguistic explanation plus revision) versus control group were 1.34 and 0.99 respectively which in Cohen's terms would be considered large effect sizes. Although both had large effect sizes, DCO (direct CF and direct CF plus revision) had a higher value than MEO (metalinguistic explanation and metalinguistic explanation plus revision). The analyses failed to detect any significant differences between DCO (direct CF and direct CF plus revision) and MEO (metalinguistic explanation and metalinguistic explanation plus revision). Cohen's d value for DCO (direct CF and direct CF plus revision) versus MEO (metalinguistic explanation and metalinguistic explanation plus revision) was 0.37 which was a small effect size.

For gain 2, a one-way ANOVA revealed significant differences between groups:  $F(2, 97) = 26.26, p < .001, \eta^2 = 0.35$ . The results reported in Table 5.11 show that for gain 2, DCO (direct CF and direct CF plus revision) and MEO (metalinguistic explanation and

metalinguistic explanation plus revision) performed significantly higher than control group. The resulting Cohen's  $d$  values for DCO (direct CF and direct CF plus revision) and MEO (i.e., metalinguistic explanation and metalinguistic explanation plus revision) versus control group were 1.66 and 1.59 respectively which in Cohen's terms would be considered large effect sizes. However, the value of DCO (direct CF and direct CF plus revision) was higher than MEO (metalinguistic explanation and metalinguistic explanation plus revision) as opposed to control group. DCO (direct CF and direct CF plus revision) was not significantly different from MEO with small effect size.

To summarize, for gains 1 and 2, the means of the DCO (direct CF and direct CF plus revision) and MEO (metalinguistic explanation and metalinguistic explanation plus revision) groups were significantly higher than the control group with large effect sizes. However, the value of DCO was higher than MEO in both gains 1 and 2. In other words, DCO was relatively more effective than MEO.

In order to measure within-groups, repeated measure ANOVAs with pre- and post-test scores were conducted to compare scores on writing of the DCO (direct CF and direct CF plus revision) and MEO (metalinguistic explanation and metalinguistic explanation plus revision) feedback groups at Time 1, Time 2 and Time 3. The means and standard deviations are presented for both groups in Table 5.9.

The findings of repeated measures ANOVA showed that all groups had a significant effect for time:  $F(2, 96) = 45.15, p < .001, \eta^2 = .48$ . Additionally, the results of repeated measures ANOVAs for each group showed a significant effect for time for the DCO (direct CF and direct CF plus revision) group:  $F(2, 38) = 53.95, p < .001, \eta^2 = .74$  and MEO (metalinguistic explanation and metalinguistic explanation plus revision) group:  $F(2, 37) = 37.04, p < .001, \eta^2 = .66$ . The findings also showed that both DCO (direct CF and direct CF plus revision) and MEO (metalinguistic explanation and metalinguistic

explanation plus revision) groups had a significant effect for time, that is, DCO (direct CF and direct CF plus revision) Group:  $F(2, 38) = 53.95, p < .001, \eta^2 = .74$ ; MEO (metalinguistic explanation and metalinguistic explanation plus revision) Group:  $F(2, 37) = 37.04, p < .001, \eta^2 = .66$ .

The findings of pairwise comparisons and the within-group effect sizes for the DCO (direct CF and direct CF plus revision), the MEO (metalinguistic explanation and metalinguistic explanation plus revision) and control groups are reported in Table 5.12.

Table 5.12. *Effect Sizes in the Form of Cohen's d for the Accuracy Scores Overtime*

Group		Time 1–Time 2		Time 2–Time 3		Time 1–Time3	
	N	$d^1$	$p^2$	$d$	$p$	$d$	$p$
DCO	40	1.90	.00	0.25	.309	1.54	.00
MEO	39	1.46	.00	0.02	1.00	1.43	.00
CN	21	.14	1.0	0.56	.008	0.40	.23

Note. 1 *Effect size (Cohen's d)*. 2 *Results of null hypothesis significance testing*. DCO group stands for the combined direct CF groups (i.e., direct CF and direct CF plus revision), MEO group stands for the combined metalinguistic explanation groups (i.e., metalinguistic explanation and metalinguistic explanation plus revision) and CN stands for the control group

The results revealed that the DCO (direct CF and direct CF plus revision) and MEO (metalinguistic explanation and metalinguistic explanation plus revision) groups significantly improved from Time 1 to Time 2. The Cohen's d values for DCO (direct CF and direct CF plus revision) and MEO (metalinguistic explanation and metalinguistic explanation plus revision) increased from Time 1 to Time 2 with large effect sizes, that is, 1.90 and 1.46 respectively, while the value of the DCO (direct CF and direct CF plus revision) group was bigger than the other group over Time 1 to Time 2. For both the DCO (direct CF and direct CF plus revision) and MEO (metalinguistic explanation and metalinguistic explanation plus revision) groups, there were no significant differences from Time 2 to Time 3 with small effect sizes, that is, 0.25 and .02 respectively. However,

there were significant differences from Time 1 to Time 3. The value of effect sizes for the DCO (direct CF and direct CF plus revision) and MEO (metalinguistic explanation and metalinguistic explanation plus revision) groups increased from Time 1 to Time 3 with large effect sizes, that is, 1.54 and 1.43 respectively. The DCO (direct CF and direct CF plus revision) group had a bigger value than the MEO (metalinguistic explanation and metalinguistic explanation plus revision) group. The control group showed no significant differences from Time 1 to Time 2 with small effect sizes, but the scores significantly decreased from Time 1 to Time 2 with medium effect size. The control group showed small effect sizes in Time 1 to Time 2 (.14) and Time 1 to Time 3 (.40), and medium effect size in Time 2 to Time 3 (.56). To summarize, both DCO and MEO were effective over time; however, DCO was relatively more effective, especially in the short term.

***5.5 RQ 4: To what extent do working memory (WM) and phonological short-term memory (PSTM) moderate the effects of the different types of feedback?***

To address RQ4, that is, the extent to which working memory and phonological short-term memory mediate the efficacy of different types of written CF, descriptive statistics, correlation and multiple regression were conducted. Table 5.13. displays the descriptive statistics of feedback type groups and working memory and phonological short-term memory scores. The direct CF group had the highest working memory and metalinguistic explanation the lowest capacity. The direct CF plus revision group had the highest phonological short-term memory and metalinguistic explanation had the lowest capacity. One way ANOVAs showed no significant differences between the feedback groups in terms working memory:  $F(3, 78) = 14.09, p = .93$ , and phonological short-term memory:  $F(3, 78) = .91, p = .43$ .



Table 5.13. Descriptive Statistics of the Feedback Groups' Working Memory and Phonological Short-term Memory.

Groups	N	WM		PSTM	
		M	SD	M	SD
DC	20	.04*	.67	40.40	9.82
DC+R	20	-.07	.53	41.50	11.84
ME	20	-.007	.60	36.60	8.42
ME+R	19	.03	.59	40.63	10.27

\*The scores of working memory are in terms of average *z* scores. Note: DCF = direct corrective feedback with no revision, ME = metalinguistic explanation with no revision, DCF+R = direct corrective feedback with revision, ME+R = metalinguistic explanation with revision, WM= working memory, PSTM= phonological short-term memory

### 5.5.1 Correlation and Regression Findings for Experimental Groups

Correlation analyses were employed to describe the direction and strength of the linear relationship between working memory or phonological short-term memory and feedback types. The findings for the experimental groups have been outlined in Table 5.14.

Table 5.14. Correlation Results for Experimental Groups

Feedback	IDs	Pre-test		Post 1		Post 2	
		<i>r</i>	<i>p</i>	<i>r</i>	<i>p</i>	<i>r</i>	<i>p</i>
DC	WM	.38	.09	.28	.22	.23	.32
	PSTM	.21	.37	-.09	.68	.29	.20
DC+R	WM	.28	.22	.23	.32	.53	.01*
	PSTM	-.35	.13	-.30	.19	-.36	.11
ME	WM	.21	.36	.55	.01*	.62	.003*
	PSTM	.20	.39	.12	.61	.23	.32
ME+R	WM	.30	.20	.19	.42	.41	.07
	PSTM	.14	.55	-.18	.46	-.21	.37

Note: DCF = direct corrective feedback with no revision, ME = metalinguistic explanation with no revision, DCF+R = direct corrective feedback with revision, ME+R = metalinguistic explanation with revision, WM= working memory, PSTM= phonological short-term memory

As shown, working memory was not correlated with post-test scores after direct CF treatments. However, working memory had a strong correlation with the delayed direct CF plus revision post-test scores ( $r=.53$ ,  $p<0.05$ ). The findings also show that working memory had a strong correlation with the immediate and delayed metalinguistic explanation post-test scores. That is, the correlation of working memory and short-term (post-test 1) scores was:  $r=.55$ ,  $p=.01$ ; and the correlation of working memory and long-term (posttest 2) scores was:  $r=.62$ ,  $p=.003$ . Table 5.14 also shows that phonological short-term memory was not correlated with the post-test scores of all groups.

In order to reveal the unique contribution of working memory and phonological short-term memory to the efficacy of different types of WCF, the scores of working memory, phonological short-term memory and feedback types were subjected to multiple regression analysis. Table 5.15 shows the standardized regression coefficient ( $\beta$ ) and significance value ( $P$ ) for each predictor – working memory and phonological short-term memory – also the R squared value ( $R^2$ ) for each regression model. The standardized coefficient represents the number of standard deviation changes in the outcome variable as a result of one standard deviation change in the predictor variable.  $R^2$  refers to the percentage of variance in the dependent variable which is explained by each independent variable (e.g.,  $R^2=.43$  means 43% of the variance in the depended variable was accounted for by the predictor or independent variable).

As shown in Table 5.15, working memory did not emerge as a significant predictor of variables for the direct CF group. However, working memory was a significant predictor of the post-test 2 scores for the direct CF plus revision group ( $R^2=.42$ ,  $Beta=.54$ ,  $p=.01$ ). Working memory was a significant predictor of the short-term scores ( $R^2=.37$ ,  $Beta=.73$ ,  $p=.006$ ) and long-term scores ( $R^2=.42$ ,  $Beta=.74$ ,  $p=.004$ ) for the metalinguistic explanation group. However, working memory was not a significant predictor for

metalinguistic explanation plus revision. With regards to phonological short-term memory, Table 5.15 shows that phonological short-term memory was not revealed to be a significant predictor of variables for the direct CF, metalinguistic explanation, metalinguistic explanation plus revision groups, however, phonological short-term memory was a significant predictor of the post-test 2 scores for the direct CF plus revision group, but with a negative coefficient ( $R^2 = .42$ ,  $Beta = -.38$ ,  $p = .05$ ). That is, the participants with lower levels of phonological short-term memory showed greater gains.

*Table 5.15. Regression Results for the Effects of Experimental Groups and Contributions of WM and PSTM*

Groups		Predictor		Predictor		
		WM		PSTM		
		<i>B</i>	<i>P</i>	<i>B</i>	<i>P</i>	<i>R</i> <sup>2</sup>
<b>DC</b>	Post-test 1	.45	.10	-.33	.21	.16
	Post-test 2	.10	.70	.24	.37	.10
<b>DC+R</b>	Post-test 1	.24	.29	-.31	.18	.15
	Post-test 2	.54	<b>.01*</b>	<b>-.38</b>	<b>.05</b>	.42
<b>ME</b>	Post-test 1	<b>.73</b>	<b>.006*</b>	-.30	.21	.37
	Post-test 2	<b>.74</b>	<b>.004*</b>	-.20	.39	.42
<b>ME+R</b>	Post-test 1	.16	.51	-.14	.55	.06
	Post-test 2	.39	.10	-.14	.54	.19

*Note: DCF = direct corrective feedback with no revision, ME = metalinguistic explanation with no revision, DCF+R = direct corrective feedback with revision, ME+R = metalinguistic explanation with revision, WM= working memory, PSTM= phonological short-term memory*

### 5.5.2 Correlation and Regression Findings for DMWR (direct CF and metalinguistic explanation) and DMPR (direct CF plus revision and metalinguistic explanation plus revision)

To address RQ4, that is, the extent to which working memory and phonological short-term memory mediate the efficacy of different types of DMWR (direct CF and metalinguistic explanation) and DMPR (direct CF plus revision and metalinguistic

explanation plus revision), descriptive statistics, correlation, and multiple regression were conducted. Table 5.16 displays the descriptive statistics of feedback type groups and working memory and phonological short-term memory scores. The DMWR group (direct CF and metalinguistic explanation) had higher capacity working memory than DMPR (direct CF plus revision and metalinguistic explanation plus revision). The DMPR group (direct CF plus revision and metalinguistic explanation plus revision) had higher capacity phonological short-term memory than DMWR (direct CF and metalinguistic explanation).

#### *5.16 Descriptive Statistics of the Feedback Groups' Working Memory and Phonological Short-term Memory*

Groups	N	WM		PSTM	
		M	SD	M	SD
<b>DMWR</b>	40	.02	.63	38.50	9.23
<b>DMPR</b>	39	-.02	.56	41.07	10.96

*Note: DMWR group stands for the combined direct CF and metalinguistic explanation groups, DMPR group stands for the combined direct CF plus revision and metalinguistic explanation plus revision groups, WM stands for working memory, PSTM stands for phonological short-term memory.*

The correlation findings for the DMWR group (a combination of direct CF and metalinguistic explanation) and the DMPR group (a combination of direct CF plus revision and metalinguistic explanation plus revision) are shown in Table 5.17.

*Table 5.17. Correlation Results for DMWR and DMPR*

Feedback	IDs	Pre-test		Post 1		Post 2	
		<i>r</i>	<i>p</i>	<i>r</i>	<i>p</i>	<i>r</i>	<i>p</i>
<b>DMWR</b>	WM	.26	.10	<b>.40</b>	<b>.01*</b>	<b>.36</b>	<b>.02*</b>
	PSTM	.12	.44	.01	.94	.21	.18
<b>DMPR</b>	WM	.30	.06	.21	.19	<b>.42</b>	<b>.002*</b>

	PSTM	-.12	.46	-.24	.13	-.30	.06
--	------	------	-----	------	-----	------	-----

*Note: DMWR group stands for the combined direct CF and metalinguistic explanation groups, DMPR group stands for the combined direct CF plus revision and metalinguistic explanation plus revision groups, WM stands for working memory and PSTM stands for phonological short-term memory*

As revealed in Table 5.17, WM had a medium significant correlation with the short-term scores ( $r=.40, p=.01$ ) and long-term scores ( $r=.36, p=.02$ ) for the DMWR group (direct CF and metalinguistic explanation). Working memory had a medium significant correlation with the long-term scores for the DMPR group ( $r=.42, p=.002$ ). Phonological short-term memory did not show any correlation with the scores for the DMWR group (direct CF and metalinguistic explanation). However, phonological short-term memory had a negative approaching significant correlation with post-test scores for the DMPR group (direct CF plus revision and metalinguistic explanation plus revision).

Table 5.18 shows that working memory was a significant predictor of the short-term scores ( $R^2=.23$ , Beta=.57,  $p=.002$ ), and approaching as a significant predictor of the long-term scores ( $R^2=.13$ , Beta=.35,  $p=.059$ ) for the DMWR group (direct CF and metalinguistic explanation).

*Table 5.18. Regression Results for the Effects of DMWR and DMPR and Contributions of WM and PSTM*

Predictors						
Groups		WM		PSTM		
		B	P	B	P	R <sup>2</sup>
<b>DMWR</b>	Post-test 1	<b>.57</b>	<b>.002*</b>	-.32	.06	.23
	Post-test 2	<b>.35</b>	<b>.059</b>	.02	.89	.13
<b>DMPR</b>	Post-test 1	.19	.23	-.23	.15	.09
	Post-test 2	<b>.45</b>	<b>.003*</b>	-.26	.06	.30

*Note: DMWR group stands for the combined direct CF and metalinguistic explanation groups, DMPR group stands for the combined direct CF plus revision and metalinguistic explanation plus revision groups, WM stands for working memory and PSTM stands for phonological short-term memory*

Phonological short-term memory was not shown to be a significant predictor of the short-term and long-term scores in the DMWR group (direct CF and metalinguistic explanation). Working memory was a significant predictor of the long-term scores for the DMWR group (direct CF plus revision and metalinguistic explanation plus revision) ( $R^2=30$ ,  $Beta=.45$ ,  $p=.003$ ). Phonological short-term memory did not show to be a significant predictor of the short-term and long-term scores for the DMWR group (direct CF plus revision and metalinguistic explanation plus revision).

### **5.5.3 Correlation and Regression Findings for DCO (direct CF and direct CF plus revision) and MEO (metalinguistic explanation and metalinguistic explanation plus revision)**

To address RQ4, that is, the extent to which working memory and phonological short-term memory mediate the efficacy of different types of DCO (direct CF and direct CF plus revision) and MEO (metalinguistic explanation and metalinguistic explanation plus revision), descriptive statistics, correlation, and multiple regression were conducted. Table 5.19 displays the descriptive statistics of feedback type groups and working memory and phonological short-term memory scores. The MEO group (metalinguistic explanation and metalinguistic explanation plus revision) had higher capacity working memory than DCO (direct CF and direct CF plus revision). The DCO (direct CF and direct CF plus revision) group had higher capacity phonological short-term memory than MEO (metalinguistic explanation and metalinguistic explanation plus revision).

*Table 5.19. Descriptive Statistics of the Feedback Groups' Working Memory and Phonological Short-term Memory*

Groups	N	WM		PSTM	
		M	SD	M	SD
<b>DCO</b>	40	-.01	.60	40.95	10.75
<b>MEO</b>	40	.01	.59	38.56	9.46

*Note: DCO group stands for the combined DC groups (i.e., direct CF and direct CF plus revision), MEO group stands for the combined ME groups (i.e., metalinguistic explanation and metalinguistic explanation plus revision), WM stands for working memory and PSTM stands for phonological short-term memory.*

The correlation results for the DCO group (direct CF and direct CF plus revision) and the MEO group (metalinguistic explanation and metalinguistic plus revision) are shown in Table 5.20.

*Table 5.20. Correlation Results for DCO and MEO*

Feedback	IDs	Pre-test		Post 1		Post 2	
		<i>r</i>	<i>p</i>	<i>r</i>	<i>p</i>	<i>r</i>	<i>p</i>
<b>DCO</b>	WM	.26	.10	.23	.14	.26	.10
	PSTM	-.07	.66	-.15	.33	.05	.72
<b>MEO</b>	WM	.25	.05	<b>.40</b>	<b>.005*</b>	<b>.54</b>	<b>.00*</b>
	PSTM	.20	.11	.02	.43	.07	.32

*Note: DCO group stands for the combined DC groups (i.e., direct CF and direct CF plus revision), MEO group stands for the combined ME groups (i.e., metalinguistic explanation and metalinguistic explanation plus revision), WM stands for working memory and PSTM stands for phonological short-term memory.*

As demonstrated in Table 5.20, working memory had a medium significant correlation with the short-term scores ( $r=.40, p=.005$ ) for the MEO group (metalinguistic explanation and metalinguistic plus revision); working memory had a strong significant correlation with the long-term scores for the MEO group ( $r=.54, p=.00$ ). Phonological short-term memory did not emerge as a correlation with the MEO group (metalinguistic explanation and metalinguistic plus revision). Working memory and phonological short-term memory

were not correlated with post-test scores for the DCO group (direct CF and direct CF plus revision).

Table 5.21 reveals that working memory was approaching as a significant predictor of the short-term scores for the DCO group ( $R^2 = .10$ , Beta=.30,  $p=.07$ ). WM was a significant predictor of the short-term scores ( $R^2 = .16$ , Beta=.41,  $p=.01$ ) and long-term scores ( $R^2 = .29$ , Beta=.54,  $p=.001$ ) for the MEO group (metalinguistic explanation and metalinguistic plus revision).

*Table 5.21. Regression Results for the Effects of DCO and MEO and Contributions of WM and PSTM*

Groups		Predictors				
		WM		PSTM		
		<i>B</i>	<i>P</i>	<i>B</i>	<i>P</i>	$R^2$
<b>DCO</b>	Post-test 1	.30	.07	-.23	.14	.10
	Post-test 2	.26	.11	-.01	.93	.06
<b>MEO</b>	Post-test 1	.41	.01*	-.04	.78	.16
	Post-test 2	.54	.001*	.01	.89	.29

*Note: DCO group stands for the combined DC groups (i.e., direct CF and direct CF plus revision), MEO group stands for the combined ME groups (i.e., metalinguistic explanation and metalinguistic explanation plus revision), WM stands for working memory and PSTM stands for phonological short-term memory.*

Phonological short-term memory did not show to be a significant predictor of variables for the DCO (direct CF and direct CF plus revision) and MEO (metalinguistic explanation and metalinguistic explanation plus revision) groups.

To summarize, the results showed that working memory was a significant predictor of both the short-term and the long-term scores for the metalinguistic explanation, MEO (metalinguistic explanation and metalinguistic explanation plus revision) and DMWR (direct CF and metalinguistic explanation) groups. Working memory was also a significant predictor of only the long-term scores for the direct CF plus revision and



DMPR (direct CF plus revision and metalinguistic explanation plus revision) groups. Phonological short-term memory was a negative significant predictor of the long-term scores for the direct CF plus revision.

## **5.6. Conclusion**

In order to address RQ1, RQ2 and RQ3, a series of one-way ANOVAs was conducted to measure between groups ANOVA. Repeated measure ANOVA was also run to measure within groups. Additionally, effect size in the form of Cohen's *d* was run to measure the size of difference between groups.

The findings of the between-groups ANOVA for RQ1, showed that all experimental groups were effective both in the short term (gains 1) and in the long term (gains 2) in comparison with Control. The results of Cohen's *d* for both gains 1 and 2 showed that all experimental groups had large effect sizes. The value of direct CF group was higher than other groups. In other words, direct CF treatment was relatively more effective than other treatments. The results of within-groups ANOVA for RQ1 showed that all experimental groups were effective over time; however, the direct CF group was relatively more effective in the short term and in the long term.

The results of between-groups ANOVA for RQ2 revealed that both DMPR (direct CF plus revision and metalinguistic explanation plus revision) and DMWR (direct CF and metalinguistic explanation) outperformed control groups. Cohen's *d* values showed that both experimental groups had large effect sizes; however, DMWR (direct CF and metalinguistic explanation) had a higher value than DMPR (direct CF plus revision and metalinguistic explanation plus revision) in the short and long term, and so DMWR (direct CF and metalinguistic explanation) was more effective. Within-groups ANOVA showed that both the DMWR (direct CF and metalinguistic explanation) and the DMPR (direct CF plus revision and metalinguistic explanation plus revision) treatments were

effective over time. Cohen's *d* showed that DMWR was relatively more effective in the short term and DMWR was longer lasting in the long term.

The results of between-groups ANOVA for RQ3 showed that both the DCO (direct CF and direct CF plus revision) and the MEO (metalinguistic explanation and metalinguistic explanation plus revision) groups were effective in the short and long term. Cohen's *d* showed that DCO was relatively more effective than MEO in both post-test 1 and post-test 2. The findings of the within-groups ANOVA showed that both experimental groups were effective over time; however, DCO was relatively more effective in the short term and longer lasting than MEO.

In order to answer RQ4, correlation and standard multiple regression were conducted. The results of correlation revealed that working memory had a strong correlation with the long-term direct CF plus revision, short and long-term metalinguistic explanation post-test scores, and with the long-term scores for MEO (metalinguistic explanation and metalinguistic explanation plus revision). Working memory had a medium significant correlation with short-term scores in the MEO group (metalinguistic explanation and metalinguistic explanation plus revision), with the short-term and long-term scores in the DMWR group (direct CF and metalinguistic explanation), and with the long-term scores for the DMWR group (direct CF plus revision and metalinguistic explanation plus revision). Phonological short-term memory had a negative approaching significant correlation with post-test scores for the direct CF plus revision group.

The results of standard regression revealed that working memory was a significant predictor of both the short-term and the long-term scores for the metalinguistic explanation, MEO (metalinguistic explanation and metalinguistic explanation plus revision) and DMWR (direct CF and metalinguistic explanation) groups. Furthermore, working memory was a significant predictor of the long-term scores for the direct plus

revision and DMPR (direct CF plus revision and metalinguistic explanation plus revision) groups. Phonological short-term memory was a negative significant predictor of the long-term scores for the direct CF plus revision group.

## Chapter 6

### DISCUSSION

#### 6.1 Introduction

This chapter will discuss the findings of the four research questions by referring to the theories and research in chapters 2 and 3. The first sections (6.2; 6.2.1; 6.2.2; 6.2.2.1; 6.2.2.2; 6.2.3) discuss research question 1. The following sections (6.3; 6.3.1; 6.3.2) discuss research question 2. After that, the chapter goes on to discuss the results for research question 3 (6.4). Then, research question 4 will be discussed (6.5). Finally, the conclusion of the chapter will be presented (6.6).

#### **6.2 RQ1: What effect does focused DCF and ME with and without revision have on learners' use of the English passive voice in an immediate text revision and in new texts over time?**

RQ1 investigated the efficacy of direct CF and metalinguistic explanation feedback with and without revision on learners' use of the English passive voice in an immediate text and in new texts over a four-week period. The findings for RQ1 are discussed in the following sections : (1) the findings for a within group comparison immediately and over time; (2) the findings for a between group comparison immediately and over time, which is discussed in two parts: (a) comparing the difference between the treatment groups and the control group immediately and over time; (b) comparing the difference between the individual treatment groups immediately and over time; and (3) the findings of the impact of written CF on the passive voice as a complex structure.

### **6.2.1 Discussion of within group comparison of findings immediately and over time**

This section provides a theoretical and empirical discussion of the within group comparison of the findings immediately and over time. The results of a series of ANOVAs showed that the experimental groups (direct CF, direct CF plus revision, metalinguistic explanation, metalinguistic explanation plus revision) significantly improved their accuracy from the pre-test to the immediate post-test. Then, from the immediate to the delayed post-test the improvement deteriorated slightly, but the decrease in accuracy was not statistically significant. This reveals that some learners retained the improvement from the immediate to the delayed post-test. Furthermore, the accuracy in the delayed post-test was significantly higher than that of the pre-test. Additionally, the accuracy rate for the control group revealed no significant improvement from the pre-test to the immediate post-test and from the immediate post-test to the delayed post-test. These findings not only support the theoretical expectations, but also confirm the results of previous studies.

Theoretically, Gass's (1997) framework identifies how a single episode of input processing (e.g., in the form of written CF) may help learners to develop their explicit knowledge. Accordingly, if noticed and comprehended, input can subsequently go through central processing (i.e., intake and integration), and result in output. Thus, the findings of the current study have shown that a single written CF treatment can help learners to improve the accuracy of a complex structure such as the English passive voice, and that accuracy is not only evident in the short term but also in the writing of a new text after a period of time.

Empirically, these results are consistent with earlier studies (e.g., Bitchener & Knoch, 2008; Bitchener & Knoch, 2010a; Rummel, 2014; Sheen 2007; Stefanou & Revesz 2015) on the effectiveness of written CF. Bitchener and Knoch's (2008) study on 144 low

intermediate ESL learners reported that written CF was effective in the functional use of articles over two months. Similarly, Bitchener and Knoch (2010a), in another longitudinal study, explored the relative effectiveness of the different written CF types (direct corrective feedback; direct corrective feedback and written meta-linguistic explanation; direct corrective feedback, oral, and written meta-linguistic explanation; the control group) on two functional uses of the English article system. They found that written CF was effective after 10 months. Written CF has also been reported to be effective in the use of English articles over nine weeks (Sheen, 2007). Likewise, in their study of 89 Greek EFL learners, Stefanou and Revesz (2015) found that written CF was effective in the use of articles with generic and specific plural referents over four weeks. Similarly, Rummel's (2014) study on 72 advanced EFL learners at Kuwait and Laos universities found that written CF helped learners improve accuracy in the use of the simple past tense and the present perfect tense over seven weeks. Thus, the current findings reveal further evidence of the potential effectiveness of written CF.

### **6.2.2 Discussion of between group comparison of findings immediately and over time**

This section provides a theoretical and empirical discussion of the findings of the between group comparison in two sections, the first (6.2.2.1) comparing the difference between the treatment groups and the control group immediately and over time, and the second (6.2.2.2) comparing the difference between individual treatment groups immediately and over time. The effectiveness of written CF on all groups in the short term and over time is discussed below.

### **6.2.2.1 Discussion comparing the effectiveness of the treatment groups and the control group immediately and over time**

The results of the between group analysis showed that all experimental groups (direct CF, direct CF plus revision, metalinguistic explanation, metalinguistic explanation plus revision) outperformed the control group in both the short term and over four weeks. There is a theoretical explanation for why written CF can be considered to play a facilitative role in L2 development. Gass's framework (1997) shows that if learners are able to modify their output based on the input they have been given, they must have proceeded successfully through the cognitive processing stages (apperceived [noticed] input, comprehended input, intake, integration) to reach that modified output.

The findings are also empirically supported (e.g., Bitchener & Knoch 2010a; Ellis et al., 2008). Bitchener and Knoch (2010a) investigated the relative effectiveness of written CF on two functional uses of the English article system. The findings revealed that all three groups that received treatments outperformed the control group on all post-tests. Similarly, Ellis et al.'s (2008) study on 49 intermediate EFL learners found that those who received written CF outperformed the control group both immediately and in post-tests over 10 weeks. However, the findings of the present study are in contrast to those of Guo (2015). In her study of 147 low intermediate EFL learners, Guo (2015) reported that learners who were provided with written CF outperformed the control group only in the immediate post-test, but not after four months. A possible reason for the difference in findings between the study by Guo and the current study could be the use of different study designs; that is, the delayed post-test in this study was conducted in week four and in Guo's study it was conducted after four months. Participants in Guo's study may not have been able to consolidate their knowledge of the target structures, which were the regular and irregular past tense and prepositions of place, after four months. As Sheen (2007) suggested, the passage of time may be a critical factor in terms of the effectiveness

of delayed post-tests. Thus, a longitudinal study is recommended in treating the passive voice as it would enable the researcher to document trends related to the effectiveness of written CF over a longer time frame. Guo (2015) also argued that a possible reason for not sustaining accuracy over time is that the participants in her study were at a low proficiency level. Thus, they needed more practice or instruction to develop their explicit knowledge. Additionally, because they had a low proficiency level, their explicit knowledge of the target structures was insufficient to draw upon. Because of this, they could not retain the improved accuracy demonstrated on their immediate post-test after four months.

#### **6.2.2.2 Discussion comparing the difference between individual treatment groups immediately and over time**

The findings also showed that direct CF treatment was relatively more effective than the other treatments in both the short term and over time, and also from the pre-test to the immediate and delayed post-tests. This can be explained theoretically in that the degree of explicitness provided to the direct CF group was higher than that of the metalinguistic explanation group. In other words, in this study, direct CF was more explicit because the direct CF group was provided with the correct form of the target structure, the English passive voice. On the other hand, metalinguistic explanation was less explicit because the feedback for the metalinguistic explanation group was in the form of a one-page handout that included when to use the English passive voice, how to construct it and some examples; however, the errors in the English passive voice were not identified in the students' written text. Thus, the metalinguistic group received no explicit feedback on their pre-test writings. As has been discussed in Chapter 2, Section 2.9, learners' existing knowledge and the salience of the input play important roles in the processing of the input,



and more explicit types of written CF are likely to draw learners' attention to a greater extent than less explicit written CF types (e.g., metalinguistic explanation as used in this study). More explicit types of feedback (e.g., direct CF) may also reduce the confusion that learners may experience if they do not understand less explicit types of CF. This is because more explicit types of feedback explicitly reveal learners' errors and may help learners to more clearly and fully comprehend their errors than less explicit CF. Additionally, more explicit types of written CF contain more linguistic information, which may lead to the formation of a new hypothesis about the target structure (e.g., the English passive voice) and the production of output. Thus, because the degree of explicitness of written CF types may impact the level of attention learners pay to the input, and because the amount of linguistic information included in the input facilitates the formation of a new hypothesis, a more explicit type of written CF (e.g., direct CF) is more likely to lead to output.

Empirically, the findings of the current study regarding the high level of effectiveness of direct CF in comparison to other types of written CF are supported by the results of earlier studies. Bitchener (2008) reported that providing only direct CF was more effective than providing direct CF plus metalinguistic explanation over two months. He argued that the possible reason was that the limited details of written metalinguistic explanation may have been insufficient to result in a significant effect. Shintani et al. (2014) also reported that direct CF was more effective than metalinguistic explanation provided in the form of a handout in the immediate post-test. The reason was that the direct CF was more explicit and provided participants with the correct form of the target structure and thus enabled them to make a comparison between the input and their own erroneous structure. However, the metalinguistic explanation feedback was less explicit as it was in the form of a handout and the errors in the use of the targeted structure were not identified on the students' written text, and thus the learners were required to apply an abstract explanation

of the target structure to recognize and correct their errors. Similarly, Van Beuningen et al. (2008) found that direct CF, as an explicit type of feedback, was more effective than other less explicit types. In their study of 62 Dutch EFL students in a secondary school, the researchers used two treatment groups (direct correction and error codes) and compared them with two control groups that had writing practice and self-correction. The findings showed that only the direct correction group showed a significant improvement in accuracy in new texts.

In her study of 147 low intermediate EFL learners, Guo (2015) also reported that learners who were provided with more explicit types of feedback (direct error correction, direct error correction plus metalinguistic explanation) outperformed those who were provided with less explicit types of written CF (error code and underlining) in the immediate post-test.

In contrast to the present study, in her study of Laotian EFL learners, Rummel (2014) found no difference between the direct error correction group and the metalinguistic explanation group in treating target structures (simple past tense and present perfect tense). A possible reason is that the metalinguistic explanation in the present study was in the form of a handout and the errors in the target structure were not identified on the students' writing. However, in Rummel's study the metalinguistic group received explicit feedback in the form of identifying the errors and providing explicit explanations on them. Thus, in the present study, direct CF and metalinguistic explanation feedback had a different degree of explicitness, that is, the direct CF was more explicit than the metalinguistic explanation in the form of a one-page handout. The direct CF and metalinguistic CF in her study were both explicit CF.

In the same study, Rummel found that for Kuwaiti learners, the group that received direct error correction outperformed the metalinguistic explanation groups. The researcher recommended that the differences in teaching and learning approaches in Kuwait and

Laos as well as the learners' beliefs towards different types of written corrective feedback may have produced the different findings in the two contexts.

### **6.2.3 The findings on the impact of written CF on the passive voice as a complex structure**

The findings of the current study also showed that written CF is effective in terms of effectively targeting a complex structure, the English passive voice, immediately and after week four. The English passive voice can be categorized as a complex structure because it is formed by both rule-based and item-based structures. In the passive voice, the verb "to be" can be categorized as rule-based while the "past participle" can be categorized as item-based. For example, in the sentences *The door is closed*. *The vegetables are grown*. *The rug picnic is spread*, "to be" is rule-based as there is a rule that identifies when learners need to use "am", "is" or "are". However, the past participle is item-based as it takes three different forms: regular verbs that end in "ed" (e.g., "closed"); irregular verbs in which the past participle changes form (e.g., "grown"); and irregular verbs where the past participle retains the same form (e.g., "spread"). Thus, structures such as the passive voice that are both rule-based and item-based require learners to employ more attentional capacity when using them.

Empirically, as far as can be ascertained, there are only two studies that have targeted complex structures: Shintani et al. (2014) and Rummel (2014) targeted the hypothetical conditional and the present perfect tense respectively. Similar to the present study, Rummel (2014) found that written CF was effective for the present perfect tense immediately and over time (seven weeks). However, in contrast to the present study, Shintani et al. (2014) reported that learners did not sustain improved accuracy in the use of the hypothetical conditional over two weeks. Bitchener (2016) argued that a number

of variables, including linguistic error type and written CF type, may interact with one another and thus impact the extent to which written CF impedes or facilitates the accuracy of learners' output. The reason for this is that the development of morphological, syntactic and lexical knowledge requires an understanding of meaning, form and use in relation to other parts of the language system and to other words (Ortega, 2009), and students may need to focus their attention on more than one linguistic element each time they use a form or structure. Additionally, it has been claimed by Young (1996) that one linguistic form or structure may be more difficult to learn than another and that different linguistic forms and structures may be learnt at different cognitive stages (Pienemann, 1998). Thus, some linguistic forms and structures may be more 'treatable' than others (Ferris, 1999). Therefore, as a result of written CF, some forms and structures may develop more easily than others. Thus, it is possible differences in the target structures employed in Shintani et al.'s (2014) study and the present study, that is, the hypothetical conditional and the English passive voice respectively, resulted in different findings. The hypothetical conditional is more complex than the passive voice as it is formed by seven components ((1) the past tense, (2) the perfect aspect, (3) the past participle (PP) form in the *if* clause, (4) the modal, (5) the past tense, (6) the perfect aspect and (7) the PP form in the main clause) compared to the passive voice which has two components (the verb "to be" and the past participle). Celce-Murcia and Larsen-Freeman (1999) argued that learning the hypothetical conditional is difficult for learners as the structure is both semantically and syntactically complex; that is, learners need to have knowledge of the tense aspect system, modal auxiliaries and negation to use it accurately. Thus, it is more demanding to attend to, comprehend, process and produce the hypothetical condition with seven components over time rather than the passive voice with two components. Additionally, it may be possible that the learners in Shintani et al.'s study developed their explicit knowledge with the aid of more explicit types of written corrective feedback (i.e. direct

CF); however, they were unable to consolidate their knowledge in order to use it after a period of time as a result of the high level of complexity of the hypothetical conditional. In addition to the reasons recommended above, the differences in the findings may also be the results of additional factors including the participants' proficiency levels and other individual and contextual factors. It is not surprising, therefore, that learners in Shintani et al.'s (2014) study retained the hypothetical conditional only in the post-test (1) whereas participants in this study were able to retain accurate use of the passive voice in both the immediate and delayed post-tests.

### **6.3 RQ2: Does the opportunity for revision influence the efficacy of DC and ME?**

In order to address research question 2, that is, investigating whether requiring learners to undertake revision has any effect on the accuracy in new pieces of writing, the two revision groups (i.e., direct CF plus revision and metalinguistic plus revision) and the two groups that did not make revisions (i.e., direct CF and metalinguistic explanation) were combined and compared. The findings for RQ2 are discussed in the following two subsections: 5.3.1 discusses the findings for the within group comparison immediately and over time and 5.3.2 discusses the findings for the between group comparison immediately and over time.

#### **6.3.1 Discussion of the within group comparison of findings immediately and over time**

Providing learners with opportunities to revise their texts may play an important role in the development process because it invites them to notice the feedback they have been provided with and to process it across the stages identified in Gass' (1997) cognitive framework (Bitchener, 2016). However, the effectiveness of revising the text while having access to written CF is arguable, because little or no cognitive processing may take place. Bitchener (2016) pointed out that when learners revise their text and have

access to the corrections (e.g. direct CF) little or no cognitive processing may be needed. Thus, this study, to the best of my knowledge, is the first study in which the learners revised their texts and had no access to the text on which they had received feedback.

The study found that both revision groups (i.e., metalinguistic explanation plus revision and direct CF plus revision) and non-revision groups (direct CF and metalinguistic explanation) had improved accuracy from the pre-test to the immediate post-test while there was no significant change for the control group. From the immediate post-test to the delayed post-test, both revision and non-revision groups' accuracy atrophied a little, but the change was not significant. In other words, from the immediate to the delayed post-test the revision and non-revision groups retained accuracy in the use of the passive voice. From the immediate to the delayed post-test there was no significant change for the control group.

Revision following feedback may lead to greater accuracy in new writing texts. This can be explained theoretically in that it is likely that written CF followed by revision leads to 'pushed output', especially if learners have no access to the corrections when they start writing the revision draft (as in the present study) (Shintani et al., 2014). Swain (1995) argued that when learners are pushed to produce language, they are likely to "notice a gap between what they want to say and what they can say, leading them to recognize what they do not know, or know only partially" (pp. 125-126). Swain maintained that pushed output helps learners to notice the grammatical forms that otherwise are likely to go unattended. Additionally, revisions require information retrieval from long-term memory and facilitate the consolidation and proceduralization of L2 knowledge.

Empirically, the findings of this study on the effectiveness of written CF followed by revision are in line with Van Beuningen et al.'s (2008, 2012) findings that reported both experimental groups (error code and direct error correction) increased their accuracy in

text revisions; however, learners in the first experimental group (i.e., direct error correction) were able to write a new piece of writing with improved accuracy a week later. In their main study (2012), the authors reported that after four weeks, all 268 learners retained the same level of accuracy in the delayed post-test as was recorded in the text revision. However, the findings of the current study differ from those of Truscott and Hu (2008), who found that the increase in accuracy shown by their experimental group when revising their texts was not shown in their writing of new texts. Thus, they concluded that “the successful error reduction during revision is not a predictor ... of learning” (p. 299). However, Bruton (2009) challenged the validity of the claim because in Truscott and Hsu’s (2008) study learners made only a few errors in their pre-test writing and therefore had little room for improvement. Additionally, one of the main differences between Truscott and Hsu’s (2008) and Van Beuningen et al.’s (2008,2012) studies and the present study is the degree of explicitness of written CF that learners received, and this may explain the contradictory findings. In contrast to Truscott and Hsu’s (2008) study in which one type of written CF with a low level of explicitness (underlining) was employed, in the present study and in Van Beuningen et al.’s (2008, 2012) studies more explicit written CF types were used (i.e. direct correction).

The findings of within group analysis also showed that even though both revision and non-revision groups had improved accuracy from the pre-test to the immediate and delayed post-tests, the non-revision group (i.e., direct CF and metalinguistic explanation) had greater accuracy in the immediate post-test but the accuracy of the revision group (i.e., direct CF plus revision and metalinguistic plus revision) was relatively longer lasting. These results may indicate the long-term advantage of requiring learners to perform revision following feedback. Requiring learners to revise their initial text may help them to process feedback more deeply and to consolidate their declarative knowledge of the passive voice.

Empirically, the findings confirm Shintani et al.'s (2014) results in which both the revision (i.e., direct CF plus revision and metalinguistic plus revision) and non-revision groups (i.e., direct CF and metalinguistic explanation) showed improved accuracy from the pre-test to the immediate post-test to the delayed post-test, and that the accuracy of the revision group (i.e., direct CF plus revision and metalinguistic plus revision) was longer lasting. Chandler's (2003) study also found that feedback followed by revision led to a significant improvement in accuracy in the first drafts of new texts over a semester, whereas the control group that did not make revisions following feedback showed no improved accuracy on first drafts over a semester.

### **6.3.2 Discussion of between group comparison of findings immediately and over time**

The results of this study also showed that both revision and non-revision groups had significantly higher accuracy than the control group in the immediate post-test and the delayed post-test. However, the non-revision group (i.e., direct CF and metalinguistic explanation) had relatively greater accuracy than the revision group (i.e., direct CF plus revision and metalinguistic plus revision ) in the immediate and the delayed post-tests.

In the current study, when learners performed revisions, they then had no access to their initial text when they completed their immediate post-test. Furthermore, both the revision and the immediate post-test were undertaken in the same session. For these reasons this finding could be theoretically explained because it is possible that conducting two tests (i.e., revision and immediate post-test) in one session was demanding, thus putting a greater cognitive load on the revision group, especially because the revision group had no access to the initial draft that they had received feedback on.

Empirically, to the best of my knowledge, the study by Shintani et al. (2014) is the only one so far that combined revision groups and non-revision groups and compared the



effectiveness of written CF types. Similar to this study, they found that both revision (i.e., direct CF plus revision and metalinguistic plus revision) and non-revision groups (i.e., direct CF and metalinguistic explanation) had greater accuracy than the control group in the short term. However, in contrast to the present study that found both the revision and non-revision groups were more accurate than the control group in the delayed post-test and that the non-revision group was more accurate than both the revision and the control groups in the delayed post-test, they found that only the revision group had greater accuracy than the control group in the delayed post-test. The possible reason for the difference in the findings could be that learners in the revision groups in Shintani et al.'s study had access to the initial draft that they had received feedback on while they were writing the revision text; however, learners in the current study had no access to the initial draft that they received feedback on. In the study by Shintani et al., access to the first draft meant that learners in the revision group had less cognitive load on their attention and memory, both of which are necessary when correcting errors, than the learners in the present study. Shintani (2017) also argued that having access to explicit instruction (e.g., the initial draft that learners had received feedback on) assisted learners in monitoring the accurate use of the target structure and enabled them to correct their errors in the writing task. Additionally, in the present study, it is possible that conducting two tests (i.e., revision and immediate post-test) in one session was demanding, thus putting a greater cognitive load on the revision group. Shintani et al. (2014) conducted post-test (1) seven days after the treatment session. Thus, it is possible that, in Shintani et al.'s (2014) study, providing the learners with the initial draft and allowing a gap between the treatment session and post-test (1) resulted in different findings.

#### **6.4 RQ3: Is there any difference in the effect of DC and ME regardless of whether there is an opportunity for revision?**

Research question 3 examined the relative efficacy of metalinguistic and direct CF regardless of whether there was an opportunity for revision. Thus, the two direct CF groups (i.e., direct CF and direct plus revision) and the two metalinguistic groups (i.e., metalinguistic and metalinguistic plus revision) were combined and compared.

The findings showed that both combined groups, that is the direct CF groups (i.e., direct CF and direct plus revision) and the metalinguistic explanation groups (i.e., metalinguistic and metalinguistic plus revision) improved significantly from the pre-test to immediate post-test. The accuracy deteriorated slightly from the immediate post-test to the delayed post-test, but the changes were not statistically significant. Moreover, the findings showed that combined direct CF was relatively more effective in the short term and longer lasting than combined metalinguistic explanation and that the combined direct CF group was relatively more effective than the combined metalinguistic explanation group in both the immediate and delayed post-tests.

Theoretically, as has been explained in RQ1, the degree of explicitness and salience of input provided to the combined direct CF group was higher than that of the combined metalinguistic explanation group because the direct CF group was provided with the correct form of the target structure, the English passive voice. On the other hand, the feedback for the metalinguistic explanation group was in the form of a handout explaining the English passive voice and the errors were not identified in their text. Thus, the direct CF (a more explicit type of written CF) is likely to draw learners' attention to a greater extent than metalinguistic explanation in the form of a handout (a less explicit written CF type).

Empirically, Shintani et al. (2014) was the only study which compared the effectiveness of the combined direct CF groups (i.e., direct CF and direct CF plus revision) and the combined metalinguistic explanation groups (i.e., metalinguistic explanation and metalinguistic explanation plus revision). They found that both the combined direct CF and the combined metalinguistic explanation treatments were effective in the short term, but only the combined direct CF was effective over time (i.e. 2 weeks). Similarly, this study found that both direct CF and metalinguistic explanation treatments were effective in short time; however, it found that treatment types were effective in the long term as well, and that the combined direct CF group was more accurate both immediately and over time (i.e. 2 weeks). Even though the metalinguistic explanation in the two studies was in the form of a handout and the errors were not identified in their text, the difference in findings, that is the ineffectiveness of metalinguistic feedback over time in Shintani et al.'s study, may be due to the different target structures in the two studies, that is the hypothetical conditional and the English passive voice respectively. As has been explained in the discussion section of RQ1, the hypothetical condition with seven components is a complex structure and likely to be less treatable than the passive voice which has two components. To put it another way, the learners in Shintani's study were required to notice, comprehend, process and produce seven components. Since in the hypothetical condition they need to notice, comprehend, process and produce seven components, this may be the reason that the combined metalinguistic group was not able to show improved accuracy in the delayed post-test. Another reason for the learners in Shintani's et al. (2014) study not sustaining the accuracy after two weeks could be that the metalinguistic feedback was not explicit. Because, metalinguistic feedback was only provided in the form of a handout and the errors in the English passive voice were not identified in the students' written text, thus processing the metalinguistic explanation of a complex/ less treatable structure (hypothetical condition) was demanding for learners.

Because they were low proficiency learners, it is unlikely that they would even have had instruction on this and therefore it is unlikely that they would have the knowledge in their memory store. Thus, it may be that learners were not able to deeply process it to employ it overtime. Even though in the present study metalinguistic explanation was also less explicit, the degree of complexity of the passive voice is less than that of the hypothetical conditional. Additionally, the metalinguistic groups in Shintani et al.'s study were given five minutes to read the handout that included an explanation of the target structure, the hypothetical conditional, and it is possible that they may have needed more time to fully understand the explanation of the hypothetical conditional as a complex structure. However, in this study, the metalinguistic explanation groups were provided 10 minutes to read the handout and so they had more time to understand the usage of the passive voice. Additionally in Shintani et al.'s study, because the learners in the metalinguistic explanation group only received feedback in the form of a handout and the errors in the use of the hypothetical conditional were not identified in the students' written text, it might have been that they did not relate the metalinguistic explanation to specific errors. With regards to the ineffectiveness of the combined ME group, Shintani et al. argued that the complexity of the hypothetical condition might have impacted the effectiveness of ME in the delayed post-test. Drawing on this, Shintani et al. (2014) recommended that "what constitutes "treatability" is not just a question of whether or not a feature is rule-based but also the complexity of the rule-based structure" (p.123). In other words, the degree of complexity of rule based forms/structures may have a varying impact on the effectiveness of written CF. Thus, as it was explained above, the reason the combined metalinguistic explanation was not effective over time in their study could be the high degree of complexity (less treatability) of the hypothetical conditional and the limited degree of explicitness of the metalinguistic explanation feedback. Shintani (2016) also suggested that metalinguistic explanation in the form of a handout mainly leads to

developing the participants' explicit knowledge, which is less durable than automatized or implicit linguistic knowledge.

The study by Shintani, Aubrey and Donnellan (2016) resulted in similar findings to those of Shintani et al. (2014). In their recently published study of 61 Japanese EFL learners, Shintani et al. (2016) investigated the comparative efficacy of pre- and post-task metalinguistic explanation on learners' improved accuracy of the hypothetical conditional. Similar to Shintani et al.'s (2014) study, learners received metalinguistic explanation in the form of a handout. The pre-task group was provided with metalinguistic explanation before the writing task and the post-task metalinguistic explanation group was provided with metalinguistic explanation following completion of the writing task. The results showed that both pre- and post-task groups had improved accuracy in the short term but only the pre-task group maintained accuracy over time (after 3 weeks). Because of the similarity of findings to those of Shintani et al. (2014), the researchers suggested that metalinguistic explanation mainly led to developing explicit knowledge, which is less durable than automatized or implicit linguistic knowledge. However, as Bitchener and Storch (2016) argued, before generalizing and drawing any conclusion regarding the efficacy of written CF types, the effectiveness of written CF types such as metalinguistic explanation needs to be investigated taking into account learners' individual and contextual factors and different linguistic structures.

#### **6.5 RQ 4 To what extent do working memory and phonological short-term memory moderate the effects of different types of feedback?**

Research question 4 investigates the extent to which working memory and phonological short-term memory moderate the efficacy of direct corrective feedback and metalinguistic explanation with and without revision in new writing texts immediately and over time.

Working memory is a cognitive device with the dual function of storing and processing information. Working memory is operationalized as either complex working memory, which refers to both storage and processing components, or phonological short-term memory, which only consists of the storage component. In order to address this research question, correlation and standard multiple regression were conducted.

The findings revealed that (1) working memory moderated the efficacy of the metalinguistic explanation and the combined metalinguistic explanation group (ME and ME+R) and the combined metalinguistic and direct CF groups (ME+DC) both immediately and over time; (2) working memory moderated the direct CF plus revision and the combined revision group (ME+R and DCF+R) only in the long term; and (3) phonological short-term memory negatively moderated the efficacy of DC plus revision (DC+R) only in the long term. The findings are discussed as follows.

The possible reason for the first finding, that is, the moderating effect of working memory on the metalinguistic explanation and the combined metalinguistic explanation group and the combined metalinguistic and direct CF groups, is related to the nature of the feedback selected in the present study. In other words, there was a greater degree of explicitness in the direct corrective feedback and a lower degree of explicitness in the metalinguistic explanation feedback. The direct corrective feedback group was provided with corrected errors, which is a more explicit type of feedback, whereas the metalinguistic explanation group received feedback in the form of a handout, which is a less explicit type of feedback. In other words, the metalinguistic explanation group received only received a handout that included an explanation of the targeted structure, which was the English passive voice and the errors in the target structure were not identified on the students' writing. This may reflect two different mechanisms of noticing. The higher level of explicitness of direct corrective feedback might lead to learners noticing the target structure (i.e., the English passive voice) with less cognitive demand compared to the

metalinguistic explanation that provides less explicit feedback and thus does not directly draw learners' attention to form (i.e., the English passive voice) when writing new texts immediately and over time. The noticing of less explicit types of CF (e.g., metalinguistic explanation in the form of a handout) is an attention demanding task and thus requires a domain-general, attention control mechanism considered as a critical component of working memory capacity (e.g., Goo, 2013; Engle, 2002; Kane, Conway, Hambrick, & Engle, 2007). This cognitive control mechanism is not needed for the noticing of direct CF, which has an obtrusive and explicit nature. Thus, this may be the reason that working memory is related to the effectiveness of metalinguistic explanation but not of direct CF in the acquisition of the English passive voice. In other words, it is the involvement of the executive attention process in the noticing of metalinguistic explanation that may distinguish metalinguistic explanation from direct CF in respect to the moderating role of working memory capacity on the efficacy of the two feedback conditions. Theoretically, this makes sense because working memory is a cognitive space in which an erroneous form is compared with the corrected form (Baralt, 2015). Thus, when direct CF is provided on the written text, there is no need for the processing and comparison of forms in the working memory. This suggests that the explicit nature of direct CF may neutralize learners' individual differences in working memory in terms of their ability to notice the target structure (i.e., the English passive voice). Thus, these findings reveal that feedback types with different levels of explicitness may have a different impact on the moderating effect of working memory on written CF.

These findings are in line with Goo's (2012) study that found working memory moderates the effectiveness of recasts as a less explicit type of CF while it is not a predictor of metalinguistic CF because it is more explicit. In Goo's study, metalinguistic CF explicitly targeted the learners' erroneous utterances of the target structure, that is, the English that-trace filter. It appears that the mechanism of working memory works similarly on more

and less explicit types of oral and written CF. However, because the present study is the first study that has investigated the moderating effect of working memory on the effectiveness of written CF, there is a need for much more research in this regard.

The second finding related to working memory also revealed that it moderated the efficacy of revision, that is, direct CF plus revision and the combined revision group (ME+R and DCF+R) only in the long term. The possible reason for this is that when revising their texts, these groups had no access to the initial draft on which they had received feedback; thus, when revising, they may have drawn on their working memory. Therefore, revision enabled the learners to process the feedback more deeply, which helped to consolidate their explicit knowledge of the English passive voice. However, the findings showed that working memory did not moderate the revision groups' output immediately. This may be because the revision groups undertook two tests (i.e., the revision and immediate post-test) in one session and had no access to the pre-test that they had previously received feedback on during revision. This may have placed too great a demand on learners' cognitive control, thus putting a greater cognitive load on the working memory, so they were unable to draw on it. Additionally, the revision groups had no access to the text they had received feedback on, which meant the revision text became a complex task. Thus, the requirement for the revision groups to conduct two tasks in one session with cognitively more complex tasks may have been too great a cognitive load to process feedback in working memory in the short term.

This study also investigated the potentially moderating effect of phonological short-term memory on direct CF and metalinguistic explanation with and without revision. Phonological short-term memory as a component of working memory is operationalized by the phonological loop, which stores information for a short time (Baddeley, 2003; Baddeley & Logie, 1999). Although the modality of the phonological short-term memory test is not the same as the modality of the higher order skill (writing), working memory



is considered a domain-general cognitive ability whose effects are not restricted to a particular domain or skill. Phonological short-term memory has been measured in various ways through non word recall (as was used in this study), digit span, letter span and so on, and all these measures have demonstrated strong predictive power for various aspects of L2 learning. Therefore, regardless of how it is measured, the underlying construct is the same, that is, the ability to store and rehearse information in the phonological loop.

The findings of this study showed that phonological short-term memory only negatively moderates the effectiveness of direct CF plus revision in the long-term. A possible explanation for the finding is that different models of working memory agree that the whole working memory system has a limited capacity, resulting in a trade-off between processing and storage (Shah & Miyake, 1999). Therefore, learners with a high phonological short-term memory are better able to store the accurate form of feedback, which consumes the attentional resources, and thus there are fewer resources available for processing the linguistic structure. However, learners with a low phonological short-term memory tend to process the input in order to find out the explicit rules of the linguistic structure. This processing consumes attentional resources and thus there are fewer resources available for storage of the linguistic structure in the phonological short-term memory. Therefore, a possible reason for the finding that phonological short-term memory moderates the efficacy of only direct CF plus revision in the long term is that when the direct CF plus revision group carried out the revision task, they had no access to the text on which they had received feedback. Therefore, they may have deeply focused on the processing of the feedback, leading to improved accuracy over time, rather than utilising storage which is central in phonological short-term memory. Additionally, the findings that those with less phonological short-term memory benefitted more from feedback can be explained based on the 'less is more' hypothesis (Miyake & Friedman, 1998; Newport, 1990). This theory was firstly proposed by Newport (1990) to explain the

advantage that children have in language learning compared to adults. The applicability of this theory to adult second language learners was investigated by Miyake and Friedman (1998) who pointed out that when applying this hypothesis to adult second language learners, researchers should be prudent because their small capacity is actually quite grate. Thus, the learners with less phonological short term memory resources were better able to attend to the details of the target structure, that is, the English passive voice.

Therefore, the findings of the present study showed that phonological short-term memory only negatively moderates the efficacy of only direct CF plus revision in the long term and that did not moderate the effectiveness of the other types of written CF used in the study, that is, direct CF, metalinguistic explanation and metalinguistic explanation plus revision. In other words, the learners' differences in their phonological short-term memory were not a main factor in their improved accuracy. Thus, the learners' performance on the non-word test as a measurement of phonological short-term memory did not appear to be an important factor in predicting the learners' improved accuracy because, with the exception of direct CF plus revision, there was no significant correlation between measures of phonological short-term memory and the written CF types used in the study.

In oral corrective feedback, several studies have been conducted on the relationship between phonological short-term memory and the efficacy of corrective feedback based on recasts. It is likely that learners with a high phonological short-term memory have a superior ability to store recasts (N. Ellis, 2005), and may benefit from recasts in the long term (Mackey et al., 2002). Revesz (2012) found a significant positive correlation between phonological short-term memory and recasts in an immediate oral post-test. Trofimovich et al. (2007) also reported that learners with superior phonological loop ability benefitted more from recasts in the longer term; however, in their study the testing was conducted between 2 and 12 minutes following the immediate post-test. Thus, the

literature shows no conclusive findings regarding the role of phonological short-term memory in processing recasts.

## **6.6 Conclusion**

To summarize, the results of the present study have answered the major questions raised in written CF studies. First, a more explicit type of written CF (i.e. direct CF) enabled the Iranian EFL learners to improve their accuracy to a higher degree compared to a less explicit type of written CF (i.e., metalinguistic explanation as used in this study) in the short term and over time. This may be because more explicit types of written CF are likely to draw learners' attention to a greater extent than less explicit written CF types (e.g., the delivery of metalinguistic explanation as used in this study). Additionally, more explicit types of written CF contain more linguistic information, which may lead to the formation of a new hypothesis about the target structure (e.g., the English passive voice) and the production of output. Besides, more explicit types of feedback (e.g., direct CF) may also reduce the confusion that learners may experience if they do not understand less explicit types of CF.

Second, written CF followed by revision resulted in deeper processing of information compared to non-revision, and thus improved accuracy was longer lasting. Revision following feedback may lead to greater accuracy in new writing texts because written CF followed by revision leads to 'pushed output', especially if learners have no access to the corrections when they start writing the revision draft (as in the present study). Swain (1985, 1995) argued that that pushed output helps learners to notice grammatical forms that otherwise are likely to go unattended.

Finally, the findings showed that working memory and phonological short term memory moderate some types of written corrective feedback. Working memory seems to moderate

less explicit types of feedback because the noticing of less explicit types of CF (e.g., metalinguistic explanation as used in this study) is an attention demanding task and thus requires a domain-general, attention control mechanism considered to be a critical component of working memory capacity. The findings of this study also showed that phonological short term memory negatively moderates the efficacy of direct CF plus revision, but only in the long term. In other words, the poorer one's phonological short term memory was, the more he/she benefited from the feedback.

## **CHAPTER 7**

### **CONCLUSION**

#### **7.1 Introduction**

In this chapter, the findings from Chapter 5 are discussed further with regard to their contribution to theory, research, methodology and pedagogy. First, in Section 7.2, the main aims of this thesis and the way the aims were achieved are reiterated. This is followed by a summary of the findings for each research question. Section 7.3 discusses the contributions of the present thesis to research, theory, methodology and pedagogy respectively. Section 7.4 presents the limitations of the thesis and then the recommendations for future research are provided in Section 7.5. Finally, the chapter ends with some final remarks in Section 7.6.

#### **7.2 The aims of the study**

The study had two main purposes. The first aim was to examine the effect of different types of written CF (direct and metalinguistic) on learners' output (immediate and delayed) in relation to the targeted structure (passive voice). The motivation for exploring this aim was drawn from the mixed findings from several recent studies (e.g., Shintani & Ellis, 2013; Shintani, et al., 2014) on the effectiveness of written CF. Additionally, it is still not clear to what extent the feedback needs to be explicit and whether certain types of CF are more effective than others in facilitating L2 development (Bitchener, 2016). For this reason, the study focused on two different types of written CF (direct and metalinguistic) in order to show to what extent the type of corrective feedback and its

explicitness can impact learners' subsequent output (immediate and delayed). In this study, direct CF was more explicit than metalinguistic explanation; that is, the direct CF group received explicit corrective feedback on their errors in the use of the target structure (the English passive voice) by being given the correct form, while the metalinguistic explanation was less explicit as this group received corrective feedback in the form of a one-page handout about the target structure only (i.e., the English passive voice) and the errors were not identified in their written texts.

The second purpose of this study was to determine the extent to which individual differences in working memory and phonological short-term memory may moderate the effectiveness of different types of written CF (direct and metalinguistic), and whether these differences have an effect on learners' subsequent output (immediate and delayed). This purpose was in response to calls from Bitchener and Storch (2016), Bitchener and Ferris (2012) and Ellis (2010) for an investigation into the moderating effect of individual factors on how learners respond to and use written CF. Understanding whether certain individual factors have the potential to moderate progress (that is, the extent to which they may have an impeding or facilitating effect on L2 development) helps researchers understand why some learners are able to develop their L2 knowledge more easily than others and why some learners are successful in learning from the feedback provided while others appear to not learn from it. Additionally, this type of study helps us to understand the relationship between cognitive abilities and treatment type so we know how different learners may benefit from different forms of instruction.

To date, very few studies have investigated the moderating effect of individual factors on the effectiveness of written CF on L2 development from a cognitive perspective. This study, therefore, investigated the moderating effect of working memory and phonological short term memory on the effectiveness of written CF types (namely, direct CF, metalinguistic explanation, direct CF plus revision and metalinguistic explanation plus

revision). Because it has been shown to impact on learning in the oral context (Goo, 2012; Li, 2013), it was important to see if it has a similar effect in the written context. This study, for the first time, has investigated the potential moderating effect of working memory and phonological short-term memory in written CF research.

This study employed a pre-test, treatment, post-test, and delayed post-test design, using intact English as a Foreign Language (EFL) classes. A pre-test-post-test design allows the researcher to assess the impact of the experimental manipulation by observing the difference between the pre-test and post-test.

One hundred Iranian EFL university students participated in the quantitative study and were assigned to one of the four previously mentioned experimental groups or the control group. The efficacy of one session of written CF on the use of the passive voice was investigated immediately and over time (four weeks). One week prior to the start of the CF treatment, participants performed a writing task as the pre-test. The immediate post-tests (revision and new test) were conducted immediately after the CF treatment session had been completed in Week 2. The delayed post-test was completed in Week 4. This quantitative study was designed to answer research questions 1-3. Additionally, participants completed a working memory test (reading span test) and a phonological short-term memory test (non-word span test) in Week 3 in order to address research question 4.

### **7.3 Summary of key findings**

#### *Research question 1*

RQ1 investigated the efficacy of direct CF and metalinguistic feedback, with and without revision, on learners' use of the English passive voice in a text written immediately after

the provision of written CF and in new texts over time (i.e., four weeks). The results showed that all experimental groups (direct CF, metalinguistic explanation, direct CF plus revision, metalinguistic explanation plus revision) outperformed the control group in both the short term and over time; however, the direct CF treatment was relatively more effective than other treatments in both cases. Additionally, the experimental groups significantly improved their accuracy from the pre-test to the immediate post-tests. Then, from the immediate to the delayed post-test the improvement deteriorated slightly, but the decrease in accuracy was not significant, which reveals the learners retained their improvement from the immediate to the delayed post-test. In addition, the improved accuracy in the delayed post-test was significantly higher than that in the pre-test. The results also showed that the accuracy for the control group revealed no significant differences from the pre-test to the immediate post-test and from the immediate post-test to the delayed post-test.

### *Research question 2*

In order to address research question 2, that is, investigating whether requiring the learners to do revision had any effect on the accuracy of new pieces of writing, the two revision groups (i.e., direct CF plus revision and metalinguistic explanation plus revision) and the two groups that did not make revisions (i.e., direct CF and metalinguistic explanation) were combined and compared. The difference between this study and the previous studies (e.g., Frear, 2012; Van Beuningen et al., 2008, 2012; Shintani et al., 2014) that investigated the effect of revision on improved accuracy is that in this study students did not have access to the text on which they had received feedback when they were doing the revision task. The findings showed that the revision groups (direct CF plus revision and metalinguistic plus revision) improved their accuracy from the pre-test to the immediate post-test. From the immediate post-test to the delayed post-test, the revision groups' accuracy deteriorated a little, but the change was not significant. In other words,



from the immediate to the delayed post-test, the revision groups retained their accuracy in using the passive voice. Similarly, the non-revision groups (direct CF and metalinguistic explanation) improved their accuracy from the pre-test to the immediate post-test. From the immediate post-test to the delayed post-test, this group's accuracy also deteriorated a little, but the change was not significant. This shows that the non-revision group also retained their accuracy after two weeks. From the immediate to the delayed post-test there was no significant change for the control group. Additionally, the results of the study showed that both the revision and non-revision groups had significantly higher improvements in accuracy than the control group in the immediate post-test and delayed post-test. However, the written CF for the non-revision groups (direct CF and metalinguistic explanation) was relatively more effective in the short term and the written CF for the revision groups (direct CF plus revision and metalinguistic plus revision) proved relatively more effective in the long term.

### *Research question 3*

Research question 3 examined the relative efficacy of direct CF and metalinguistic explanation regardless of whether there was an opportunity for revision. In order to address this question, the two direct CF groups (i.e., direct CF and direct CF plus revision) and the two metalinguistic explanation groups (i.e., metalinguistic explanation and metalinguistic explanation plus revision) were combined and compared. The findings showed that both combined groups (that is, the combined direct CF groups and the combined metalinguistic explanation groups) improved significantly from pre-test to immediate post-test. The accuracy deteriorated a little from the immediate to the delayed post-test, but the changes were not significant. In addition, the findings showed that the combined direct CF group improved its accuracy relatively more than the combined metalinguistic explanation group in both the immediate and delayed tests.

#### *Research question 4*

Research question 4 investigated the extent to which working memory and phonological short term memory may moderate the use of direct corrective feedback and metalinguistic explanation, with and without revision, in new writing texts over time. The findings revealed that working memory moderated the effect of the metalinguistic explanation group and the combined metalinguistic explanation group both immediately and over time (i.e. four weeks). Additionally, working memory moderated the effect of the direct CF plus revision group and the combined revision group (direct CF plus revision and metalinguistic explanation plus revision) only in the long-term (i.e. four weeks). The results also showed that phonological short term memory negatively moderated the effect of direct CF plus revision in only the long term (i.e. after two weeks).

### **7.4 Contributions of the study**

This section presents a discussion of the contributions. It firstly discusses the study's contribution to new knowledge, followed by a discussion of its contribution to theory, methodology and pedagogy.

#### **7.4.1 Contributions to new empirical knowledge**

Some of the findings of the study support the findings of other studies. In addition this study has added to existing findings.

The findings of this study corroborate previous studies that revealed that written CF may improve the short and long-term development of certain targeted grammatical features (e.g., Bitchener & Knoch, 2010a, 2010b; Sheen, 2007; Shintani et al., 2014). The findings of this study also support Shintani et al.'s (2014) study which found that direct CF was more effective than metalinguistic explanation. In both studies direct CF was more

explicit than metalinguistic explanation because the metalinguistic explanation group received feedback in the form of a handout and the errors were not identified in their written text.

As far as new knowledge, this thesis adds to our knowledge from existing studies by targeting a new complex structure, that is the English passive voice. To date, a number of the focused written CF studies have investigated the effectiveness of written CF on simple rule-based forms such as the English article (e.g., Bitchener, 2008; Bitchener & Knoch, 2009; Shintani & Ellis, 2013), the past tense (e.g., Rummel, 2014; Bitchener et al. 2005) and prepositions (e.g., Bitchener et al., 2005; Guo, 2015). However, as far as the researcher knows, there are only two studies that have targeted complex structures, that is, the present perfect (Rummel, 2014) and the hypothetical conditional (Shintani et al, 2014). This study is the first to target the English passive voice as a complex structure. The study found that written CF on the use of the English passive voice can increase linguistic accuracy in the short term and over time.

Additionally, when the combined revision (direct CF plus revision and metalinguistic explanation plus revision) and non-revision (direct CF and metalinguistic explanation) groups were compared, the results showed that improved accuracy was retained longer by the revision group than by the non-revision group. The difference between this study and previous studies that compared revision and non-revision groups (e.g., Frear, 2012; Van Beuningen et al., 2008, 2012; Shintani et al., 2014) is that this study was the first to not permit learners to have access to the text on which they had received feedback. Bitchener (2016) pointed out that when learners revise their text and have access to the corrections (e.g. direct CF), little or no cognitive processing may be needed. Thus, another contribution of the study is that providing learners with revision tasks, without allowing access to the text on which they received feedback, may improve their accuracy in the long term.

A further contribution of the study to empirical knowledge is that, for the first time, working memory and phonological short-term memory were investigated as potential moderating factors in written CF research. The moderating effect of individual differences has started to be investigated in written CF research (e.g., Rummel, 2014; Shintani & Ellis, 2015; Sheen, 2007a). However, prior to the current study, the potential moderating effect of working memory and phonological short-term memory had yet to be investigated in terms of their impact on the effectiveness of written CF. Oral CF studies have shown that working memory impacts on the process of L2 development (Goo, 2012; Li, 2013) and it has been argued that working memory may moderate the learners' use of written CF (Kormos, 2012). Thus, there has been a need to examine if working memory and phonological short-term memory, as cognitive factors, may also moderate the effectiveness of written CF. The findings of the study showed that phonological short term memory negatively moderated the effect of direct CF plus revision in only the long term. Besides, working memory moderated the effect of less explicit types of feedback, namely metalinguistic explanation and metalinguistic explanation plus revision, both immediately and over time. Additionally, working memory moderated direct CF plus revision for the combined revision groups (metalinguistic explanation plus revision and direct CF plus revision), but only in the long term (A discussion of the theoretical contribution of these findings is presented in Section 7.4.2.4 of this chapter).

#### **7.4.2 Contributions to theory**

This section reports on the extent to which this study validates existing cognitive theories as an explanation of performance as well as its contributions towards new knowledge about theory. It begins by discussing the contributions to our understanding of the value of explicit input in the form of written CF to the cognitive processing of written CF

(section 7.4.2.1). This is followed by a theoretical discussion of the study's contributions to the development of a complex linguistic structure (section 7.4.2.2) and the value of revision for improving the effectiveness of written CF for L2 development (section 7.4.2.3). Finally, section 7.4.2.4 discusses why working memory may impact on the use of written CF for L2 development.

#### **7.4.2.1 The value of explicit written input for L2 development**

The extent to which written CF as an explicit form of input plays a role in learners' acquisition of a target language has been a subject of inquiry for both researchers and teachers of second language writing and acquisition. For instance, questions remain as to whether focused written CF facilitates L2 development immediately and over time and whether certain types of written CF are more effective in improving accuracy than other types. This study has attempted to answer these questions and has subsequently validated existing theory which explains that an explicit form of input such as written CF can lead to the development of explicit knowledge.

The findings of this study validate aspects of Gass's (1997) cognitive processing model for the written context. As explained in Chapter 2, Gass's (1997) framework identifies how a single episode of explicit input processing may help learners to develop their explicit knowledge. Accordingly, if noticed and comprehended, explicit input can subsequently go through central processing (i.e., intake and integration) and result in output. Bitchener and Storch (2016) pointed out that even though the five stages were designed to discuss the cognitive processing of oral corrective feedback, the same stages can be broadly applied to the cognitive processing of written CF. The first stage explains that learners are required to notice or apperceive the explicit input. In other words, learners are required to notice the gap between the explicit corrective feedback input and their erroneous production in order to modify it accurately. Similarly, learners in the

written context are required to notice the explicit input in the same way as in the oral context. The second stage of information processing explains that input needs to be ‘comprehended’ by learners described by Gass. This stage is also required in the written context for explicit input to go through central processing. Written CF can also facilitate intake. At this stage, learners match the information provided by the input/CF with their existing explicit knowledge. In the writing context, as opposed to the speaking context, learners may have a better opportunity to match the input/written CF with their existing knowledge because of (1) the writing process which enables them to refer back as often as desired to their written text and the feedback provided on the text, and (2) having sufficient time to analyze and re-analyze their text. The next stage, *integration*, deals with “storage of new information for later use, hypothesis formulation, and confirmation or reformulation of existing hypotheses” (Gass, 1997, p. 25). In the context of writing, hypothesis-testing is facilitated by time, in that learners have sufficient time to make a cognitive comparison between the written CF and existing knowledge retrieved from their long-term memory.

Because this study showed how explicit written CF led to improved accurate output, it means that the feedback helped students to (1) notice the difference between what they had produced and the information provided by the CF and (2) produce improved accuracy of the targeted structure, the English passive voice, during output.

These findings of the study also validate skill acquisition theories for the written context. These theories posit that declarative knowledge must be processed with conscious attention (i.e., controlled processing) (McLaughlin, 1983) and that declarative knowledge must be proceduralized (Anderson, 2000). These theories present the stages (from controlled to automatic) that learners are required to pass through to acquire any skill, including a language skill. Skill acquisition theories maintain that intentional learning (e.g., by means of explicit written CF) can play an important role in the controlled stage,

when learners receive *explicit input* and draw on their procedural knowledge, and that such learning can ultimately result in more automatic processing of their procedural knowledge and, ultimately, their automatic knowledge. Thus, skill acquisition theories can be used to explain how learners may improve their L2 development of a target structure (e.g., English passive voice) immediately and over time. The findings of the study showed improved accuracy not only immediately, but also after two weeks. Even though the accuracy deteriorated slightly from the immediate to the delayed post-test, the decrease was not statistically significant. These findings may indicate that written CF raised the learners' conscious attention to the target structure and may have helped them to establish procedural knowledge of the structure in the immediate post-test. Because significant improvement in accuracy was also found in the delayed post-test, it seems that the learners retained that procedural knowledge and thus were able to retrieve it more effectively after two weeks. In other words, these findings show that one session of written CF treatment enabled the learners to process the declarative knowledge of the English passive voice in a controlled manner with conscious attention, after which they retained and proceduralized their knowledge. As the delayed post-test was conducted two weeks after the treatment session, whether their procedural knowledge was converted to automatize knowledge cannot be determined from this study.

Additionally, the absence of improvement by the control group suggests that the improved accuracy of the treatment groups was the result of the explicit written CF provided. These findings also support those of other studies (e.g., Bitchener & Knoch, 2010a; Rummel, 2014; Sheen 2007; Stefanou & Revesz 2015) and refute Krashen's (1985) and Truscott's (1996) theories that explicit knowledge only results in superficial learning and cannot facilitate L2 development over time.

The findings of this study also support existing theory about how more explicit types of feedback may be more likely to draw learners' (especially those with a lower level of

proficiency) attention to the target structure than less explicit types of corrective feedback. In the written CF literature, there have been theoretical arguments on the efficacy of different types of written CF on L2 development. Those who favour of the effectiveness of less explicit types of written CF suggest that it is most useful because learners can engage learners in problem solving and guided learning (Lalande, 1982). Those supporting more explicit types of written CF suggest that they may reduce learners' confusion in understanding the feedback and offer more explicit feedback on the hypotheses being tested.

Additionally, written CF types may play different roles in terms of getting the learners attention and helping them to comprehend written CF. This is because different types of written CF differ in their degree of explicitness and as a result may draw learners' attention to different levels of cognitive processing. For instance, at the stage of knowledge modification, learners (especially those with a lower level of proficiency) may notice and then comprehend more explicit types of written CF, resulting in improved accuracy. Schmidt (1994, 2001), by drawing on Tomlin and Villa's (1994) attention theory, explained that learners have three levels of attention: alertness, orientation and detection, and believed that detection (the highest level of attention) is necessary for further processing of input and subsequent learning to take place. More explicit types of written CF may be more likely to draw the highest level of attention of a learner because the target linguistic structure is clearly or explicitly stated, and thus, they may process it further. Therefore, more explicit types of written CF, because they may raise the highest level of attention to explicit input and provide additional linguistic information, may result in a new, accurate hypothesis and thus may lead to accurate output. However, if less explicit types of written CF are provided, the learners may not be able to form a correct hypothesis and, consequently, the result may not be immediate improved accuracy.



Although less explicit types of written CF may not lead learners (especially those with a lower level of proficiency) to form a correct hypothesis, they may enable learners to store the information and wait for more input. At the stage of knowledge consolidation, learners (especially those at an advanced level) may benefit from less explicit types of written CF because they are able to retrieve the relevant explicit knowledge from their long-term memory. Consequently, learners may undertake deeper controlled processing because the feedback is not as clearly explained (Bitchener, 2012).

The findings of RQ1 and RQ3 showed that even though both direct CF and metalinguistic explanation were effective immediately and after two weeks, direct CF was relatively more effective than metalinguistic explanation in improving the accurate use of a complex structure (i.e., the English passive voice). As stated before, in this study, the degree of explicitness of direct CF was more than the degree of explicitness of metalinguistic explanation. In other words, for direct CF learners were explicitly provided with the correct form of their errors, while the metalinguistic explanation groups were only provided with a handout that explicitly explained and exemplified the target structure (i.e., the English passive voice), meaning the errors were not identified for this group. This finding supports DeKeyser's (2003) recommendation that arbitrary form-function mappings will need to be explained more explicitly. In other words, when learners cannot depend on a clear rule when attempting to understand a target structure, less explicit types of written CF might not be explicit enough to be beneficial. Similar findings regarding the efficacy of more explicit types of feedback in drawing learners' attention to the target structure in comparison to less explicit types add further support to this theory (e.g., Shintani et al., 2014; Van Beuningen et al., 2008; Guo, 2015). However, as explained above, less explicit types of written CF may be sufficient for advanced learners because they are able to retrieve the relevant explicit knowledge from their long-term memory.

#### **7.4.2.2 The value of revision for improving the effectiveness of written CF**

This section discusses the extent to which existing theory about revision following feedback can help learners improve accuracy in their use of the target structure.

Theoretically, providing learners with an opportunity to revise a text plays an important role in the development process. Asking learners to perform revision tasks may foster learning because “producing the correct form may help learners automatize their production” (Loewen, 2004, p. 157). In other words, when learners are revising their text with no access to the text on which they received feedback (as was the case in this study), they retrieve information from their long-term memory and this helps facilitate the consolidation of L2 knowledge. In addition, revision can also be argued from a skill-learning perspective because modifying output through revising and correcting an initial draft can provide the practice required for the proceduralization of explicit knowledge (Frear, 2012).

Additionally, it can be argued that regardless of whether there is a single opportunity to revise a text following written CF (Frear, 2012; Van Beuningen et al., 2008) or multiple opportunities (Chandler, 2003; Hartshorn et al., 2010), revision following feedback leads to greater accuracy in new writing. This is because revision following written CF leads to ‘pushed output’ (Shintani et al., 2014). Swain (1995) argued that when learners are pushed to produce language, they are likely to “notice a gap between what they want to say and what they can say, leading them to recognize what they do not know, or know only partially” (pp. 125-126). Swain (1995) hypothesized that pushed output helps learners to notice the grammatical forms that otherwise may go unattended. Thus, revision following feedback pushes learners to attend to the feedback they have been provided with and to process it across the cognitive processing stages as determined by Gass (1997), before making a hypothesis on an accurate modification.

This study found that even though both the revision and the non-revision groups improved their accuracy from the immediate to the delayed post-tests, the knowledge gained from written CF when its provision was followed by revision appeared to be retained longer than when no opportunity to revise after receiving written CF was provided. The finding suggests that the extra attention learners were required to give to the feedback by revising their work may have resulted in pushed output because learners noticed the gap.

Although in a number of studies learners did not undertake revision, they still improved their accuracy in both the immediate and delayed post-tests (e.g., Bitchener, 2008; Rummel, 2014; Sheen, 2007). However, it can be argued that in these studies, the absence of a revision group provided no opportunity to determine whether a revision group could achieve greater accuracy than other experimental groups. Both this study and a number of previous studies (Chandler, 2003; Frear, 2012; Hartshorn et al., 2010; Shintani et al., 2014; Van Beuningen et al., 2008, 2012) compared the improved accuracy of revision and non-revision groups in their use of the target structure and found that even though non-revision groups improved their accuracy, revision groups retained accuracy longer than non-revision groups. The lack of revision groups in many other written CF studies provided no opportunity to identify whether revision groups would be able to improve their accuracy to a greater extent than non-revision groups or vice versa.

#### **7.4.2.3 Contribution of written CF to the development of a complex linguistic structure**

This section presents a theoretical discussion of the study's contribution to the development of a complex linguistic structure (i.e., the English passive voice).

Even though there is a growing body of evidence that written CF does improve accuracy over time, there has been limited research to investigate the effectiveness of written CF with specific linguistic errors such as those involving complex linguistic structures.

Yang and Lyster (2010) categorized linguistic error types into rule-based and item-based errors. Rule-based errors are those that occur in a rule-governed way and can be corrected (for example, by referring to grammar books) while item-based errors are those for which there are no set of rules students can refer to or those for which the rules vary according to the way the structure is used in different linguistic environments. For instance, the rules for employing the simple past tense are relatively straightforward; however, the rules for some structures such as the passive voice can be difficult to acquire as they involve complex structures, which means that learners need to attend to and process a number of components that make up the complex structure. There are two components to the passive voice structure, that is, the verb “to be” and the past participle. The verb “to be” is rule-based as there is a rule that identifies when learners need to use “am”, “is” or “are”. However, the past participle is item-based as it takes three different forms: regular verbs that end in “ed” (e.g., “opened”); irregular verbs in which the past participle changes form (e.g., “lost”); and irregular verbs where the past participle retains the same form (e.g., “read”). Thus, structures such as the passive voice that are both rule-based and item-based are likely to require learners to employ more attentional capacity when using them.

There are theoretical arguments concerning the extent to which written CF can effectively improve different types of linguistic errors. The reason for this is that morphological, syntactic, and lexical errors represent different domains of linguistic knowledge (Ellis, 2008; Ortega, 2009) and learners may need to focus their attention on more than just one linguistic element each time they make a hypothesis on the use of the correct linguistic form for a certain linguistic error type. Additionally, it has been argued (Young, 1996) that one linguistic form/structure may be more difficult to learn than another. Thus, some linguistic forms /structures may be more ‘treatable’ than others (Ferris, 1999). Therefore, as a result of written CF, some complex structures may develop more easily than others. However, Truscott (2007) argued that written CF can only treat errors that “are relatively

simple and can be treated as a discrete item” (p. 258) such as non-grammatical errors. Additionally, he argued that when learners are not confident, they avoid using complex structures; moreover, because the knowledge of complex structures is more than a collection of discrete items, written CF may not facilitate the development of knowledge of these structures. However, avoiding complex structures does not mean that written CF cannot facilitate improved accuracy in the use of complex structures. Additionally, when treating complex structures, there are several factors such as the type of feedback and the learners’ level of proficiency that may impact on the effectiveness of written CF in developing knowledge of complex structures. For instance, structures such as the English passive voice that are both semantically and syntactically complex require learners to use more attentional capacity when employing them. When targeting complex structures, less explicit types of feedback are unlikely to be beneficial for low proficiency learners; rather, such learners require more explicit types of written CF that explicitly explain and illustrate the complex structures. Thus, it may be that for some learners more explicit feedback may need to be provided on complex structure errors such as the English passive voice and the hypothetical conditional. However, less explicit types of written CF on complex structures may be effective for advanced learners, because they may have some knowledge of the complex structure and can retrieve it from their long-term memory.

A number of studies have indicated that written CF can improve the accuracy of simple rule-based categories such as English articles and the past tense (e.g., Frear, 2012; Sheen, 2007a). To the best of my knowledge, there are only two studies that have focused on a complex structure: Shintani et al. (2014) and Rummel (2014) targeted the hypothetical conditional and the present perfect tense respectively. They are complex structures because the hypothetical conditional comprises seven components (i.e. (1) the past tense, (2) the perfect aspect, (3) the past participle (PP) form in the *if* clause, (4) the modal, (5) the past tense, (6) the perfect aspect and (7) the PP form in the main clause) and the

present perfect tense includes two components (i.e., the verb “to be” and the past participle). Rummel (2014) found that written CF was effective for improving the accurate use of the present perfect tense immediately and over time (seven weeks). However, Shintani et al. (2014) reported that learners who received metalinguistic explanation did not sustain improved accuracy in the use of the hypothetical conditional over two weeks.

The findings of the current study showed that different types of written CF (direct CF, direct CF plus revision, metalinguistic explanation, metalinguistic explanation plus revision) were effective in terms of targeting a complex structure, the English passive voice, immediately and after two weeks. From the immediate to the delayed post-test the improvement atrophied slightly; however, the decrease in accuracy was not statistically significant. This showed that some learners retained the improvement from the immediate to the delayed post-test. Furthermore, the accuracy in the delayed post-test was significantly higher than that of the pre-test. Additionally, the accuracy rate for the control group revealed no significant improvement from the pre-test to the immediate post-test and from the immediate post-test to the delayed post-test. Thus, these findings revealed that written CF not only improves accuracy in the use of rule based errors such as the simple past tense ( e.g., Frear, 2012; Sheen, 2007a.), but it can also improve the accurate use of complex structures such as the English passive voice. These findings also support Rummel’s findings on the effectiveness of written CF in the use of complex structures and refute Truscott’s (2007) claim that written CF cannot facilitate the development of complex structures. However, it is important that future research investigates the effectiveness of written CF in relation to different types of complex structures.

#### **7.4.2.4 The moderating effect of working memory on the use of written CF for L2 development**

As the first study that investigated the moderating effect of working memory in written CF, this study contributes to new knowledge about theory. This section discusses why working memory may moderate the different explicit types of written CF differently.

Individual factors have been recognized as important in the process of L2 development (Ellis, 2010; Bitchener & Ferris, 2012; Bichener & Storch, 2016). It has been argued that the stages of cognitive processing of a single episode of input may either be facilitated or impeded by the presence of individual cognitive factors. One such factor may be the learners' working memory (Ellis, 2010; Bitchener, 2012, 2016). Because oral CF studies have shown that working memory moderates the process of L2 development (Li, 2013) there was a need to investigate whether working memory, as a cognitive factor, may moderate the effectiveness of written CF (Bichener & Storch, 2016). Kormos (2012) highlighted the importance of working memory in the written context. She pointed out that working memory may moderate how learners learn from different types of written feedback. She explained that in contrast to the oral context, learning opportunities through feedback in the written context are less constrained by time pressure; however, she stated that because learners are dependent on their working memory capacity while writing, they may respond differently to feedback. Thus, for the first time, this study investigated the extent to which working memory may moderate learners' use of different types of written CF (direct CF and metalinguistic explanation). In this study, there were differing degrees of explicitness in the types of feedback, with the direct CF being more explicit than the metalinguistic explanation as explained in previous sections.

Working memory is a cognitive device with the dual function of storing and processing information. Working memory is operationalized as either phonological short-term memory, which only consists of the storage component, or working memory capacity,

which refers to both storage and processing components. These components of working memory contribute to learning effects: attention control (noticing), storage of information, simultaneous storage and processing of information, and retrieval of information from long-term memory.

Working memory capacity impacts on the extent to which input is noticed. Learners with a high working memory capacity are better at noticing than those with a low working memory capacity (Kane & Engle, 2003). Robinson (1995) defined “noticing” as a mechanism of “detection plus rehearsal in short-term memory” (p. 296). He stated that detection without awareness is insufficient; moreover, rehearsal in working memory results in a level of awareness that is essential for L2 development. Working memory capacity also involves the simultaneous storage and processing of information; however, the working memory system has a limited capacity, resulting in a trade-off between processing and storage (Shah & Miyake, 1999). Working memory as a unitary construct assumes that both storage and processing make use of the same resources (Just & Carpenter, 1992). Thus, if a task requires a learner to focus mostly on the processing of information, there are less resources for the storage of the information and vice versa. Working memory can also contribute to the retrieval of information from long-term memory. In other words, learners can recall the information they have already stored in long-term memory through working memory.

Based on the capacity-limited model (e.g., Skehan, 1998), working memory has limited capacity and can process a limited amount of information at one time; therefore, learners with greater working memory capacity can better attend to and process input/CF (Ortega, 2009). Additionally, the model explains that learners with a lower level of proficiency might find it difficult to simultaneously attend to more than one aspect of language (e.g., meaning and form). Because learners with a lower level of proficiency need to process new knowledge consciously, they need to pay greater attention to and put more effort into



processing new information in their working memory. Although written CF allows more time to process information than oral CF, a learner's working memory still needs to coordinate the learner's attention to cognitive processing. On each new occasion during the consolidation stage, after written CF has been given and accurate output has been produced, learners draw upon their working memory to retrieve the new information from their long-term memory and produce it. Over time, when new knowledge is proceduralized, the learner may put less effort into the working memory to process new information.

Additionally, different types of explicit written CF may reflect two different mechanisms of noticing. The higher level of explicitness of corrective feedback might lead to learners noticing the target structure with less cognitive demand compared to a less explicit type of written CF that does not directly draw learners' attention to form when writing new texts immediately and over time. The noticing of CF is an attention demanding task and thus requires a domain-general, attention control mechanism considered as a critical component of working memory capacity (e.g., Goo, 2013; Engle, 2002; Kane et al., 2007). This cognitive control mechanism is not needed for the noticing of more explicit types of CF. This study found that working memory moderated the effect of the less explicit type of feedback (i.e., metalinguistic explanation) but not that of the more explicit one (i.e., direct CF). As discussed above, this finding shows that the degree of explicitness of written CF may determine the extent to which working memory has a moderating effect on written CF. This suggests that the explicit nature of direct CF may minimize learners' individual differences in working memory. Thus, these findings reveal that the feedback types of direct CF and metalinguistic explanation may have a different impact on the moderating effect of working memory on written CF as they have different levels of explicitness.

Additionally, this study explored the potentially moderating effect of phonological short-term memory on direct CF and metalinguistic explanation with and without revision. The results of this study revealed that phonological short-term memory negatively moderates the effectiveness of direct CF plus revision in the long term. A possible reason for this finding is that different models of working memory agree that the entire working memory system has a limited capacity, resulting in a trade-off between processing and storage (Shah & Miyake, 1999). Thus, a possible explanation for the finding that phonological short-term memory moderates the effectiveness of direct CF plus revision in the long term is that when the direct CF plus revision group carried out the revision task, they had no access to the text on which they had received feedback. Therefore, they may have deeply focused on the processing of the feedback, resulting in improved accuracy over time rather than utilising storage which is essential in phonological short-term memory. This mechanism resulted in a trade-off between process and storage. Additionally, the findings that those with less phonological short-term memory benefitted more from feedback can be explained based on the 'less is more' hypothesis (Miyake & Friedman, 1998; Newport, 1990). Miyake and Friedman (1998) pointed out that researchers need to consider that adults' small phonological short-term memory capacity is actually great. Thus, the findings of this study showed that learners with less phonological short-term memory capacity who received direct CF plus revision were able to attend to the details of the target structure, that is, the English passive voice.

Therefore, this study found that phonological short-term memory moderates the effectiveness of direct CF plus revision in the long term but does not moderate the effectiveness of other types of written CF as used in this study, that is, direct CF, metalinguistic explanation, and metalinguistic explanation plus revision. This reveals that the learners' differences in their phonological short-term memory are not a main factor in

their improved accuracy in written CF. However, as this is quite possibly the first study of its kind, further studies are required before coming to a conclusion.

Oral CF studies previously found that working memory moderates the effect of less explicit types of oral CF (Goo, 2012). Similarly, this study adds new knowledge about theory by finding that working memory moderates the effect of less explicit types of feedback in written CF.

### **7.4.3 Contributions to research methodology**

This section discusses the contributions of this research to the understanding of methodological approaches that may help further advance our knowledge of the role of written CF in L2 development. There are two new contributions: first, providing the revision group with no access to the text on which they received feedback when revising or writing new texts and second, using a reading span test through the DMDX system to measure the learners' working memory. The contributions are discussed as follows.

The first new contribution of the study is related to its design. In order to measure the effectiveness of revision following feedback on improved accuracy, the learners, for the first time, revised their texts without access to the text on which they had received feedback. They had 10 minutes to study the feedback on their pre-test before conducting the revision task. This is the main difference between this study and previous studies that have compared revision and non-revision groups (e.g., Frear, 2012; Van Beuningen et al., 2008, 2012; Shintani et al., 2014). Bitchener (2016) pointed out that when learners revise their text and have access to the corrections (e.g. direct CF) little or no cognitive processing may be needed. Shintani et al. (2014) also suggested that revision may lead to pushed output, especially if learners have no access to the text on which they received

feedback. Thus, when learners revise their text with no access to the corrections, they appear to undertake more cognitive processing which leads to pushed output.

The second contribution to methodology is related to the measurement of learners' working memory. Working memory is a cognitive device with the dual function of storing and processing information. This is the first study, to the best of my knowledge, to utilize a reading span test through the DMDX system to measure participants' working memory in either the oral or the written context. DMDX is a Win 32-based display system employed to measure reaction times to auditory and visual stimuli. Accordingly, participants read a set of unrelated sentences and judged whether they made sense or were nonsense (processing assessment). Their reaction time during the judgment was measured. They were then asked to recall the final word of each sentence at the end of the set (storage assessment).

Several studies have investigated the relationship between complex working memory and the efficacy of recasts in oral CF studies (Li, 2013; Goo, 2012; Mackey et al., 2002; Re'v'e'sz, 2012; Sagarra, 2007; Trofimovich et al., 2007). However, one of the main limitations of the above studies (except for that of Li [2013]) is that the tests they employed only measured the storage component and did not measure reaction time and veracity judgment (indicators of the processing component of working memory) (Li, 2013). As far as I am aware, only Li (2013) attempted to remove the limitations of the above studies by using the DMDX system.

This study used the DMDX system to measure both the storage and processing of information as well as reaction time. This helped to measure the learners' working memory more accurately than had been done in previous oral CF studies, which in turn impacted on the reliability of the test. However, the difference between this study and Li's study is that he used a listening span test and this study, for the first time, used a

reading span test. This measurement of working memory is not only applicable to written CF studies, but to oral CF studies as well.

Thus, as the findings of this study have shown, revising a task without having access to the text on which the learners received feedback helped them to improve their accuracy. Additionally, using the DMDX system improved the reliability of the test.

#### **7.4.4 Contributions to pedagogy**

This section reports the contributions this study makes to pedagogy, including how the findings of this study may encourage teachers to provide focused written CF to facilitate learners' L2 development.

The findings of this study showed that written CF improved learners' accuracy of the target structure (the English passive voice) immediately and after two weeks. By drawing on the findings of the present study, the most important observation is that L2 teachers can use one single written CF treatment to help promote learners' L2 development not only in the short term, but also over time.

In this study the participants were at an intermediate level and the findings showed that a more explicit type of written CF (e.g., direct CF) was more effective than a less explicit type of written CF (e.g., metalinguistic explanation as used in this study) on the complex structure of the English passive voice. A more explicit type of feedback could be especially useful for those learners who are at a lower level of proficiency because it may reduce confusion about the feedback that they may experience if they do not understand the less explicit types of CF, especially on complex structures. This is because more explicit types of feedback explicitly reveal learners' errors and may help them to more clearly and fully comprehend their errors than less explicit CF. Additionally, the learners who are at a lower level of proficiency are less likely to have a more developed long-term

memory than learners at a higher level of proficiency (Ortega, 2009). Thus, when lower level learners receive written CF on a complex structure, they may not already have the required level of knowledge of that structure in their long term memory to draw upon to understand the provided feedback. Thus, for learners at lower levels of proficiency, a more explicit type of written CF may be more likely to lead to accurate output than less explicit types of written CF. It is therefore recommended that teachers of those learners who are not at an advanced level of proficiency use more explicit types of written CF when providing feedback on complex structures rather than less explicit types of written CF (e.g., metalinguistic explanation as used in this study). However, if metalinguistic explanation is in a form that explicitly explains and provides instances of the target structure, then it may be more effective for lower proficiency learners because it explicitly explains their error and how they can correct it.

Additionally, teachers should try to identify the best way to give feedback on their students' writing. For instance, providing written CF on individual learners' errors may be time consuming for some teachers. However, providing metalinguistic explanation in the form of a handout on one error was also found to positively affect learners' accuracy immediately and over time; therefore, it could prove successful for teachers to use metalinguistic explanation in the form of a handout on one target structure. It is especially recommended for advanced learners because, as it was discussed above, they may have a more developed long term memory that can be drawn upon to understand the target language. These handouts, once prepared, can be used repeatedly when teachers need them.

Another contribution of the study is that it showed that the knowledge gained from written CF when followed by revision may be retained longer than when learners are not provided with an opportunity to revise after receiving written CF. Thus, by asking students to revise, teachers can help them to develop accuracy, leading to "pushed output" and the

triggering of noticing and hypothesis testing (Swain, 1995), which may help learners to notice the grammatical forms that otherwise may go unattended. The finding suggests that the extra attention learners were required to give to the feedback by revising their work may have resulted in pushed output because learners noticed the gap.

This study also found that working memory may have a moderating effect on less explicit types of feedback because the noticing of less explicit types of CF (e.g., metalinguistic explanation as used in this study) is an attention-demanding task and thus requires the domain-general, attention control mechanism considered to be a critical component of working memory capacity. Thus, by using more explicit types of written CF (e.g., direct CF) a teacher can minimize the moderating effect of individual differences in working memory.

### **7.5 Limitations of the study**

Even though the present study has achieved its aims, a number of limitations and shortcomings are acknowledged. These limitations are related to certain weaknesses and the scope of the study. They are discussed in detail below.

In quasi-experimental research such as the present study, it is often necessary to employ intact classes and this may bring some limitations. The first limitation concerns the sample size; while acceptable, it was smaller than the size that had been originally planned. Initially, 135 participants were recruited. However, as 35 of them did not complete all the tasks, their data were excluded. Even though 100 students participated in the study, it included five groups, so only 20 students were available for each group. Cohen (1988) reported that "depending upon the statistics in question, and the specific statistical model on which the test is based, reliability [i.e., precision] may or may not be directly dependent upon the unit of measurement, the population value, and the shape of

the population distribution. However, it is always dependent upon the size of the sample" (p. 6). Additionally, Brown (2006) suggested that when the number of participants increases, the standard error decreases, or becomes more precise (i.e., reliable). Thus, to increase reliability, a larger number of test-taking participants should be aimed for in future studies.

In addition, the participants of this study were from one proficiency level. Participants with different proficiency levels may have different levels of understanding of the English passive voice, which may result in different levels of effectiveness of written CF. For example, more advanced learners may be more successful in using less explicit types of written CF than learners with lower levels of proficiency. In this study the participants were at an intermediate level and the findings showed that more explicit types of feedback were more effective than less explicit types. However, less explicit types of feedback may be more effective for advanced learners because they may have existing knowledge of the target structure and can easily retrieve it from their long term memory. Thus, these findings may not be applicable to different proficiency levels.

A further limitation concerns the use of the same type of writing tasks, reconstruction tasks, for carrying out the pre-test, immediate post-test and delayed post-test. The tasks deal with a number of stages that are in time order. Accordingly, it is necessary to start at the beginning and describe each stage through to the last one. In this study, reconstruction tasks were used to encourage participants to use the target structure. For example, the participants reconstructed the process of making chocolate. While participants showed improved accuracy when undertaking this type of writing task, it is not clear if they would show the same improvement when writing in other genres. The tasks used in this study require reconstruction, which is different from other genres such as free writing. In other words, the nature of reconstruction tasks may create a context for the participants to use



the target structure more accurately than in free writing, as the requirement to reconstruct a text could act as scaffolding for the learners.

The duration of this research may also be seen as a limitation. The study was conducted over four weeks and the delayed post-test was conducted two weeks after the immediate post-test. The findings showed statistically significant improved accuracy for the experimental groups over four weeks. However, it is unclear if learners would have been able to retain this accuracy over a longer period of time.

Furthermore, the results of the ANOVA on the pre-test showed that the scores of one of the groups was significantly different from the others. Thus, gain scores were used to control the impact of other factors. In experimental studies, as covariate adjustment for a pre-test can cause biased results, it is recommended to employ gain scores because they are an unbiased estimate of true change (Rogosa, 1988). It is possible the placement test (an internal English placement test) used in the study was unable to measure the learners' proficiency level accurately. Thus, for future studies, it is recommended that a more recognized international placement test be used such as *Cambridge English: Proficiency (CPE)*. This may help to measure the learners' proficiency level more accurately.

There is also a limitation in the scope of the study. This study investigated the effectiveness of written CF on the learners' improved accuracy (i.e., the learners' output, the product). However, the study did not investigate the processes of L2 development (e.g., noticing, hypothesis testing etc.) that may underlie the effectiveness of written CF (Stefanou & Revesz 2015; Swain & Lapkin, 2002). Understanding the processes involved in L2 development is important, as it will help us understand more about the cognitive stages that need to be traversed before output is provided. If any of the stages considered important for output are not reached by learners, it may help us to find out why they failed to benefit from the feedback (Bitchener, 2017).

## **7.6 Recommendations for future research**

This section presents recommendations for further research that may help to address the limitations of this study, along with potential avenues for research development based on the findings of this study.

As generalizability is one of the limitations of the study, it is important that the findings of this study regarding the effectiveness of written CF be confirmed in different contexts. There is a great focus on English grammar in the context of Iran that may have led to the result that direct CF was explicit enough for them to improve their accuracy on a complex structure. Because they might have some knowledge of the passive voice in their long term memory that they drew upon to understand the English passive voice. However, in other contexts where there is less focus on grammar, maybe learners need the type of feedback that is more explicit than direct CF. For instance, metalinguistic explanation that identifies the errors and explains how they can correct their errors. This type of metalinguistic explanation is different from that used in this study.

It is also suggested that for future research, delayed post-tests be conducted after a longer time period than the two weeks that was used in the present study. This study revealed that the treatment groups significantly improved their accuracy in the use of the English passive voice after two weeks; however, a longer time period (e.g. 6 months) before conducting delayed post-tests would provide an opportunity to determine if participants were in the process of consolidating the English passive voice and if they were proceduralizing it.

This study only examined the effectiveness of a single written CF episode on the use of the English passive voice. However, the findings do not show how effective written CF would be if the learners who did not benefit from one written CF episode were provided with more written CF episodes. Thus, it is recommended that future studies investigate

whether learners who do not improve their accuracy of the target structure in the immediate or delayed post-tests may benefit from written CF when they are provided with more than one written CF treatment session.

This study is the first to investigate the extent to which working memory may moderate the effects of different types of written CF. Therefore, it is important to carry out further investigations on the extent to which working memory may moderate other types of written CF. This study found that working memory moderates the effectiveness of a less explicit type of written CF (i.e., metalinguistic explanation as used in this study). Thus, it is recommended that future researchers investigate the moderating effect of the other types of written CF such as indirect CF.

An investigation of the moderating effect of other individual differences, such as motivation, is also suggested as individual differences have been recognized as an important factor impacting the processes of L2 development. Understanding whether certain individual factors have the potential to moderate progress (i.e., the extent to which they may have an impeding or facilitating effect on L2 development) will help researchers understand why some learners are able to develop their L2 knowledge more easily than others and why some learners are successful in learning from the feedback provided while others are not (Bitchener & Ferris, 2012; Bitchener & Storch, 2016).

## References

- Andersen, R. W. (1983). Transfer to somewhere. In S. M. Gass and L. Selinker (Eds), *Language learning* (pp.177-201). Rowley, MA: Newbury House.
- Anderson, J. R. (2000). *Cognitive psychology and its implications* (5th ed.). New York: Worth Publishers.
- Ashwell, T. (2000). Patterns of teacher response to student writing in a multiple-draft composition classroom: Is content feedback followed by form feedback the best method? *Journal of Second Language Writing*, 9(3), 227-257. doi:10.1016/S1060-3743(00)00027-8
- Bachman, L. F. (1995). *An investigation into the comparability of two tests of English as a foreign language* (Vol. 1). New York: Cambridge University Press.
- Baddeley, A. (2000). The episodic buffer: a new component of working memory? *Trends in cognitive sciences*, 4(11), 417-423. doi: 10.1016/ S1364-6 6 1 3 (00)01538-2
- Baddeley, A. (2003). Working memory: looking back and looking forward. *Nature reviews neuroscience*, 4(10), 829-839. doi: 10.1038/nrn1201
- Baddeley, A. D. (1986). *Working memory*. New York: Oxford University Press.
- Baddeley, A. D., & Hitch, G. (1974). Working memory. *Psychology of learning and motivation*, 8, 47-89. [http://dx.doi.org/10.1016/S0079-7421\(08\)60452-1](http://dx.doi.org/10.1016/S0079-7421(08)60452-1)
- Baddeley, A. D., & Hitch, G. (1993). The recency effect: Implicit learning with explicit retrieval? *Memory & Cognition*, 21(2), 146-155. doi:10.3758/ BF 0 3202726
- Baddeley, A.D., Logie, R.H. (1999). Working Memory: The multiple-component model. In A. Miyake & P. Shah (Eds) *Models of Working Memory* (pp 28-61). New York: Cambridge University Press.
- Baddeley, A., Gathercole, S., & Papagno, C. (1998). The phonological loop as a language learning device. *Psychological review*, 105(1), 158. [http://dx. doi. org/10.1037/0033 295X. 105.1.158](http://dx.doi.org/10.1037/0033-295X.105.1.158)
- Baralt, M. (2015). Working memory capacity, cognitive complexity and L2 recasts in online language teaching. In Wen, Z., Borges Mota, M., McNeill, A. (2015). *Working memory in second language acquisition and processing* (pp 248-269). Bristol: Multilingual Matters
- Birjandi, P., Maftoon, P., & Rahemi, J. (2011). VanPatten's processing instruction: Links to the acquisition of the English passive structure by Iranian EFL learners. *European Journal of Scientific Research*, 64(4), 598-609.

- Bitchener, J. (2008). Evidence in support of written corrective feedback. *Journal of Second Language Writing*, 17 (2), 102–118. doi:10.1016/j.jslw.2007.11.004
- Bitchener, J. (2012). A reflection on ‘the language learning potential’ of written CF. *Journal of Second Language Writing*, 21 (2) 348–363. doi:10.1016/j.jslw.2012.09.006
- Bitchener, J. (2016). To what extent has the published written CF research aided our understanding of its potential for L2 development? *ITL International Journal of Applied Linguistics*, 167(2), 111-131. doi: 10.1075/itl.167.2.01bit.
- Bitchener, J. (2017). Why some L2 learners fail to benefit from written corrective feedback. In Nassaji, H., & Karchava, E. (Eds.), *Corrective feedback in second language teaching and learning* (pp. 129-140). New York: Routledge.
- Bitchener, J., & Ferris, D. R. (2012). *Written corrective feedback in second language acquisition and writing*. New York, NY: Routledge.
- Bitchener, J., & Knoch, U. (2008a). The value of written corrective feedback for migrant and international students. *Language Teaching Research Journal*, 12 (3), 409–431.
- Bitchener, J., & Knoch, U. (2009). The value of a focused approach to written corrective feedback. *ELT Journal*, 63(3), 204–211. <https://doi.org/10.1093/elt/ccn043>
- Bitchener, J., & Knoch, U. (2010a). The contribution of written corrective feedback to language development: A ten month investigation. *Applied Linguistics*, 31(2), 193–214. doi: 10.1093/applin/amp016
- Bitchener, J., & Knoch, U. (2010b). Raising the linguistic accuracy level of advanced L2 writers with written corrective feedback. *Journal of Second Language Writing*, 19(4), 207–217. doi: 10.1016/j.jslw.2010.10.002
- Bitchener, J., & Storch, N. (2016). *Written Corrective Feedback for L2 Development*. Bristol, UK: Multilingual Matters.
- 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(3), 191-205. doi:10.1016/j.jslw.2005.08.001
- Brown, R.B. (2006). *Doing Your Dissertation in Business and Management: The Reality of Research and Writing*. London: Sage Publication
- Bruton, A., (2009). Designing research into the effect of error correction in L2 writing: not so straightforward. *Journal of Second Language Writing* 18 (2), 136-140. <https://doi.org/10.1016/j.system.2010.07.001>

- Carroll, J. B. (1981). Twenty-five years of research on foreign language aptitude. In K. C. Diller (Ed.), *Individual differences and universals in language learning aptitude* (pp. 83-118). Amsterdam: John Benjamins Publication CO.
- Celce-Murcia, M., Larsen-Freeman, D., & Williams, H. A. (1999). *The grammar book: An ESL/EFL teacher's course* (2nd ed.). Boston, UK: Heinle & Heinle.
- Handler, J. (2003). The efficacy of various kinds of error feedback for improvement in the accuracy and fluency of L2 student writing. *Journal of Second Language Writing*, 12(3), 267-296. doi:10.1016/S1060-3743(03)00038-9
- Chun, D. M., & Payne, J. S. (2004). What makes students click: Working memory and look-up behaviour. *System*, 32(4), 481-503. doi: 10.1016/j.system.2004.09.008
- Creswell, J. W. & Clark, V. L. P. (2011). *Designing and conducting mixed methods research*. London: SAGE Publication Ltd.
- Creswell, J. W. (2013). *Qualitative inquiry & research design: Choosing among five approaches* (3rd ed.). Thousand Oaks, Calif: SAGE Publications.
- Crocker, R. A. (2009). An introduction to qualitative research. In J. Heigham & R. A. Crocker (Eds.), *Qualitative research in applied linguistics: A practical introduction* (pp. 3-24). New York: Palgrave Macmillan.
- Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of verbal learning and verbal behavior*, 19(4), 450-466. doi: 10.1016/S0022-5371(80)90312-6
- Daneman, M., & Merikle, P. M. (1996). Working memory and language comprehension: A meta-analysis. *Psychonomic Bulletin & Review*, 3(4), 422-433. doi: 10.3758 /BF 032 4546
- DeKeyser, R. (1998). Beyond focus on form: Cognitive perspectives on learning and practicing second language grammar. In C. Doughty & M.H. Long (Eds.), *Focus on form in classroom Second language acquisition* (pp.42-63). New York: Cambridge university press.
- DeKeyser, R. (Ed.). (2007). *Practice in a second language: Perspectives from applied linguistics and cognitive psychology*. New York: Cambridge University Press.
- Dekeyser, R. (1997). Beyond explicit rule learning: Automatizing second language morphosyntax. *Studies in Second Language Acquisition* 19, 195-221.
- Diab, N. M. (2015). Effectiveness of written corrective feedback: Does type of error and type of correction matter? *Assessing Writing*, 24, 16-34.
- Dörnyei, Z. (2005). *The psychology of the language learner*. Mahwah, NJ: Lawrence Erlbaum Associates.

- Duff, P. (2008). *Case study research in applied linguistics*. Taylor & Francis.
- Egi, T. (2007). Recasts, learners' perceptions, and L2 development. In A. Mackey (Ed.), *Conversational interaction in second language acquisition: A series of empirical studies* (pp. 249–267). Oxford: Oxford University Press.
- Ellis, N. C. & Sinclair, S. (1996). Working memory in the acquisition of vocabulary and syntax: Putting Language in Good Order. *Quarterly Journal of Experimental Psychology*, 49, 234–250. <http://dx.doi.org/10.1080/713755604>
- Ellis, N. C. (2005). At the interface: Dynamic interactions of explicit and implicit language knowledge. *Studies in second language acquisition*, 27(02), 305–352. <https://doi.org/10.1017/S027226310505014X>
- Ellis, R. (2008). *The Study of Second Language Acquisition*. (2<sup>nd</sup> ed.). Oxford: Oxford University Press.
- Ellis, R. (2009). A typology of written corrective feedback types. *ELT Journal*, 63(2), 97–107. doi:10.1093/elt/ccn023
- Ellis, R. (2010). A Framework for Investigating Oral and Written Corrective Feedback. *Studies in Second Language Acquisition*, 32, 335–349. doi:10.1017/S0272263109990544
- Ellis, R., Sheen, Y., Murakami, M., & Takashima, H. (2008). The effects of focused and unfocused written corrective feedback in an English as a foreign language context. *System*, 36(3), 353–371. <https://doi.org/10.1016/j.system.2008.02.001>
- Engle, R. W. (2002). Working memory capacity as executive attention. *Current Directions in psychological Science*, 11(1), 19–23.
- Fathman, A., & Whalley, E. (1990). Teacher response to student writing: Focus on form versus content. *Second language writing: Research insights for the classroom*, 9, 178–190. <https://doi.org/10.1017/CBO9781139524551.016>
- Feak, C., & Swales, J. M. (2000). *English in today's research world: A writing guide* (No. 428 S71). University of Michigan, Ann Arbor, MI (EUA).
- Ferris, D. R. (2004). The “grammar correction” debate in L2 writing: Where are we, and where do we go from here? (and what do we do in the meantime?). *Journal of Second Language Writing* 13(1), 49–62. <https://doi.org/10.1016/j.jslw.2004.04.005>
- Ferris, D. R. (1999). The case for grammar correction in L2 writing classes: A response to Truscott. *Journal of second language writing*, 8(1), 1–11. [https://doi.org/10.1016/S1060-3743\(99\)80110-6](https://doi.org/10.1016/S1060-3743(99)80110-6)
- Ferris, D. R. (2002). *Treatment of error in second language student writing*. Michigan: The University of Michigan Press.

- Ferris, D. (2003). *Response to student writing: Implications for second language students*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Ferris, D. R. (2006). Does error feedback help student writers? New evidence on the short- and long-term effects of written error correction. In F. H. Ken Hyland (Ed.), *Feedback in second language writing: Contexts and issues* (pp. 81-104). New York, NY: Cambridge University Press.
- Ferris, D. R., & Roberts, B. (2001). Error feedback in L2 writing classes: How explicit does it need to be? *Journal of second language writing*, 10(3), 161-184. [https://doi.org/10.1016/S1060-3743\(01\)00039-X](https://doi.org/10.1016/S1060-3743(01)00039-X)
- Ferris, D. R., Chaney, S. J., Komura, K., Roberts, B. J., & McKee, S. (2000, March). Perspectives, problems, and practices in treating written error. In *Colloquium presented at International TESOL Convention* (pp. 14-18). Vancouver, BC.
- Frear, D. (2012). *The effect of written corrective feedback and revision on intermediate Chinese learners' acquisition of English*. Unpublished doctoral dissertation, University of Auckland, Auckland, New Zealand.
- Gass, S. M. (1997). *Input, interaction and the second language learner*. Mahwah: Lawrence Erlbaum Associates.
- Gathercole, S. E., & Baddeley, A. D. (1993). Phonological working memory: A critical building block for reading development and vocabulary acquisition? *European Journal of Psychology of Education*, 8(3), 259. doi:10.1007/BF03174081
- Goldschneider, J. M., & DeKeyser, R. M. (2001). Explaining the "Natural Order of L2 Morpheme Acquisition" in English: A Meta-analysis of multiple determinants. *Language learning*, 51(1), 1-50. doi: 10.1111/1467-9922.00147
- Goo, J. (2012). Corrective feedback and working memory capacity in interaction-driven L2 learning. *Studies in Second Language Acquisition*, 34(3), 445-474. <https://doi.org/10.1017/S0272263112000149>
- Guo, Q. (2015). *The effectiveness of written CF for L2 development: A mixed method study of written CF types, error categories and proficiency levels*. Unpublished doctoral dissertation, AUT University, Auckland, New Zealand.
- Harrington, M., & Sawyer, M. (1992). L2 working memory capacity and L2 reading skill. *Studies in Second Language Acquisition*, 14(1), 25-38. <https://doi.org/10.1017/S0272263100010457>
- Hinkel, E. (2002). *Second language writers' text: Linguistic and rhetorical features*. New York: Routledge.



- Hulstijn, J. (1995). Not all grammar rules are equal: Giving grammar its proper place in foreign language teaching. In R. Schmidt (Ed.), *Attention and awareness in foreign language teaching and learning* (pp. 359-386). Honolulu: University of Hawai'i at Manoa, Second Language Teaching and Curriculum Center.
- Hyland, K. (1998). *Hedging in scientific research articles*. Amsterdam: John Benjamins Publishing.
- Hyland, K. (2000). *Disciplinary discourses: Social interactions in academic writing*. London: Longman.
- Hyland, K. (2003). Genre-based pedagogies: A social response to process. *Journal of second language writing*, 12(1), 17-29. [https://doi.org/10.1016/S1060-3743\(02\)00124-8](https://doi.org/10.1016/S1060-3743(02)00124-8)
- Juffs, A. (2004). Representation, processing and working memory in a second language. *Transactions of the Philological Society*, 102(2), 199-225.
- Juffs, A., & Harrington, M. (2011) Aspects of working memory in L2 learning. *Language Teaching*, 44(02), 137-166. <https://doi.org/10.1017/S0261444810000509>
- Just, M. A., & Carpenter, P. A. (1992). A capacity theory of comprehension: individual differences in working memory. *Psychological review*, 99(1), 122.
- Kane, M. J., Brown, L. H., McVay, J. C., Silvia, P. J., Myin-Germeys, I., & Kwapil, T. R. (2007). For whom the mind wanders, and when: An experience-sampling study of working memory and executive control in daily life. *Psychological science*, 18(7), 614-621.
- Kane, M. J., & Engle, R. W. (2003). Working-memory capacity and the control of attention: The contributions of goal neglect, response competition, and task set to Stroop interference. *Journal of Experimental Psychology: General*, 132(1), 47-70.
- Kempe, V., & Brooks, P. J. (2008). Second language learning of complex inflectional. *Language Learning*, 58(4), 703-746.
- 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. doi: 10.1111/j.1540-4781.1991.tb05359.x
- Kormos, J. (2012). The role of individual differences in L2 writing. *Journal of Second Language Writing*, 21(4), 390-403. <https://doi.org/10.1016/j.jslw.2012.09.003>
- Kormos, J., & Sáfár, A. (2008). Phonological short-term memory, working memory and foreign language performance in intensive language learning.

*Bilingualism: Language and Cognition*, 11(2), 261-271.<https://doi.org/10.1017/S1366728908003416>

Krashen S.D. (2003). *Explorations in language acquisition and use*. Portsmouth, NH: Heinemann.

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

Krashen, S.D. (1994). The input hypothesis and its rivals. In N. Ellis (Ed.), *Implicit and explicit learning of languages* (pp.45-77). London, UK: Academic Press.

Krause, K. L., & Coates, H. (2008). Students' engagement in first-year university. *Assessment & Evaluation in Higher Education*, 33(5), 493-505.

Lalande, J. F. (1982). Reducing composition errors: An experiment. *The Modern Language Journal*, 66(2), 140-149.doi: 10.1111/j.1540-4781. 1982. tb0 69 73.x

Leeser, M. J. (2007). Learner-based factors in L2 reading comprehension and processing grammatical form: Topic familiarity and working memory. *Language Learning* 57, 229–270.doi: 10. 11 1 1 /j.1467-9922.2007.00408.x

Li, S. (2013). The interface between feedback type, L2 proficiency, and the nature of the linguistic target. *Language Teaching Research*, 97(3), 634-654.doi: 10. 1177 /1362 1 688 13510384

Light, L. L., & Anderson, P. A. (1985). Working-memory capacity, age, and memory for discourse. *Journal of Gerontology*, 40(6), 737-747. <https://doi.org/10.1093/geronj/40.6.737>

Linck, J. A., Osthus, P., Koeth, J. T., & Bunting, M. F. (2014). Working memory and second language comprehension and production: A meta-analysis. *Psychonomic Bulletin & Review*, 21(4), 861-883. doi: 10.3758/s13423-013-0565-2.

Mackey, A. (2006). Feedback, noticing and instructed second language learning. *Applied linguistics*, 27(3), 405-430.

Mackey, A., & Gass, S. M. (2005). *Second language research: Methodology and design*. New York: Routledge.

Mackey, A., & Sachs, R. (2012). Older learners in SLA research: A first look at working memory, feedback, and L2 development. *Language Learning*, 62(3), 704-740.doi: 10 .1111/j.1467-9922.2011.00649.x

Mackey, A., Adams, R., Stafford, C. & Winke, P. (2010). Exploring the relationship between modified output and working memory capacity. *Language Learning*, 60(3), 501–533.doi: 10.1111/j.1467-9922. 2010. 00 56 5.x

- Masoura, E. V., & Gathercole, S. E. (1999). Phonological Short-term Memory and Foreign Language Learning. *International Journal of Psychology*, 34(5-6), 383-388. [https://doi.org/10.1016/0749-596X\(91\)90040-Q](https://doi.org/10.1016/0749-596X(91)90040-Q)
- Masoura, E. V., & Gathercole, S. E. (2005). Phonological short-term memory skills and new word learning in young Greek children. *Memory*, 13, 422-429.
- McDonough, K., & Mackey, A. (2006). Responses to recasts: Repetitions, primed production, and linguistic development. *Language Learning*, 56(4), 693-720.
- McLaughlin, B. (1987) *Theories of Second-Language Learning*, London: Edward Arnold.
- McLaughlin, B. (1990). Restructuring. *Applied linguistics*, 11(2), 113-128. <https://doi.org/10.1093/applin/11.2.113>
- Miyake, A., & Friedman, N. P. (1998). Individual differences in second language proficiency: Working memory as language aptitude. Foreign language learning: *Psycholinguistic studies on training and retention*, 339-364.
- Nopashani, K. M., & Azari, H. (2013). The effects of interaction on the acquisition of the passive English structure by Iranian EFL learners. *Asian Journal of Management Sciences and Education [AJMSE]*, 2(4), 169–176
- O'Brien, I., Segalowitz, N., Freed, B., & Collentine, J. (2007). Phonological memory predicts second language oral fluency gains in adults. *Studies in Second Language Acquisition*, 29(04), 557-581. doi:10.1017/S027226310707043X
- Ortega, L. (2009). *Understanding second language acquisition*. London: Hodder Arnold.
- Osaka, M., & Osaka, N. (1992). Language-independent working memory as measured by Japanese and English reading span tests. *Bulletin of the Psychonomic Society*, 30(4), 287-289. doi: 10.3758/BF03330466
- Pallant, J. (2001). *SPSS Survival Manual: A Step by Step Guide to Data Analysis Using SPSS for Windows (Versions 10 and 11): SPSS Student Version 11.0 for Windows*. London: Open University Press.
- Papagno, C., Valentine, T., Baddeley, A. (1991). Phonological short-term memory and foreign-language vocabulary learning. *Journal of Memory and Language*, 30(3), 331-347. [https://doi.org/10.1016/0749-596X\(91\)90040-Q](https://doi.org/10.1016/0749-596X(91)90040-Q)
- Pawlak, M. (2014). Investigating learner engagement with oral corrective feedback: aims, methodology, outcomes. In *Awareness in Action* (pp. 69-84). Springer International Publishing
- Pienemann, M. (1998). *Language processing and second language development: Processability theory*. Amsterdam: John Benjamins Publishing.

- Posner, M. I., & Petersen, S. E. (1990). The attention system of the human brain. *Annual review of neuroscience*, 13(1), 25-42.
- Révész, A. (2012). Working memory and the observed effectiveness of recasts on different L2 outcome measures. *Language Learning*, 62(1), 93–132. doi:10.1111/j.1467-9922.201.0690.x
- Robb, T., Ross, S., & Shortreed, I. (1986). Salience of feedback on error and its effect on EFL writing quality. *TESOL Quarterly*, 20(1), 83-96. doi:10.2307/3586390
- Robinson, P. (1995). Task complexity and second language narrative discourse. *Language Learning*, 45(1), 99-140. doi: 10.1111/j.1467-1770.1995.tb00964.x
- Rummel, S. (2014). *Student and Teacher Beliefs about Written CF and the Effect those Beliefs Have on Uptake: A Multiple Case Study of Laos and Kuwait*. Unpublished doctoral dissertation, AUT University, Auckland, New Zealand.
- Rummel, S., & Bitchener, J. (2015). The effectiveness of written corrective feedback and the impact Lao learners' beliefs have on uptake. *Australian Review of Applied Linguistics*, 38(1), 66-84.
- Sachs, R., & Polio, C. (2007). Learners' uses of two types of written feedback on a L2 writing revision task. *Studies in Second Language Acquisition*, 29(1), 67-100.
- Sagarra, N. (2007). Working memory and L2 processing of redundant grammatical forms. In Z. Han & E. S. Park (eds.), *Understanding second language process* (pp. 133–147). Bristol, UK: Multilingual Matters.
- Schmidt, R. (1990). The role of consciousness in second language learning. *Applied Linguistics* 11(2), 129–158.
- Schmidt, R. (1994). Implicit learning and cognitive unconscious: Of artificial grammars and SLA. In N. Ellis (Ed.), *Implicit and explicit learning of language* (pp.165-209). London: Academic Press.
- Schmidt, R. (1995). Consciousness and foreign language learning: A tutorial on attention and awareness in learning. In R. Schmidt, (Ed.), *Attention and awareness in foreign language learning* (pp. 1-63). Honolulu, HI: University of Hawai'i, National Foreign Language Resource Center.
- Schmidt, R. (2001). Attention. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 3-32). Cambridge: Cambridge University Press.
- Semke, H. D. (1984). Effects of the red pen. *Foreign Language Annals*, 17(3), 195-202. doi:10.1111/j.1944-9720.1984.tb01727.x

- Service, E. (1992). Phonology, working memory, and foreign-language learning. *The Quarterly Journal of Experimental Psychology*, 45(1), 21-50. <http://dx.doi.org/10.1080/14640749208401314>
- Service, E., & Craik, F. I. M. (1993). Differences between young and older adults in learning a foreign vocabulary. *Journal of Memory and Language*, 32(5), 608-623. <https://doi.org/10.1006/jmla.1993.1031>
- Service, E. & Kohonen, V. (1995). Is the relation between phonological memory and foreign language learning accounted for by vocabulary acquisition? *Applied Psycholinguistics*, 16(2), 155–172. <https://doi.org/10.1017/S0142716400007062>
- Shahnazari, M. (2012). *The role of working memory in second language reading comprehension*. Unpublished doctoral dissertation, University of Auckland, Auckland, New Zealand.
- Shah, P., & Miyake, A. (1999). *Models of working memory: Mechanisms of active maintenance and executive control*. New York, NY: Cambridge University Press.
- Sheen, Y. (2006). Exploring the relationship between characteristics of recasts and learner uptake. *Language Teaching Research*, 10(4), 361-392.
- Sheen, Y. (2007 a). The effect of focused written corrective feedback and language aptitude on ESL learners' acquisition of articles. *TESOL Quarterly*, 41(2), 255–283. doi: 10.1002/j.1545-7249.2007.tb00059.
- Sheen, Y. (2008). Recasts, language anxiety, modified output, and L2 learning. *Language Learning*, 58(4), 835-874. doi:10.1111/j.1467-9922.2008.00480.x
- Sheen, Y. (2010 a). Introduction: The Role of Oral and Written Corrective Feedback in SLA. *Studies in Second Language Acquisition*, 32(2), 169– 179. <http://dx.doi.org/10.1017/S027226310990489>
- Sheen, Y. (2010 b). Differential effects of oral and written corrective feedback in the ESL classroom. *Studies in Second Language Acquisition*, 32(2), 203– 234. <https://doi.org/10.1017/S0272263109990507>
- Sheen, Y. (2011). *Corrective feedback, individual differences and second language learning*. The Netherlands: Springer.
- Sheen, Y., Wright, D., & Moldawa, A. (2009). Differential effects of focused and unfocused written correction on the accurate use of grammatical forms by adult ESL learners. *System*, 37(4), 556-569. doi:10.1016/j.system.2009.09.002.
- Sheppard, K. (1992). Two feedback types: Do they make a difference? *RELC Journal*, 23(1), 103-110. doi: 10.1177/003368829202300107
- Shintani, N. (2017). The Effects of the Timing of Isolated FFI on the Explicit Knowledge and Written Accuracy of Learners with Different Prior

- Knowledge of the Linguistic Target. *Studies in Second Language Acquisition*, 39(1), 129-166. doi:10.1017/S0272263116000127
- Shintani, N., & Ellis, R. (2013). The comparative effect of metalinguistic explanation and direct written corrective feedback on learners' explicit and implicit knowledge of the English indefinite article. *Journal of Second Language Writing*, 23(3), 286-306. doi: 10. 1016/ j. jslw. 2013.03.011
- Shintani, N., & Ellis, R. (2015). Does language analytical ability mediate the effect of written feedback on grammatical accuracy in second language writing? *System*, 49, 110-119. <https://doi.org/10.1016/j.system.2015.01.006>
- Shintani, N., Aubrey, S., & Donnellan, M. (2016). The Effects of Pre-Task and Post-Task Metalinguistic Explanations on Accuracy in Second Language Writing. *TESOL Quarterly*, 50(4), 945-955. doi: 10.1002/tesq.323
- 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(1), 103-131. doi:10.1111/lang.12029
- Skehan, P. (1998). *A cognitive approach to language learning*. Oxford: Oxford University Press.
- Slife, B. D., & Williams, R. N. (1995). *What's behind the research? Discovering hidden assumptions in the behavioural sciences*. London: Sage.
- Speciale, G., Ellis, N. C., & Bywater, T. (2004). Phonological sequence learning and short-term store capacity determine second language vocabulary acquisition. *Applied psycholinguistics*, 25(2), 293-321. doi: <https://doi.org/10.1017/S0142716404001146>.
- Stefanou, C., & Revesz, A. (2015). Direct written corrective feedback, learner differences, and the acquisition of second language article use for generic and specific plural reference. *The Modern Language Journal*, 99(2), 263-282. doi: 10.1111/modl.12212.
- Storch, N., & Wigglesworth, G. (2010). Learners' processing, uptake, and retention of corrective feedback on writing. *Studies in Second Language Acquisition*, 32(02), 303-334. <https://doi.org/10.1017/S0272263109990532>.
- Swain, M. (1985). Communicative competence: Some roles of comprehensible input and comprehensible output in its development. In S. Gass, & C. Madden (Eds.), *Input in second language acquisition* (pp. 235-253). Rowley, MA: Newbury House.
- Swain, M. (1995). Three functions of output in second language learning. In G. Cook & B. Seidlhofer (Eds), *Principle and practice in applied linguistics: Studies in honour of H.G. Widdowson* (pp. 125-144). Oxford, UK: Oxford University Press.

- Swain, M. (2009). Languaging, agency and collaboration in advanced second language proficiency. In Byrnes, H. (Ed.), *Advanced language learning: The contribution of Halliday and Vygotsky* (pp. 95-108). A&C Black.
- Swain, M., & Lapkin, S. (1995). Problems in output and the cognitive processes they generate: A step towards second language learning. *Applied linguistics*, 16(3), 371-391. doi: 10.1093/applin/16.3.37.
- Swain, M., & Lapkin, S. (2002). Talking it through: Two French immersion learners' response to reformulation. *International Journal of Educational Research*, 37(3), 285-304. [https://doi.org/10.1016/S0883-0355\(03\)00006-5](https://doi.org/10.1016/S0883-0355(03)00006-5)
- Swanson, H. L. (1993). Working memory in learning disability subgroups. *Journal of experimental child psychology*, 56(1), 87-114.
- Tomlin, R. S., & Villa, V. (1994). Attention in cognitive science and second language acquisition. *Studies in second language acquisition*, 16(02), 183-203.
- Trofimovich, P., & Baker, W. (2007). Learning prosody and fluency characteristics of second language speech: The effect of experience on child learners' acquisition of five suprasegmentals. *Applied Psycholinguistics*, 28(2), 251-276. <https://doi.org/10.1017/S0142716407070130>
- Truscott, J. (1996). The case against grammar correction in L2 writing classes. *Language learning*, 46(2), 327-369. doi:10.1111/j.1467-1770.1996.tb01238.x
- Truscott, J. (1998). Noticing in second language acquisition: A critical review. *SLA Research* 14(2), 103-135.
- Truscott, J. (1999). The case for "The case against grammar correction in L2 writing classes": A response to Ferris. *Journal of Second Language Writing*, 8(2), 111-122. [https://doi.org/10.1016/S1060-3743\(99\)80124-6](https://doi.org/10.1016/S1060-3743(99)80124-6)
- Truscott, J. (2004). Evidence and conjecture on the effects of correction: A response to Chandler. *Journal of second language writing*, 13(4), 337-343. 10.1016/j.jslw.2004.05.002
- Truscott, J. (2007). The effect of error correction on learners' ability to write accurately. *Journal of second language Writing*, 16(4), 255-272. <https://doi.org/10.1016/j.jslw.2007.06.003>
- Truscott, J. (2010). Some thoughts on Anthony Bruton's critique of the correction debate. *System*, 38(2), 329-335. <https://doi.org/10.1016/j.system.2010.03.014>.
- Truscott, J., & Hsu, A. Y. p. (2008). Error correction, revision, and learning. *Journal of Second Language Writing*, 17(4), 292-305. Doi:10.1016/j.jslw.2007.05.003.
- Van Beuningen, C. G. (2011). The effectiveness of comprehensive corrective feedback in second. *Studies*, 10(2), 1-27.



- Van Beuningen, C. G., De Jong, N. H., & Kuiken, F. (2008). The effect of direct and indirect corrective feedback on L2 learners' written accuracy. *ITL-Review of Applied Linguistics*, 156, 279-296. doi: [10.1075/itl.156.24beu](https://doi.org/10.1075/itl.156.24beu)
- 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), 1-41. doi: 10.1111/j.1467-9922.2011.00674.x
- VanPatten, B. (Ed.). (2004). *Processing instruction: Theory, research, and commentary*. Routledge.
- VanPatten, B., & Benati, A. G. (2010). *Key terms in second language acquisition*. London: Bloomsbury Publishing.
- VanPatten, B., & Oikarinen, S. (1996). Explanation versus structured input in processing instruction. *Studies in Second Language Acquisition*, 18(04), 495-510. <https://doi.org/10.10272263100015394>
- Vyatkina, N. (2010). The effectiveness of written corrective feedback in teaching beginning German. *Foreign Language Annals*, 43(4), 671-689.
- Walter, C. (2004). Transfer of reading comprehension skills to L2 is linked to mental representations of text and to L2 working memory. *Applied Linguistics*, 25(3), 315-339.
- Waters, G.S., & Caplan, D. (1996). The measurement of verbal working memory capacity and its relation to reading comprehension. *The Quarterly Journal of Experimental Psychology*, 49A, 51-79. <http://dx.doi.org/10.1080/713755607>
- Williams, J. (2012). The potential role (s) of writing in second language development. *Journal of Second Language Writing*, 21(4), 321-331. doi: 10.1016/j.jslw.2012.09.007
- Yang, Y., & Lyster, R. (2010). Effects of form- practice and feedback on Chinese EFL learners' acquisition of regular and irregular past tense forms. *Studies in Second Language Acquisition*, 32(2), 235-263. doi:10.1017/S0272263109990519



# Appendix A



A U T E C  
S E C R E T A R I A T

15 June 2015

John Bitchener  
Faculty of Culture and Society

Dear John

Re Ethics Application: **15/184 Written corrective feedback, individual differences and second language acquisition of the English passive voice.**

Thank you for providing evidence as requested, which satisfies the points raised by the Auckland University of Technology Ethics Subcommittee (AUTEC).

Your ethics application has been approved for three years until 15 June 2018.

As part of the ethics approval process, you are required to submit the following to AUTEC:

- A brief annual progress report using form EA2, which is available online through <http://www.aut.ac.nz/researchethics>. When necessary this form may also be used to request an extension of the approval at least one month prior to its expiry on 15 June 2018;
- A brief report on the status of the project using form EA3, which is available online through <http://www.aut.ac.nz/researchethics>. This report is to be submitted either when the approval expires on 15 June 2018 or on completion of the project.

It is a condition of approval that AUTEC is notified of any adverse events or if the research does not commence. AUTEC approval needs to be sought for any alteration to the research, including any alteration of or addition to any documents that are provided to participants. You are responsible for ensuring that research undertaken under this approval occurs within the parameters outlined in the approved application.

AUTEC grants ethical approval only. If you require management approval from an institution or organisation for your research, then you will need to obtain this. If your research is undertaken within a jurisdiction outside New Zealand, you will need to make the arrangements necessary to meet the legal and ethical requirements that apply there.

To enable us to provide you with efficient service, please use the application number and study title in all correspondence with us. If you have any enquiries about this application, or anything else, please do contact us at [ethics@aut.ac.nz](mailto:ethics@aut.ac.nz).

All the very best with your research,

Kate O'Connor  
Executive Secretary

Auckland University of Technology Ethics Committee Cc:

Saeed Roshan [saeedrosh10@gmail.com](mailto:saeedrosh10@gmail.com)

## Appendix B: Information Sheet for students

# Participant Information Sheet



### Date Information Sheet Produced:

20.05.2015

### Project Title

Written Corrective Feedback, Individual Differences and Second Language Acquisition of the English Passive Voice

### An Invitation

My name is Saeed Roshan. I am a PhD Candidate at Auckland University of Technology. I am inviting you to participate in my research programme, which will form the basis of a PhD thesis. I am investigating the efficacy of two types of written corrective feedback on learners' immediate and delayed output as well as determining if individual differences affect learners' response to written corrective feedback. Your participation in this research is voluntary and you may withdraw at any time.

### What is the purpose of this research?

The purpose of this research is to learn which types of written corrective feedback (direct corrective feedback, metalinguistic corrective feedback) help learners to improve their writing. The second aim of the study is to investigate if working memory and phonological short-term memory, as individual differences, affect learners' response to written corrective feedback. The results of this project will be written up in the form of a PhD thesis and may also be presented to conferences and published in journals which discuss English language learning issues. Your identity will always be kept confidential.

### How was I chosen for this research?

You have been approached because you are at an upper-intermediate level in the language school; however, you are in no way obligated to participate. Your participation in this project is completely voluntary, and you can withdraw from the research at any time up to the end of data collection (June, 2016).

### What will happen in this research?

In order to conduct the research, participants will be invited to participate in the following stages of the project:

- 1) Writing tasks
- 2) Working memory test
- 3) Phonological short-term memory test

The first stage of the project will involve your participation in four writing tasks. Each writing task will take about 40 minutes. Writing tasks will be conducted over eight weeks, that is, week 1, week 2, week 5 and week 8. In week 2, you will also be asked to revise the writing text that you conducted in week 1. This will take between 10 and 25 minutes.

The second stage of the project will involve your participation in the working memory test. This will take 10 minutes of your time.

The third stage of the project will involve your participation in the phonological short-term memory test. This test will also take 10 minutes of your time.

**What are the discomforts and risks?**

You might feel uncomfortable if you have done poorly in the language tests. However, the purpose of this project is to know about your learning needs and to help you to improve your writing.

**How will these discomforts and risks be alleviated?**

I (the primary researcher) will keep all your information confidential. However, if you experience any discomfort you can consult your school consultation office.

**What are the benefits?**

If you agree to take part in this research project, you will be helping me to find out more information about which types of written corrective feedback help learners to develop their writing skills, and if learners working memory and phonological short-term memory impact on the use of written CF for L2 development. I also hope to publish the research findings of this study in language journals. The finding of this research will fulfil the requirement of my doctoral study.

**How will my privacy be protected?**

All data collected from participants will be kept in a locked cabinet in Saeed Roshan's office (WT1105C). These include a memory stick and hard copies of transcriptions. All participants will be identified only by code. The researcher will not share any information of the study with the class teacher, family, friends or outsiders in any way which could identify the participants of the study. The teacher will have no access to the test result data.

**What are the costs of participating in this research?**

The research will involve your time. In total, you will be asked to use approximately 210 to 225 minutes of your time, spread across 6 classes over 4 weeks. In fact, the testing is part of normal classroom activity and therefore not a cost-of-time for this research.

**What opportunity do I have to consider this invitation?**

Please let me know if you are willing to participate within 1 week of receiving the invitation.

**How do I agree to participate in this research?**

I will give you a Consent Form to sign.

**Will I receive feedback on the results of this research?**

A summary of the research will be given to all participants who indicate their interest on the Consent Form. In addition, any journal articles published will be forwarded to you.

**What do I do if I have concerns about this research?**

Any concerns regarding the nature of this project should be notified in the first instance to the Project Supervisor, Prof. John Bitchener, [jbitchen@aut.ac.nz](mailto:jbitchen@aut.ac.nz) ,+64921 9999, ext. 7830.

Concerns regarding the conduct of the research should be notified to the Executive Secretary of AUTECH, Kate O'Connor, [ethics@aut.ac.nz](mailto:ethics@aut.ac.nz) ,0064921 9999 ext 6038.

**Whom do I contact for further information about this research?**

***Researcher Contact Details:***

Saeed Roshan

[Saeedrosh10@gmail.com](mailto:Saeedrosh10@gmail.com) , contact number in Iran 00989163730719, in New Zealand +642102312216

***Project Supervisor Contact Details:***

Prof. John Bitchener, [jbitchen@aut.ac.nz](mailto:jbitchen@aut.ac.nz), +64921 9999, ext. 7830.

**Approved by the Auckland University of Technology Ethics Committee on 15.Jun. 2015, AUTECH Reference number 15/184.**

## Appendix C: Consent Form

### Consent Form



Project title: **Written Corrective Feedback, Individual Differences and Second Language Acquisition of the English Passive Voice**

Project Supervisor: **Prof. John Bitchener**

Researcher: **Saeed Roshan**

- I have read and understood the information provided about this research project in the Information Sheet dated 20/05/2015
- I agree to participate in the following activities: tick Yes or No

Tasks	Yes	No
1) Writing tasks		
2) Working memory test		
3) Phonological short term memory		

- I understand that the Memory Tests and Writing Tasks will be given in class time as part of the class curriculum. The researcher will send a consent form to me prior to the Memory Tests and Writing Tasks. This consent form will then be collected before I do the Memory Tests and Writing Tasks. I consent for the test results to be given to the researcher for the purpose of his qualification.
- I understand that the information obtained from my participation in the project will be reported or published in a way that does not identify me as its source.
- I understand that data will be kept for 6 years, after which they will be destroyed.
- I understand that the head of school has given assurance that my participation or nonparticipation in this research will in no way influence my grades in the course nor my relationship with my teachers or the school.
- I understand that I may withdraw myself or any information that I have provided for this project at any time prior to completion of data collection, without being disadvantaged in any way.
- If I withdraw, I understand that all relevant information including tapes and transcripts, or parts thereof, will be destroyed.

I wish / do not wish to receive the summary.

Participant's signature:.....

Participant's name:.....

Participant's Contact Details:                      Date: .....

Approved by the Auckland University of Technology Ethics Committee on

.....AUTEC Reference number

## Appendix D: Non word span test

**Listen carefully and repeat.**

### **Practise session**

- a. git fim pil teg
- b. sog kor mol por
- c. mor gom rin lok tor

### **Test session**

1. peb kib bon deet
2. peeb kol goob mab
3. pib kom gook tam
4. neeg gop doob jat
5. pim goot neeb kig doog
6. meb teeb dook cam jawn
7. teel nog gub pem chad
8. jep cham tudge meech pag
9. noog teed gadge pab chud
10. mep teeg keb chim nup jit
11. jick mip chool lod nug tep
12. teeg chan mig padge dop nam

- 13.geed mun peb cheem tep nuck
- 14.bick meep tooch leck nam gab
- 15.choom mit gab tidge pag nool
- 16.jeck leem gan chut bock mon tud
17. mitch tem jeeg lib cug bup neb
18. pock mun tob juck lidge ged coom
19. toock jeel peeb modge dack lig neeb
20. lon cam deech mot jooch ked gock
21. dook mip chon teep jal noog goot
22. kom chen meb lud tam dit loog



## Appendix E: Reading Span Test

هر جمله را بخوانید و تعیین کنید که آیا مفهوم درستی دارد یا خیر  
بر روی کیبورد بله یا خیر را فشار دهید.  
آخرین کلمه هر جمله را بنویسید.  
تا جایی که می توانید سریع و صحیح پاسخ دهید.  
حالا برای شروع تمرین، کلید فاصله را فشار دهید

### تمرین

گاه فردی تصمیم به انجام کاری می گیرد که فوق قدرت و توانایی او است.  
احتمالا ناتوان ترین افراد بشر کسی است که نتواند کسی دیگری با دوست شود.  
این نشانه ی رشد فکری است که یادآوری عیب هایمان را به عنوان هدیه از دوست بپذیریم.

آخرین کلمه ی هر جمله را بنویسید.  
برای ادامه کلید فاصله را فشار دهید.

بر عهده هر انسانی در این دنیا است که استعدادهای را خدادادی باز شناسد.  
خوب است که ما در یک ورزش خاص برای رشد استعدادهای خود ظاهر شویم.  
در گرفتاری ها به تو امید بسته ام و در نا امنی ها تو پناه گاه را خویش یافته ام.

آخرین کلمه ی هر جمله را بنویسید.  
برای ادامه کلید فاصله را فشار دهید.

دوستان خود را با دیگر ارزیابی می کنم و آنها را با معیارهای جدید می سنجم.  
اگر من تجربه ی این چند سال را داشتم شاید دیگری را راه بر می گزیدم.  
رنگین کمان پس از بارش باران در آسمان ظاهر می گردد و بعد از اندکی ناپدید می شود.  
گیاهان نور و دی اکسید کربن را جذب می کنند تا با آن برای غذا خود بسازند.

آخرین کلمه ی هر جمله را بنویسید  
پایان تمرین  
برای ادامه کلید فاصله را فشار دهید.

اگر در آمد بهتری در ماه های آینده داشته باشم، شاید امسال خانه ی بزرگتری بخرم.  
گل نرگس زیبایی من بر اثر سهل انگاری خودم جلوی من چشمان پژمرد.  
تازه وارد اتاق او شده بودم که حمید سراسیمه از جای خود برخاست.

آخرین کلمه ی هر جمله را بنویسید.  
برای ادامه کلید فاصله را فشار دهید.

تیم ایران تلاش زیادی برای پیروزی کرد، اما حریف متأسفانه به دوباره باخت.  
با شیلنگ آب زیادی روی زمین پاشیدم به این امید که گرد و خاک بخوابد.  
من با آشنا شدن با فلسفه بسیاری از افکار قدیمی خود را به دور انداختم.

آخرين كلمه ي هر جمله را بنويسيد.  
براي ادامه كليد فاصله را فشار دهيد.

من در كودكي از تاريخي زياد مي ترسيدم هنوز هم گاه گاهي تاريخي مرا مي ترساند.  
زنبور زيبايي ديدم كه با ظرافت هر چه تمام تر شهد گل ها را مي نوشيد.  
پدر هميشه به ما مي گفت: بايد با بعضي تا دردها آخر عمر ساخت.

آخرين كلمه ي هر جمله را بنويسيد.  
براي ادامه كليد فاصله را فشار دهيد.

ابو علي سينا حكيم معروف ايراني قرن هاست كه در آسمان علم جهان مي درخشد.  
بعضي از مارها مي توانند زهر خود به را تا دو متر طرف دشمن بپاشند.  
اگر مانند سال هاي گذشته امسال هم بارندگي كم باشد دچار كم آبي مي شويم.  
دكتور دست شكسته من را گچ گرفت و با يك پارچه آن به گردنم آويخت.

آخرين كلمه ي هر جمله را بنويسيد.  
براي ادامه كليد فاصله را فشار دهيد.

هوشنگ هنوز مانند قديم هر روز شير بيست راس با گاو را دست مي دوشد.  
پدرم هميشه مي گفت به دليل مرگ مادرت اينگونه از پيرو پا افتاده شده ام.  
اگر باغبان پير و مهربان اين جا بود كسي اين گلهاي زيبا را نمي چيد.  
اي كاش هر كودكي در زمان نياز در آغوش پر مهر مادر خود آرام بگيرد.

آخرين كلمه ي هر جمله را بنويسيد.  
براي ادامه كليد فاصله را فشار دهيد.

من حاضرم براي بدست آ و ردن مدرك علمي بالاتر با همه ي مشكلات روبرو بجنم.  
با اينكه روي لوله ي آب را پوشانيدم، باز هم زمستان امسال لوله تركيد.  
در مهماني به ظاهر آرام نشسته بودم اما دلم سر كه مثل سير و مي جوشيد.  
حسن پرچم را به دست گرفته بود و به نشانه را پيروي آن مي چرخاند.

آخرين كلمه ي هر جمله را بنويسيد.  
براي ادامه كليد فاصله را فشار دهيد.

اشتباه لفظي احمد در كلاس باعث شد سال گذشته همه به او بخندند.  
پسرم لوله بخاري را كشيد با اين كار از لوله جاي خود در رفت.  
حميده يكي از جواهرات خود را كه پارسال گم کرده بود، پشت كمد لباس ها يافت.  
مادر با لبخند گفت، بسيار خوب اميدوارم كه راه همه كارها رو به باشد.  
هنوز بسياري از دختران روستايي با زحمت زياد هر سال يك فرش مي بافند.

آخرين كلمه ي هر جمله را بنويسيد.  
براي ادامه كليد فاصله را فشار دهيد.

در كتابخانه ي آن شهر كتاب هاي بسياري بودند كه در دشمنان همه را آتش سوزانده اند.  
سالهاي زيادي است كه من هر روز صبح با صورت خود را تيغ مي تراشم  
به نوشته اي هنري كه نمايش دقيقا از روي آن اجرا مي شود نمايشنامه مي گویند.

فرهاد با افتخار مي گفت كه از پارسال تاكنون دو هزار متر زمين كشاورزي خريده ام. ديشب درهاي اتاق باز بود، به همين دليل پشه ها بدن تا صبح مرا گزیدند.

آخرين كلمه ي هر جمله را بنويسيد.  
براي ادامه كليد فاصله را فشار دهيد.

پرستار مهربان و خوشرو هر روز صبح قطره چشم پدر من را مي چکاند. بسياري از كارمندان هستند كه براي امضاي يك پرونده ماهها مردم را مي دواند. بعضي ها براي اين كه ثروت بيشتري بدست آورند دست به هر كاري مي زنند. دكتور به محمد گفته بود بايد چند بار در ماه با سر خود را سدر بشويي. با اينكه راننده همه تلاش خود را كرد، باز به هم ما جلسه نرسيديم

آخرين كلمه ي هر جمله را بنويسيد.  
براي ادامه كليد فاصله را فشار دهيد.

بدون اينكه قصدي داشته باشم از مادر را با سخنان نسنجیده خود رنجاندم. خاله زهرا مرتب اصرار مي كرد، كمي استراحت كن، تو از راه دور آمده اي. تنها تفاوت فيلم نامه با نمايش نامه اين است كه را تمام حرركات بازيگران در فيلم نامه مي نويسد. زمستان كه مي شد هر شب تا سحر پدر بزرگ براي ما شاهنامه مي خواند. از پنجره اتاقم او را در حياط خانه اش مي ديدم كه آب قدم مي زد و باغچه ها را مي داد. بي شك آنان كه خوب مي نويسند، كساني هستند كه خوب علوم مختلف را مطالعه کرده اند.

آخرين كلمه ي هر جمله را بنويسيد.  
براي ادامه كليد فاصله را فشار دهيد.

اختر فيزيك شاخه اي از علم فيزيك است كه به مطالعه اجرام آسماني مي پردازد. من براي كنترل وزن خود مجبور شده ام هر روز چند كيلومتر راه بپيمايم. به محض اينكه آب جوش را در ليوان ريختم، ليوان با صداي بلند شكست. متاسفانه دستم به ظرف غذا خورد و ظرف پايين غذا به ميز افتاد. ديروز هرچه تلاش كردي نتوانستي خوي توپ بسكتبال سبد درون را بيندازي. هرچه با مينا حرف مي زدیم به ما نگاه مي كرد و هم دندانهايش را به مي سايبید.

آخرين كلمه ي هر جمله را بنويسيد .  
براي ادامه كليد فاصله را فشار دهيد.

هميشه شوق ديدن كشوري بيگانه دلهره اي شگفت در رگ من و پي بدن مي آفريند. گربه كوچك خانگي ما گاهي با همان پنجه هاي مرا ظريف و شكندنده اش صورت مي خراشد. قرآن كتاب آسماني ما مي گويد ارزشمندترين مردم نزد خدا با تقواترين آنها است. مردم محله ي ما با وجود بدن كوچك، حاج مرتضي را پهلوان مرتضي مي ناميدند. هنوز صداي سرود ملي بلند نشده بود كه پرچم در سربازان حاضر ميدان را برافراشتند . من بر اساس عادت گذشته هنگام غذا به خوردن روي نمك آن مي پاشم.

آخرين كلمه ي هر جمله را بنويسيد  
پايان آزمائش  
با تشكر