EFFECTS OF FOCUSED DIRECT AND METALINGUISTIC CORRECTIVE FEEDBACK ON L2 LEARNERS' WRITING

GRINGO MAGBANUA JUSA

FACUTLY OF LANGUAGES AND LINGUISTICS
UNIVERSITY OF MALAYA
KUALA LUMPUR

2017

EFFECTS OF FOCUSED DIRECT AND METALINGUISTIC CORRECTIVE FEEDBACK ON L2 LEARNERS' WRITING

GRINGO MAGBANUA JUSA

THESIS SUBMITTED IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF MASTER OF ENGLISH AS A SECOND LANGUAGE (MESL)

FACULTY OF LANGUAGES AND LINGUISTICS
UNIVERSITY OF MALAYA
KUALA LUMPUR

2017

UNIVERSITY OF MALAYA

ORIGINAL LITERARY WORK DECLARATION

Name of Candidate: Gringo Magbanua Jusa

Matric No: TGB 140041

Name of Degree: Master of English as a Second Language (MESL)

Title of Thesis: EFFECTS OF FOCUSED DIRECT AND METALINGUISTIC CORRECTIVE FEEDBACK ON L2 LEARNERS' WRITING

Field of Study: Language Acquisition

I do solemnly and sincerely declare that:

(1) I am the sole author/writer of this Work;

(2) This Work is original;

(3) Any use of any work in which copyright exists was done by way of fair dealing and for permitted purposes and any excerpt or extract from, or reference to or reproduction of any copyright work has been disclosed expressly and sufficiently and the title of the Work and its authorship have been acknowledged in this Work;

(4) I do not have any actual knowledge nor do I ought reasonably to know that the making of this work constitutes an infringement of any copyright work;

(5) I hereby assign all and every rights in the copyright to this Work to the University of Malaya ("UM"), who henceforth shall be owner of the copyright in this Work and that any reproduction or use in any form or by any means whatsoever is prohibited without the written consent of UM having been first had and obtained;

(6) I am fully aware that if in the course of making this Work I have infringed any copyright whether intentionally or otherwise, I may be subject to legal action or any other action as may be determined by UM.

Candid[redacted] Date: 22/2/2017

Subscribed and solemnly declared before,

[redacted] Date: 22/2/2017

Name:

Designation: Jabatan Bahasa Inggeris
Fakulti Bahasa & Linguistik
Universiti Malaya

ABSTRACT

Three intact groups (CG group, DCF group, and MCF group) were placed into treatment procedures to compare direct (DCF) and metalinguistic corrective feedback (MCF) and its effects on knowledge acquisition in basic tenses in writing. Three tests given in three-testing time were employed to measure the CF efficacy and its effects on the learners' implicit and explicit knowledge. Picture Narrative Writing Test was used to measure the efficacy of corrective feedback in writing, while Metalinguistic Knowledge Test and Timed Grammatical Judgment Test were used to measure explicit and implicit knowledge respectively. Through mixed-method design, SPSS (two-way mixed ANOVA) was used to analyse the efficacy of CF, while thematic analysis was used for the qualitative interview data. Quantitative findings revealed that MCF is more effective than DCF in improving the writing accuracy of the students. This result was further extended to the improvement of implicit and explicit knowledge of basic English tenses. On the other hand, the control group did not show any progression in any of the three tests. The qualitative thematic analysis also revealed the benefits of employing MCF as a corrective feedback which further validated the statistical result. Specifically, most of the learners in the MCF group claimed that the operationalized feedback encouraged critical thinking. Through these findings, the study has provided pedagogical implications and future research recommendations which could further enhance the study and validate the existing claims in this area.

ABSTRAK

Tiga kumpulan intek (kumpulan CG, kumpulan DCF dan kumpulan MCF) telah diletakkan di bawah prosedur pengolahan untuk membuat perbandingan di antara maklum balas pembetulan langsung (DCF) dan metalinguistik (MCF) serta kesan-kesannya terhadap pemerolehan pengetahuan kala asas dalam penulisan. Tiga ujian yang berlangsung secara berkumpulan dalam masa yang berlainan digunakan untuk mengukur keberkesanan CF dan kesan-kesannya terhadap pengetahuan eksplisit dan implisit pelajar-pelajar. Ujian Penulisan Narratif Gambar digunakan untuk mengukur keberkesanan maklum balas pembetulan dalam penulisan, manakala Ujian Pengetahuan Metalinguistik dan Ujian Pilihan Tatabahasa Ditetap Masa masing-masing digunakan untuk mengukur pengetahuan eksplisit dan implisit. Melalui reka bentuk kaedah bercampur, SPSS (ANOVA bercampur dua-hala) digunakan untuk menganalisis keberkesanan CF, manakala analisis tematik digunakan untuk menganalisis data temubual kualitatif. Penemuan `kuantitatif menunjukkan bahawa MCF adalah lebih berkesan daripada DCF dalam peningkatan ketepatan penulisan pelajar. Keputusan ini telah dilanjutkan lagi kepada peningkatan pengetahuan eksplisit dan implisit dalam kala asas Bahasa Inggeris. Sebaliknya, kelompok kawalan tidak menunjukkan sebarang perkembangan dalam mana-mana tiga ujian ini. Analisis tematik kualitatif juga menunjukkan manfaat-manfaat penggunaan MCF sebagai suatu maklum balas pembetulan yang selanjutnya mengesahkan hasil statistik. Khususnya, kebanyakan pelajar dalam kumpulan MCF mendakwa bahawa maklum balas yang beroperasi menggalakkan pemikiran kritikal. Melalui penemuan-penemuan ini, kajian ini telah memberikan implikasi-implikasi pedagogi dan cadangan-cadangan penyelidikan masa depan yang boleh mempertingkatkan lagi kajian dan mengesahkan dakwaan-dakwaan yang sedia ada dalam bidang ini.

ACKNOWLEDGEMENTS

Countless thanks are due to the graciousness and benevolence of the Almighty God who continuously guide me to accomplish great things in life. He is the one who provided me with wonderful people whom have shared their precious time and effort to make sure I am on the right track. Everything became possible because of my faith to YOU.

My family is surely a treasure in my life. I would like to say a big THANK YOU for keeping me in your prayer all the time, and for the never-ending support in all the struggles I have faced and will be facing on. Your love is unfading and so are mine to you. This craft is dedicated to you all.

A big thank you also goes to my ever-supportive supervisor, Dr. Kuang Ching Hei, for sparing her precious time to make sure that I accomplish my research tasks on time and with quality. Thank you also for being a mother-like figure who cares about my well-being. Your personal and professional advices are greatly appreciated.

I am also very grateful of my wonderful friends who cared and scolded me in times I go off the right path. To Ali, Ernie, Ilde and Dedi, Arjay, Kenjie, Thevathy and Nithya, thank you for the stress-free happy-go-lucky crazy friendship. The Starbucks sessions we had created preciously expensive memories that even the holes it made in our wallets did not matter.

A sincere thank you also goes to Madam Sherly Sharif. For letting me conduct study at the college, you have become the instrument to fulfilling this study. I am also grateful for your motherly figure to us in Malaysia. Thank you for believing in my skills and expertise.

I am also grateful to the people I met and befriended at the university. Kaneson, Amira, Karima, Aishah, Chen, and many more, they are one of the reasons why my university study became colourful. We learn from each other, and this learning has never been boring.

Lastly, I thank University of Malaya for letting me study in this prestigious university. Thank you for pushing me to my limits; that even I had to sleep late to finish my projects, I know it benefits me now and in the future.

TABLE OF CONTENTS

ORIGINAL LITERARY WORK DECLARATION ii

ABSTRACT iii

ABSTRAK iv

ACKNOWLEDGEMENTS v

TABLE OF CONTENTS vi

LIST OF FIGURES xi

LIST OF TABLES xii

LIST OF SYMBOLS AND ABBREVIATIONS xiv

LIST OF APPENDICES xv

CHAPTER 1: INTRODUCTION 1-13

1.1 Background 1

1.2 Statement of the Research Problem 5

1.3 Significance of the Study 7

1.4 Research Objectives 9

1.5 Research Questions 10

1.6 Definition of Terms 10

1.7 Chapter Summary 13

CHAPTER 2: LITERATURE REVIEW 14-52

2.1 Introduction 14

2.2 Second Language Learners' Problems in Writing Accuracy 15

2.3 Malaysian Language Learners and their use of Basic English Tenses 20

2.4 Explicit and Implicit Knowledge in SLA 24

2.5 The Measurement of Implici and Explicit Knowledge 30

2.6 Disagreement in Corrective Feedback in SLA 34

2.7 Direct Corrective Feedback in SLA 38

2.8 Metalinguistic Corrective Feedback in SLA 44

2.9 Theoretical Framework of The Study 48

2.10 Chapter Summary 51

CHAPTER 3: RESEARCH METHODOLOGY 53-76

3.1 Introduction 53

3.2 Research Design 53

3.3 Ethical Considerations 55

3.4 Research Site 56

3.5 Subjects/Participants 57

3.6 Selection Criteria 58

3.7 Validity and Reliability 59

3.7.1 Test Validity 59

3.7.2 Test Reliability 60

3.8 Research Procedures 62

3.8.1 Phase 1: Quantitative Non-Equivalent Experimental Design 62

3.8.1.1 Testing Instrument 62

3.8.1.2 Data Collection and Treatment Set-up Operationalization 68

3.8.1.3 Target Structure 71

3.8.1.4 Data Analysis 72

3.8.2 Phase 2: Sequential Explanatory Qualitative Design 72

3.8.2.1 Interview Instrument 72

3.8.2.1 Data Collection 73

3.8.2.1 Data Analysis and Procedure 74

3.9 Chapter Summary 75

CHAPTER 4: RESULT AND DISCUSSION.....77-127

4.1 Introduction.....77

4.2 Effects of MCF and DCF on SLL's Use of Basic English Tenses in Writing 77

4.2.1 Picture Narrative Writing Test (PNT).....78

4.2.1.1 Descriptive Statistics of the Tree Groups in PNT 81

4.2.1.2 Within-Subject Effects between test groups and scores for PNT .82

4.2.1.3 Between-Subject Effects for PNT 84

4.2.1.4 Multiple Comparison Between Different Groups 85

4.2.1.5 Pairwise Comparison of Three Groups 86

4.3 Effects of MCF and DCF on SLL's Use of Explicit and Implicit Knowledge.....88

4.3.1 Metalinguistic Knowledge Test (MKT).....88

4.3.1.1 Descriptive Statistics of the Tree Groups in MKT.....90

4.3.1.2 Within-Subject Effects between test groups and scores (MKT) ...91

4.3.1.3 Between-Subject Effects for MKT.....93

4.3.1.4 Multiple Comparison Between Different Groups 94

4.3.1.5 Pairwise Comparison of Three Groups 95

4.3.2 Timed Grammatical Judgment Test (TGJT) 96

4.3.2.1 Descriptive Statistics of the Tree Groups in TGJT 98

4.3.2.2 Within-Subject Effects between test groups and scores (TGJT)...99

4.3.2.3 Between-Subject Effects for TGJT 100

4.3.2.4 Multiple Comparison Between Different Groups 101

4.3.2.5 Pairwise Comparison of Three Groups 102

4.4 Second Language Learners' View of the Received Corrective Feedback 104

4.4.1 Control Group.....104

4.4.2 Direct Corrective Feedback Group.....106

4.4.3 Metalinguistic Corretive Feedback Group 109

4.5 Analysis and Discussion of Quantitative and Qualitative Result 111

4.5.1 Discussing The Effects of Corrective Feedback on PNT Scores 114

4.5.2 Discussing The Effects of Corrective Feedback on TGJT and MKT 117

4.5.2.1 The TGJT and the Three Groups 120

4.5.2.2 The MKT and the Three Groups 121

4.5.3 Discussing The Participants' Experienced Experimental Procedure 123

4.6 Chapter Summary 127

CHAPTER 5: SUMMARY AND CONCLUSION, IMPLICATIONS AND RECOMMENDATIONS 128-135

5.1 Introduction 128

5.2 Summary and Conclusion 128

5.3 Pedagogical Implications 129

5.4 Limitation of the Study 132

5.5 Suggestions for Future Research 133

References 136

Appendices 155

LIST OF FIGURES

Figure 3.1: Study Design Paradigm 55

Figure 3.2: Sample Picture used in PNT 67

Figure 4.1: PNT Box plot results for 3 test groups 79

Figure 4.2: PNT Estimated marginal mean values of different groups 79

Figure 4.3: Box plot for MKT conducted at three different occasions 88

Figure 4.4: MKT Results of the marginal value plots for the three groups 89

Figure 4.5: Box plot for TGJT conducted at three different occasions 96

Figure 4.6: TGJT Results of the marginal value plots for the three groups 97

Figure 4.7: Control Group themes 105

Figure 4.8: Direct Corrective Feedback Themes 107

Figure 4.9: Metalinguistic and Handout Corrective Feedback Themes 110

Figure 4.10: Mean Performance Of The Three Groups 113

Figure 4.11: Mean scores of the Picture Narrative Writing test 114

Figure 4.12: Mean scores of the metalinguistic knowledge test 118

Figure 4.13: Mean scores of the Timed Grammatical Judgment test 119

LIST OF TABLES

Table 3.1: Students' demographics 58

Table 3.2: Content Validity Measurementa 60

Table 3.3: Timed Grammatical Judgment Test Reliability 61

Table 3.4: Metalinguistic Knowledge Test Reliability 61

Table 3.5: Picture Narrative Writing test Reliability 62

Table 3.6: Sample Metalinguistic Handout 66

Table 3.7: Schedule Information for the Different Group of Learners 69

Table 3.8: Sample Open-ended Questionnaire 73

Table 4.1:Different Tests of Normality 80

Table 4.2: Levene's test of equality of error variances 81

Table 4.3: Descriptive statistics for PNT 82

Table 4.4: Within Subject Effects between test groups and scores for PNT 83

Table 4.5: Within Subject Effects between different tests for PNT 84

Table 4.6: Tests of Between-Subject Effects for PNT 85

Table 4.7: Multiple Comparison between different groups for PNT 86

Table 4.8: Pairwise Comparison for PNT 87

Table 4.9: MKT Levene's Test 89

Table 4.10: Descriptive Statistics for MKT 90

Table 4.11: Within Subject Effects between test groups and scores for MKT 91

Table 4.12: Within Subject Effects of three groups for MKT 92

Table 4.13: Tests of Between-Subject Effects for MKT 93

Table 4.14: Multiple Comparison between different groups for MKT 94

Table 4.15: Pairwise comparison between different tests and groups for MKT 95

Table 4.16: TGJT Levene's Test 98

Table 4.17: Descriptive Statistics for TGJT 99

Table 4.18: Within Subject Effects between test groups and scores for TGJT 99

Table 4.19: Within Subject Effects of three groups for TGJT 100

Table 4.20: Tests of Between-Subject Effects for TGJT 101

Table 4.21: Multiple Comparison between different groups for TGJT 102

Table 4.22: Pairwise comparison between different tests and groups for TGJT 103

Table 4.23: Descriptive Statistics for all of the tests and groups 112

LIST OF SYMBOLS AND ABBREVIATIONS

ANOVA	:	Analysis of Variance
CEFR	:	Common European Framework of Reference
CG	:	Control group
EFL	:	English as a Foreign Language
ESL	:	English as a Second Language
DCF	:	Direct Corrective Feedback
FLA	:	First Language Acquisition
L1	:	First language
L2	:	Second Language
MCF	:	Metalinguistic Corrective Feedback
MKT	:	Metalinguistic Knowledge Test
SLL	:	Second Language Learners
TGJT	:	Timed Grammatical Judgment Test
PNT	:	Picture Narrative Writing Test

LIST OF APPENDICES

Appendix A: Letter of Permission and Consultation 155

Appendix B: Proposal Defence Approval 157

Appendix C: Candidature Defence Approval 158

Appendix D: Test Reliability 159

Appendix E: Two-Way Mixed Method Anova Result 161

Appendix F: Sample Tests 168

Appendix F: Sample Metalinguistic Handout 171

Appendix H: Statistical Method Consultation Reference 172

CHAPTER I

INTRODUCTION

This chapter provides the introduction for the conducted study. It sheds light on the salient features of this research by posing relevant research questions, which have been academically pursued and investigated in the different sections of this study. Prior to presenting research questions, the introduction provides an initial background of the study, providing the background and the problem statement. It also discusses the significance of the study and its contribution to language learning and to the academic institutions. Aside from that, important terms have also been defined to understand the concept throughout the study. The last section of this chapter introduces some initial information regarding the context and basic features of the research methodology, which has been elaborated in depth in chapter 3.

1.1 Background

The research area of Second Language Acquisition (SLA), English as Foreign Language (EFL), Interaction Research and English as a Second Language (ESL) has greatly expanded in the past few decades, owing to the increased global emphasis on English language learning for individuals with very limited exposure or no access at all to the language being taught. The literature on language learning outline a wide array of different types of methods for improving learners' ability to acquire and learn new language skills. Some of the different recently proposed methods include autonomous ESL learning (Chou & Chan-Lin, 2015), mobile-assisted language learning (Soleimani, E., Ismail, K. & Mustaffa, R., 2014), formative assessment for learning (Sardareh & Saad, 2012), brainstorming (Unin & Bearing, 2016) and the varying direct and indirect feedback-based language learning techniques (Mackey, Oliver & Leeman, 2003; Loewen, 2004; Ellis, Loewen & Erlam, 2006; Farid & Abdul Samad, 2012), to name a few.

Prior to 1970s, ESL and EFL teaching methodologies stressed on the importance of behaviourism and structuralism for teaching language writing to students (Raimes, 1991). However, the emphasis has shifted considerably with the introduction of the interactionist perceptive within SLA (Long, 1996; Gass, 2003). This emphasized the role of interaction between the learners and teachers in the classroom, specifically in language acquisition (Long, 1980, 1983a; Varonis & Gass, 1985; Pica, 1987; Mackey, 1999). In terms of SLA and language learning, errors have been deemed as natural part of the learning process as they allow researches and teachers to have a better insight into the processes underlying language acquisition (Hendrickson, 1978). Corrective feedback is one of the methods of ensuring that students are able to learn from their mistakes within language learning courses. Corrective feedback can be termed as the immediate response of the teacher to the learners' error (Gitsaki & Althobaiti, 2010). This allows learners to rectify their mistakes and improve information retention which could prevent the repetition of those mistakes in the future. By providing feedback, it supports language acquisition and prepare students for the practical world, where accuracy of writing is given considerable importance (Ferris, 2011).

A vast amount of literature (both descriptive and experimental) has been devoted to examining corrective feedback and its various aspects that include different types of feedback (Lyster, 1998a; Lyster & Ranta, 1997; Eslami, 2014), their overall effects on language learning (Oliver & Mackey, 2003; McDonough, 2005; Lyster, 2004; Ellis et al., 2006; Loewen & Philp, 2006), as well as learners' perception and uptake of feedback (Mackey, Gass & McDonough, K., 2000; Mackey & Oliver, 2002; Sheen, 2007; Egi, 2010). Some of the most widely used identified feedback techniques which are adopted by various researches (Sheen, 2004; Ellis, 2009; Lee, 2013) were originally proposed by Lyster and Ranta (1997) which divide them into explicit correction (this indicates the error, its type as well as the correct form), elicitation (prompt by a teacher to allow the

learners to fill the erroneous part with the correct form), repetition (emphatically stressing error through repetition), meta-linguistic feedback (provide technical information without directly providing correct answer), clarification request (request to reformulate the phrase or sentence) and recast (reformulate incorrect part without identifying the error).

Nevertheless, the advent of corrective feedback and its efficacy in second language acquisition (SLA) has been in a positional nature. Disagreements started when Truscott (1996) penned a review that warned negative impacts of feedback in SLA. Using evidence from a series of different researches (e.g. Kepner, 1991; Semke, 1984, and Sheppard, 1992), Truscott (1996, 1999, 2004, 2007, 2010) contended that grammar correction should not be used within language writing courses, while proposing the following reasons to defend his viewpoint: (i) grammar correction has been deemed as a time-consuming process, (ii) inconclusive evidence linking error corrections with improvements in the writing skills of students, and (iii) the supposition that error correction through feedback benefits language learning (especially errors related to grammar) has been termed as erroneous in nature, as it lacks theoretical and empirical justification. Truscott's (1996) rigid stance pertaining to the efficacy of error correction was received by a barrage of studies that negated his viewpoints and attempted to provide relevant evidence to empirically refute his claims (Bitchener, Young, & Cameron, 2005; Ferris, 1999; Ferris & Hedgcock, 1998; Rezaei, S., & Derakhshan, A., 2011). To date, many studies continue to investigate its efficacy in writing and oral production (Bitchener & Knoch, 2009; Ellis, 2009; Eslami, 2014; Lyster, 1994; Stefanou & Revesz, 2015). However, dearth in the studies has been observed in using error coding plus metalinguistic explanation as a feedback in writing.

In a addition, the stringent remarks of Truscott (1996) on error correction and feedback efficacy has been debated by various researchers, particularly by Ferris (1999). Ferris (1999) believed that it was wrong to completely disregard error correction, as it has

been a valuable pedagogical tool that has enabled learners to improve their overall writing skills. It was further asserted that the overall effectiveness of error correction was dependent on the quality of correction. This refutes Truscott's (1996) assumptions and were described as premature and overly rigid in nature (Ferris, 1999). The observations outlined by Ferris (1999) have also been supported by various other researchers. Polio (2012) and Bitchener (2012) believed that corrective feedback, not only improved writing accuracy, but it also enhanced the development of explicit knowledge. Similarly, Chandler (2003) revealed that Truscott (1996) failed to consider the fact that statistically sound empirical evidence is required in order to prove error corrections' lack of worth in terms of improving learners' writing skills. It is also important to remember Ferris' (1999) suggestion for future research which emphasized the need of thorough investigations of different techniques, methods and approaches to error correction in order to come up with a valid and credible result. This includes deep analysis of short- and long-term improvements brought on by each of the correction techniques employed in research studies independently and collectively.

Corrective feedback is also claimed to improve language learners' accuracy in writing. Using metalinguistic corrective feedback, Sheen (2007) claimed that it is helpful in improving the language accuracy of the students; however, with specifications to students with high English proficiency. Rassaei, Moinzadeh, and Youhannaee (2012) have also studied MCF and claimed its efficacy in improving explicit and implicit knowledge. However, some contradicting results were also found in other studies. As an example, Sanz (2003) was not able to find any significance in the performance of the metalinguistic feedback group when compared to other groups who received different feedback procedure. Hence, this calls for validation and the need a clearer perspective on how to improve writing, and develop explicit and implicit knowledge of the common error of students learning a language (Abdullah, 2013).

1.2. Statement of the research problem

Accuracy in writing is one of the current problems in Malaysia, particularly in tenses, articles, and other grammatical features (Maros, Huan & Khazriyati, 2007; Abdullah, 2013). Thus, many teachers and researchers conducted studies to address this problem, specifically what and how to give correction (Velayutham, 2013). However, most relevant studies have only delved into common feedback such as direct and indirect feedback. The former, according to Ferris and Roberts (2001), is only beneficial for short-term acquisition. The latter, however, does not allow learners to notice the target language structure (Frear & Chiu, 2015). Moreover, feedback is also claimed to promote explicit and implicit knowledge (Rohollahzadeh Ebadi, Mohd Saad, & Abedalaziz, 2014). The blatant reproach of error correction techniques by Truscott (1996, 1999, 2007, 2010) has significantly created doubts in the academic community, specifically with regards to the actual efficacy of feedback-based strategies within the language learning perspective. Despite researches claiming learning benefits of corrective feedback techniques (Hosseiny, 2014; Li, 2010; Mackey & Goo, 2007) as well as additional clarifications and criticism to Truscott's viewpoints by various researchers (Bitchener et al., 2005; Lyster, 1998; Ferris, 1999; Ferris and Hegcock, 1998), it is important to examine the empirical evidences (Chandler, 2003; Ferris, 2006; Ferris & Helt, 2000; Lee, 1997; Kepner, 1991; Kim & Mathes, 2001; Sanz, 2003; Semke, 1984) that failed to reveal the efficacy of corrective feedback techniques on language learning.

Nevertheless, results were varied due to differences in methods and instrumentations, thus further test validation is highly recommended (Rassaei et al., 2012). Kassim and Ng (2014) also emphasized on the need to study feedback in a different linguistic feature as it cannot be generalized in a single study. Hence, introducing MCF through error coding in another grammatical item will be beneficial for teachers and learners in SLA. Sheen (2007) and Velayutham (2013) claimed that MCF has improved students' writing, as it

allows them to self-edit and reflect on their work (Ferris, 2006). In view of this aforementioned evidence, the importance of conducting research to test and validate the efficacy of corrective feedback techniques (e.g. direct corrective and metalinguistic feedback) has been repeatedly highlighted, which ultimately necessitates investigation to clarify some of the existing issues and remove doubts for future researchers that currently plague this research area.

Nevertheless, the issue of language acquisition in the research site has also been observed by the teachers and researchers. The research site accommodates Malaysian learners who were hardly exposed to English language. Many of the students were from rural areas, and access to English communication, aside from online media, was limited. The Head of the Department with some lecturers, through casual conversation, also claimed that students in the beginner level are weak and have been facing difficulty in developing their writing skills. It has also been observed that the college's lecturers were only familiar in the traditional approach of error correction, and using metalinguistic feedback is unknown or has never been an option. Additionally, as part of its revised curriculum and as a requirement of Common European Framework of Reference, students enrolled in this level are required to develop specific language competency and acquire specific grammatical feature such as the basic tenses, English articles, prepositions, and other simple word class features. Accuracy in these areas is deemed important and is necessary to pass and obtain international certification.

As such, facing issues in the literature and the language concerns of the target site, this study was conducted in attempt to validate previous results and introduce corrective feedback to enhance language acquisition.

1.3. Significance of the Study

Several research studies aimed to provide solutions on how to improve writing accuracy of students in SLA. To become a proficient language user of second language, a good knowledge of vocabulary and grammatical rules need to be possessed. While Malaysia was once a colony of England, the problem in using English language in written and verbal communication never ceased. As an evidence, there are 400,000 Malaysian graduates, mostly Malay, who cannot communicate properly in written and oral English language which resulted to their failure in securing employment in local and international companies (Hussaini Abdul Karim , 2016). Furthermore, it has always been the goal of Prime Minister Najib Razak to alleviate the rising problem of English communication among Malaysian learners (Naidu, 2015). The emphasis of being the language of the world urges the Malaysian government, especially the Prime Minister, to motivate learners, increase their confidence in speaking the language, and provide them opportunity to apply it in their daily lives (Naidu, 2015). Nevertheless, while most of the learners in the urban areas, particularly in big cities, have an ease of access to English language, learners in remote or rural areas suffer from a limited resource of English language (Kamalanathan as cited in Naidu, 2015).

In the hope to provide solutions to the problematic issues in English, experts in language learning and acquisition have investigated various empirical studies rooted to the very aim of improving learners' ability to learn a second and/or foreign language. Corrective feedback, as studied and proposed by many experts (i.e., Al-jaarah, 2016; Ellis, 2009; Motlagh, 2015), has been found to improve the accuracy of students' use of target language, and improve their acquisition of implicit and explicit knowledge (Shinatni & Ellis, 2013; Rassaei, Moinzadeh & Youhannaee, 2012*)*. However, the differentiation in results highlighting opposing views on corrective feedback efficacy has made considerable effects in the area of SLA. Several studies were conducted in the hope

to validate past results, however, one finding leads to another (i.e Al Ajmi & Ahmed Ali Saleh, 2014; Ferris 1999; Truscott 1996, 2004,2007).

In this regard, this study is seen as an important opportunity to fill the gaps and contribute to the body of knowledge through involving the traditional corrective feedback over the less researched metalinguistic corrective feedback with metalinguistic handout (Motlagh, 2015; Rezaei & Derakhshan, 2011). A careful selection of corrective feedback is imporant in order to promote independent learning among students, and promote accuracy in written and verbal production. As cited by Azrinda (2013), Malaysian learners continue to seek effective strategy to improve written accucacy in English as demanded by Malaysian pubclic examinination, especially the Sijil Pelajaran Malaysia (SPM). With this study, we could provide an innovative teaching and learning process to learners and teachers by employing an empirically effective corrective feedback as has been evident in the results of this study. This also helps teachers reflect their use of corrective feedback and its suitability in enhancing learners' didactic errors. By enhancing grammatical competency, the long-term effects would lead to a lessened, if not error-free, writing productions in SLA.

Apart from that, as teachers in writing continuously seek effective teaching writing techniques, the outcome of this study could also ease teachers' burnout. Teachers' ultimate goal to improve learners' accuracy in writing has always been a dilemna (Azrinda, 2013), specifically responding to the many erros has been tedious and time consuming. Thus, these teachers hope to provide an efficient yet corrective effective feedback. To operationalize a new corrective feedback inside the classroom, applicable in different settings, may answer their plight to solve this never ending issue in classrooms all over the world. Nevetheless, the output of this research would offer a great and practical contributiion to influence the teaching practices of the research site, which may and could be applicable in the education system of Malaysia and in other countries. To

elaborate, this study may be a starting point for the administrators and policy makers to consider the kind of corrective feedback to be given in response to a particular learners' linguistic error, and ultimately consider learners' uptake on corrective feedback as what has been also undertaken in this study.

1.4. Research objectives

There is a plethora of existing evidence in terms of assessing corrective feedback techniques and their application to different educational settings for learners with varying contexts and backgrounds. However, differentiated evidences from different researchers and their studies lead to confusion and inconsistencies within the literature. One of the primary objectives of this research is to thoroughly examine the different perspectives, learn from the highlighted shortcomings and limitations of existing researches and focus on formation of research design and methodology. The study also aims to compare the effects of metalinguistic and direct corrective feedback, and gain insights on their efficacy in students' accuracy in writing, and in the explicit and implicit knowledge acquisition of basic English tenses. Consequently, this research provides evidence that could contribute effectively to the debate regarding actual benefits of corrective feedback on language learning and acquisition and provide data and information from non-native English learners in Malaysia. This can assist and guide future research in the realm of interaction research and ESL. Nevertheless, this study provides evidence from the Malaysian non-native English language learners' perspective, while presenting latest evidence, which would add to the existing evidence to either support or refute the importance of corrective feedback techniques for language learning.

1.5. Research questions

This is one of the most crucial aspect of the research as it allows identification of the overall expectations and outcomes of the study. In order to facilitate that, research questions are posed at the beginning of the research so it can assist in guiding and focusing the research efforts in the designated direction. In any case, this study aimed to arrive with answers for the formulated questions below:

1. What are the comparative effects of focused Direct Corrective Feedback (DCF) and Metalinguistic Corrective Feedback (MCF) on the learners' accuracy of English tenses in writing?

2. What are the effects of focused DCF (direct corrective feedback) and MCF (metalinguistic corrective feedback) on the learners' explicit and implicit knowledge acquisition of the basic English tenses?

3. How do students view the teacher's use of corrective feedback?

1.6. Definition of Terms

Throughout this study, the thesis makes use of important terminologies that help us better understand concepts in SLA. Most of the terminologies are provided with brief detail, while others are given further attention in the rest of the chapters. As such, to avoid misconceptions, the outline below provided could give us a clear direction of the study. Terms have been defined with appropriate reference and on how these were used in the study.

1. **Feedback**: Given information regarding contact, perception, understanding, and attitude, communicated through standard linguistic means (Allwood, 1992). In this study, it refers to an evaluative information given to students in their writing output.

2. **Corrective Feedback (CF)**: An indication of correction for learners' errors in verbal of oral tasks (El Tatawy, 2002). This also refers to teacher's strategies in providing error correction to students (Ellis, 2012).

3. **Written Corrective Feedback (WCF)**: In this study, it refers to feedback given to students to correct students' mistakes in written production (Jimenez, 2013). This includes all reaction, including comments of the teacher to students' write-up from draft to final revision (Ferris, 2002).

4. **Control Group (CG)**: This acts as a baseline of measurement in an experimental research (Dictionary.com, 2015). In this study, it is the group that does not receive experimental manipulation. The assumption points that if this group performs lower than experimental groups, the feedback provided is deemed effective.

5. **Experimental Group**: "groups of subjects who are exposed to the variable of study" (Dictionary.com, 2015). In this study, it is the group of participants that receives metalinguistic and direct corrective feedback as interventions.

6. **Error:** This refers to learners' obvious deviation in use of the target grammatical features as opposed to the standard system (Brown, 1994). In this study, errors committed by experimental groups in their use of English basic tenses are provided with corrective feedback.

7. **Beginner's Level of English proficiency:** Learners who have very limited English and cannot produce oral and written productions independently. According to CEFR ("Levels of the Common European Framework of Reference", 2016), learners in this particular level can use very basic phrases to meet every day demands. They can also interact, however pacing must be slow and very clear. In this study, this refers to students who fall into

primary/beginners level based on research site's placement test with reference to CEFR.

8. **Basic English Tenses:** These are the simple tenses that are required by the CEFR guideline for the Beginner English level proficiency: present simple, present progressive, past simple, and past progressive.

9. **Direct Corrective Feedback (DCF):** Teachers provide the correct usage form of the specific target language in response to the errors of the learners (Ellis, 2009). In this study, direct correct feedback is given to students who made errors in their use of basic tenses in English.

10. **Metalinguistic corrective feedback (MCF):** Feedback is given without providing the correct form (Ellis, 2009). In this study, error codes served as an intervention to one experimental group.

11. **Metalinguistic Handout:** This provides an explanation on grammatical rules as used in the study of Rezazadeh, Tavakoli, & Rasekh (2015) and Shintani and Ellis (2013). In this research study, metalinguistic handout provides description of the rules in using the basic English tenses. The handout also contains error codes matched to the metalinguistic corrective feedback.

12. **Implicit Knowledge:** This refers to unconscious knowledge where learners rely mostly to their intuition, feeling, and procedural knowledge (Rassaei et al., 2012). In this study, implicit knowledge is solicited from the students by answering timed grammatical judgment test.

13. **Explicit Knowledge:** This refers to the conscious knowledge that learners use to access when needed. Learners who use explicit knowledge rely on declarative knowledge (Hulstijn, 2003). In this study, explicit knowledge makes learners remember grammatical rules in answering the metalinguistic knowledge test.

1.7. Chapter Summary

This chapter provides the overview of the study. Through outlining past studies and presenting the current situation of language learning and acquisition in SLA classrooms in Malaysia and in other countries, the study structures the need of conducting this research study. It also provides the significance to the improvement of language learning processes that could enhance the current practices in this research area. The problem statement details the necessity of conducting the study. While the traditional corrective feedback and the less research corrective feedback have gaps in literature, the current situation at the research site and in other Malaysian classrooms calls for an innovative strategy. It calls for a strategy that could establish a strong foothold for learners to achieve deeper understanding of target language structures and use them with accuracy in verbal and written tasks.

As the goals of this research provide, this study compares the differential effects of two corrective feedback types in learners' writing tasks to strengthen previous findings regarding their efficacy, and their relationship in building explicit and implicit knowledge in the target structures. Nevertheless, definition of terms is provided to understand any technical terms that are used to introduce concepts in SLA. With this understanding, the course of the study would be much easier, especially if the study will be read and replicated.

CHAPTER 2

LITERATURE REVIEW

2.1. Introduction

This chapter examines all the relevant literature directly linked to the research area being studied. This is to gain valuable insight from some of the leading researchers in the field. Some of the most reliable research sources have been read and consulted to provide academically, theoretically and empirically sound and valid information. The present study has also examined current researches dealing with problems faced by second language learners in writing accuracy with an emphasis on verb tenses. This research highlights Malaysian language learners and the difficulty they had experienced with the Basic English tenses in writing. The Basic English tenses include the present simple, past simple, present progressive, and past progressive. Following are some of the concepts which have been used to analyse several findings in the present study: (i) the two types of knowledge (implicit and explicit knowledge) within the research realm of language acquisition, (ii) disagreements with respect to usage of corrective feedback techniques in second language acquisition (SLA), (iii) usage of direct corrective feedback in SLA, (iv) metalinguistic corrective feedback in the research area of SLA (error coding) and (v) the theoretical framework detailing the salient features of Noticing and Output Hypotheses (Swain, 1985; Schmidt, 2001) in second language learning and acquisition. An examination of the previously stated thematic developments put forth by each respective study on the value of metalinguistic corrective feedback over direct corrective feedback in language acquisition can answer the following question: What existing research is able to support the efficacy of corrective feedback (CF) in language acquisition?

2.2. Second Language learners' problems in writing accuracy

This section details several studies that have utilized different strategies to evaluate the problems that second language learners encounter in writing accuracy with a focus on verb tenses. The analysis provides the precedence and benefit of corrective feedback towards SLA specifically for written English grammatical problems.

In a study conducted by Mourssi (2013), he analysed the overall efficacy of error analysis of Arab second language learners of English language. The study revealed the ability of learners to highlight complexities in grammar within the written text during error identification and analysis tasks. According to him, it is the responsibility of the teacher to assess the type of errors being made by the learners and provide relevant feedback, especially when the learners are unable to produce target-like form of tense. His further investigation revealed that the main errors in verb tense are inter-lingual, intra-lingual and in-between errors (Mourssi, 2013). He also concluded that Arab learners at intermediate and pre-intermediate levels initially composed their thoughts in Arabic before performing the assigned task in English.

Abdullah (2013) also supported Mourssi's (2013) findings. Abdullah (2013) was able to establish the need for applicable research in the field of SLA. His research focused on using basic tenses from English language, along with specifying the type and location of the errors in the written text. The findings of his study corroborated to Ellis's (1990) behaviorist learning theory. This theory suggested that old habits remain a big influence in specifying the manner in which learners are able to develop new habits. As a result of the pre-conceived learning patterns of the first language developed within the minds of the language learners, the grammar rules learnt from the first language hinder and interfere with the acquisition of rules concerning the target language (Abdullah, 2013). This interference is directly the result of proactive inhibition, which essentially prevents the process of acquiring skills and learning habits related to the target language. For

second language learners, the first and second language share some level of meaning but this is often expressed much differently. Thus, it results in errors within SLA because learners typically fail to transfer the realization device from the first language to the target language (Ellis, 1990). Nevertheless, Abdullah (2013) found subject-verb agreement (tenses) as a common error of the students. The result also identified parts of speech and vocabulary as one of the problems among the participants. Abdullah (2013) attributed the causes of errors to incompetence in the target language, L1 interference, use of loan words, limited exposure to the target language. As a remedy, some suggestions include giving more practice, and devising new teaching techniques to facilitate errors of the students (Abdullah, 2013).

Thai students were also investigated by Sukasame, N., Kantho, S., & Narrot, P. (2014) on their use of English language. They collected data on the grammatical errors in learning English on tenses using three distinct tools of measurement. The tools include multiple choice tests examining grammatical errors, a table that recorded each students' errors and a survey that interviewed students who made errors. The results of the findings revealed that: (i) 87.1% of students had made errors in using the past perfect tense, (ii) 74.2% of students made errors in the past simple tense, (iii) 67.4% of the students made errors in the present perfect tense, (iv) 54.8% of the students made errors in the past continuous tense, (v) 48.4% of the students made errors in the present simple, (vi) 41.7% of the students made errors in future simple tense, and (vii) 32.3% of the students made errors in the present continuous tense. The findings of the case study indicate that Thai students have considerable difficulty with tense selection. Tenses may be used comfortable by the students, but many instances proved that there was a lack of confidence in selecting a specific tense. Nevertheless, tense is considered to be the most valuable aspect of the English grammatical structure, but also the most difficult to understand and use correctly. It has been shown that Thai students have little background

knowledge about the English language (Sukasame, Kantho, & Narrot, 2014). Another main reason for the persisting error rate of Thai students is due to the inability of the learners to realize when they had committed grammatical or tense errors. This consequently impedes their overall progress towards language acquisition (Sukasame, et. al., 2014). The insightful result provides a picture of mother tongue influence and the impact it has on the learning process. If learners of a target language are unable to realize their error, then they cannot be expected to improve their overall language acquisition progress. Hence, pointing out errors and using proper intervention can improve learners' performance of language use (Sukasame, et. al., 2014).

Some studies also focused on the formal characteristics of second language learning processes (Paradis, Crago & Genesee, 2011), while others such as Mariko (2007) proposed developmental indices for evaluating language learners' progress over time. Mariko (2007) investigated the first aspect of grammatical research in second-language acquisition on Japanese learners of English language in terms of noun and verb related errors. The researcher acknowledges that many prior studies have been limited by small sampling sizes and thus put forth a substantial amount of written and spoken data to detail the differences between spoken and written production (Mariko, 2007). Using the Standard Speaking Test (SST), the results of this test were used to categorize the students into their respective proficiency levels for the English language. This consisted of five stages starting from warm-up questions, picture description task, role-playing task, picture sequence task and wind-down questions. The researcher noted the accuracy rate for the parts of speech which ranged 90% to 100%. Although its accuracy rate is high, it was also determined that most of the sentences are simple, thus contributed to the accuracy of the lower level students. Moreover, the study has identified that verb-related errors could be expected to be applicable only for learners in the lower level, while noun-related errors are for those students in the higher level. It has also been suggested that

when teachers recognize and understand the error, the better they become in providing effective instruction in the classroom (Mariko, 2007). This supports prior Sukasame et al's study that highlighting errors for students makes it easy for them to understand the target language.

Studying verb forms is considered to be a challenging aspects of English language (Cowan, 2008). Second-language learners are typically unable to employ verb forms in a correct fashion, specifically when dealing with expression in terms of specifying time duration of any action within the English language (Cowan, 2008). In a study reported by Rahman and Ali (2015), Bengali learners of English had issues in using past tense forms of verb. The research describes this occurrence as a phenomenon, which was first depicted by Cowan (2008) in his study of German second language learners. Due to the variations in the use of tenses between German and English language, the German second language learners were having troubles conforming to English language tense. This is because German simple sentences conform to both English present simple and present progressive only (Cowan, 2008). Nevertheless, the instructors are given the opportunity to help learners adjust and learn from the common errors to improve their English language skills. As detailed by Rahman and Ali (2015), errors made are a result of the influence from the learners' native language. Thus, these differences must be addressed by the instructor in order to produce the highest possible achievement rates. Exposing them to the contrast between these two languages would help, and the provision of instructional activities within the context could facilitate in addressing this issue (Collins, 2007). It also follows that course content and second-language acquisition materials should emphasize these differences between the second-language and native language (Rahman & Ali, 2015). Learners must also be made aware of the specific areas that contain consistent errors due to the influence of the native language. Based on evidence provided by Rahman and Ali

(2015), it can be concluded that learners tend to overgeneralize the actions having progressive aspect (Rahman & Ali, 2015).

Aiming to understand why students commit errors in writing, Salima (2012) also conducted a study among teachers to understand the errors by the English language learners. Through observation and questionnaires, it revealed that 60% of the students neglected the idea of good writing skills, together with their poor level of skills in writing. Also, with learners' poor background knowledge of English, they are not motivated to practice or write in the target language. As a remedy, the study recommended the provision of more input, more practice, and provide immediate feedback for corrections and revisions. Harmer (2001) also agrees the necessity of feedback so students will understand their mistakes, and eventually correct them. Harmer (2007) also introduced the use of symbols as feedback or the use of underline so it would appear "less damaging" (Harmer, 2007, p. 121).

Since understanding of errors proved to be significant in lessening mistakes in writing, Kirgoz (2010) investigated the errors of 86 Turkish students with primary level of English proficiency. In a corpus of 120 essays, he classified the errors to determine its possible sources. As a result, most of the errors have been attributed to the interference of the mother tongue. It was also found that simple present tense is used in general situations; however, it is being expressed in present continuous in Turkish. Hence, feedback may be beneficial to lessen the error and overcome their learning problems (Kirkgoz, 2010). Teachers can also use symbols as a means of error correction, providing that symbol meaning is introduced and explained to the learners. It is also important to note the stand of Corder (1967) regarding errors. Corder (1967) provided two serving purposes of errors in language learning: (1) diagnostic-to diagnose the learners' level of language, and (ii) as a prognostic- informing teachers how to design their materials to meet the needs of the learners.

In summary, the above outlines issues in second language writing, particularly the grammatical errors of students with different backgrounds in different contexts. Errors have been identified to be beneficial in learning as it provides opportunity for learners to determine their gaps and helps teachers to effectively design methods and techniques to improve these situations.

2.3. Malaysian Language learners and their use of Basic English Tenses in Writing

Malay language is one such language that shares few similarities in structure to the English language. In fact, Malaysia was originally a British colony; all students there study English from preschool until the fifth form of secondary school (Bakar, K. Hamid, N. Z., Mat Awal, N. & Jalaluddin, N. H.,2007). This made many researchers become interested in how Malaysian students learn English, and how their English skills may be improved (Ghabool, Mariadass, & Kashef, 2012; Hijjo, 2013; Stapa & Izahar, 2010; Zainal, 1990). Thus, this section describes several research studies that have investigated the Malaysian language learners and their use of basic English tenses in writing. Specifically, this section details evidence of Malaysian learners having problems with past, present, present continuous and past continuous tenses.

Studying the errors is a vital process of educational studies as it constitutes language acquisition and production of language by learners in both oral and written forms (Noor, 1985). In a study way back in 1980s, Noor (1985) examined the verb tense problems that Malay speakers encounter in writing their exercises. Using Contrastive Analysis, she identified two of the most problematic grammatical feature for the native Malay speakers. These are the simple present and the simple past tense. Her study also reported of some students who attempted to use higher/complex tense forms but failed to use them appropriately. Difficulty in acquiring these tenses was attributed to the interference effect of the first language when writing in the target language. Suggestion was given to concentrate on teaching these tenses, especially in discourse application,

with the provision of appropriate feedback (Noor, 1985). Similarly, Maros, Hua & Salehuddin (2007) examined the interferences in SLA of English language among Malay students, specifically the written grammatical errors. Through error analysis, they were able to reveal that despite having six years of English language experience, the students still showed difficulty in using correct use of articles (21%), subject-verb agreement (46.8%), and copula 'be' (30.8%). Understanding both the linguistic and nonlinguistic sources of the errors, while relying on both contrastive and error analyses, would be beneficial for the instructors to understand their learners and design a remedy (Maros et al., (2007).

Considering implicit and explicit knowledge in the target language, Loftie, Salleh, & Kadir (2015) reported on 72 Malay graduate students' production of past-time inflections (-en, -ed) and its effect on linguistic knowledge using three tests: the Grammatical Judgement Test (GJT), Metalinguistic Knowledge Test (MKT), and Written Production Test (WPT). The general findings revealed that implicit knowledge had adequate impact on learners' output; similarly, the analysis also revealed positive correlation between implicit and explicit knowledge (Loftie at al., 2015). Nevertheless, it was highlighted in the study that Malay language had limited usage when it comes to inflecting past-tense of word forms. This suggests the learner's difficulty in acquiring the target grammatical feature. Pedagogical approach has also been suggested by including implicit and explicit instruction and use of relevant materials to fill in the learning gaps of the learners (Loftie at al., 2015).

In the case of past tense, Manokaran, J., Ramalingam, C., & Adriana, K. (2013) explored a corpus of argumentative essays and studies on how Malaysian learners used past tense in their writing, particularly the auxiliary 'be'. Findings indicated the seven types of errors committed including the tense shift and agreement. As suggested, the findings of their study can be used to improve teaching practices in the classroom,

particularly the approach in teaching writing. Material development should also be taken into consideration to target the errors of the students.

While the study of Loftie et al. (2015) was concerned on the Pas-time inflections of English language, a morpho-syntactic analysis of the present tense with inflection -s was studied by Hijjo (2013). Highlighted in the study was that morpho-syntactic issues are important topics that need to be discussed to clear out any misconceptions in this particular domain, and consequently improve our knowledge on language rules (Hijjo, 2013). In his study, it was found out that Malaysian secondary school students were not aware of how to use to the -s marker for present tense and noun plurality. Likewise, this issue has been attributed due to the influence of their native language, particularly the word order and sentence structure of Malay language. Such interpretation was similar to the previous studies (e.g. Noor, 1985; Maros, Hua and Salehuddin, 2007; Loftie, Salleh, & Kadir, 2015).

Nevertheless, since most of the relevant studies above investigated Malaysian secondary school students, it is also important to explore similar issues with adult learners, specifically post-graduate students. This would help us understand if the discussed issue is applicable to higher level students. This would help us deepen our understanding on why such cases occur and what solutions can be suggested. Stapa and Izahar (2010) studied twenty post graduate students enrolled in a teacher training college. Tasked with two types of compositions, their writing was analysed using error analysis as patterned to the framework of Corder (1967). Parallel to the results in the previous studies on Malaysian secondary students, the post graduate students committed majority of their errors on subject-verb agreement (SVA). It was also found out that these participants were conscious of how they make use of SVA complex sub rules, particularly with indefinite forms and use of complex subjects with dangling modifiers. It is also important to note in this study that the participants' level of proficiency was not indicated

in the study. Proficiency level serves as an important factor to complete this task, therefore it should be considered and mentioned.

Most of the discussions above have been concentrated to how students make errors in their second language learning; however, in the paper presented by Musa, Yew Lie, and Azman (2012), it aimed at identifying the problems encountered with teaching English in Malaysia. This includes various stakeholders in the education sector such as language learners, teachers and policy makers with respect to the teaching methods being practiced for English language learning and acquisition. Findings concluded that the national language or Bahasa Malaysia had a considerable amount of linguistic influence on the learning of English among the native learners. It has also been highlighted that the learning system emphasized on rote-learning and the mastery of specific language skills is tested using standardized examination (Musa, et. al., 2012). It was also suggested that a curriculum based on inclusiveness and active participation of learners with reflective learning pedagogy would encourage a more meaningful learning development (Musa et al., 2012). This provides a change in the performance of the students in and out of their classroom. The researchers further indicated the need of using corrective feedback strategy inside the class. This will allow learners to linguistically evolve from first language to the target language by addressing the students in a manner that satisfies the learners' interest and willingness to engage in second-language learning (Musa et al., 2012).

In summary, this section provided evidences of the issues of Malaysian learners in writing accuracy, specifically when dealing with English tenses. Recommended solutions include error correction and feedback, understanding learners' background, devising new teaching techniques, and incorporating relevant activities and materials in the classroom.

2.4. Explicit and Implicit Knowledge in SLA

The concepts of implicit and explicit knowledge have a central position within the context of language learning and acquisition. Although it became important in SLA, a vast amount of literature from diverging fields of research (e.g. cognitive psychology, second language acquisition and neurobiology, to name a few) has been dedicated to distinctly separate the two types of knowledge (Dienes and Perner, 1999; Paradis, 2004; Ellis, 2005; Williams, 2009; Rebuschat, 2013). Nevertheless, the contribution of each type of knowledge towards language learning and acquisition has never been confirmed, as language learners possess both types of knowledge when it comes to learning the target language (Ellis, 1994; Hulstijn and Ellis, 2005; Ellis et al., 2009; Sanz and Leow, 2011; Rebuschat and Williams, 2012). One of the ways of improving language acquisition and these types of knowledge is through using different corrective feedback techniques (Lyster, 1998; Mackey, 2002; Oliver and Mackey, 2003; McDonough, 2005; Lyster, 2004; Ellis et al., 2006; Sheen, 2006). However, their overall efficacy remains a heavily debated and contested claim, where at one end, different evidences support the importance of feedback (Sheen, 2006; Bitchener, 2012; Bitchener and Knoch, 2008, 2009, 2010a, 2010b; Polio, 2012). On the other end, different sources have failed to find adequate empirical and theoretical support in relation to use of corrective technique on SLA (Kepner, 1991; Semke, 1984; and Sheppard, 1992; Truscott, 1998, 1999, 2004, 2007, 2010). To properly understand the various aspects of implicit and explicit knowledge and corrective feedback in SLA, there is a need to: (i) identify the difference between the two types of knowledge, (ii) examine the evidence revealed by researchers regarding the different roles played by both types of knowledge, (iii) theoretical evidence in support of implicit and explicit knowledge, and (iv) the different manners in which both these knowledge types promote and enhance language learning and acquisition.

The comparative distinction of these types of knowledge is grounded on the level of awareness of learners regarding the knowledge. For example, explicit knowledge is within the conscious awareness of language learners, so it can be applied to improve written and oral language skills. Conversely, implicit knowledge cannot be consciously applied to improve the overall language proficiency. When learners use implicit knowledge, they are not consciously aware of it. Explicit knowledge is "conscious and declarative", while implicit knowledge can be termed as: "tacit, intuitive and procedural knowledge" in nature (Rassaei et al., 2012, p. 61-62). Ellis (1994) uses the following words in his seminal work to clarify the distinction between the two types of knowledge in the most concise and clear manner:

> "Some things we just come able to do, like walking recognizing happiness in others … We have little insight into the nature of the processing involved… Others of our abilities depend on knowing how to do them, like multiplication, playing chess… We learn these abilities explicitly…" (p. 1)

Implicit knowledge has also been termed as 'tacit knowing' by Polanyi (1967). This is the ability to recognize something without being able to verbally describe it in a contextual manner, and the diagnostic skills within a clinical encounter between patient and psychiatrist. This viewpoint of tacit knowledge in linguistics was further extended by Chomsky (1965), and specifically defined tacit knowledge as internalization of grammar and other linguistic rules or principles. He further claimed that knowledge regarding target language is comprised of knowledge based on the specific features of language that are based on positive evidence (in the form of input) and universal grammatical rules (Chomsky, 1965). Along similar lines, Schacter (1987) defined implicit memory in terms of its lack of intentional recollection, which implies that learners cannot recall anything they have learned within the implicit capacity. This makes learners unable to use implicit knowledge intentionally to improve their overall language learning. Based on Ellis' (1994) terminology, implicit learning refers to the acquisition of knowledge through

natural processes without conscious after-thought, while explicit learning refers to activities in which “the individual makes and tests hypotheses in a search for structure” (p. 1).

As for explicit knowledge. Hulstijn (2003) defined it as “a conscious, deliberative process of concept formation and concept linking” (p. 206). Another important issue linked to these knowledge types is related to accessing specific forms of knowledge to enhance learners’ ability towards language learning and acquisition. According to Godfroid et al. (2015), access to implicit knowledge is possible through automatic processing and procedural representations; while explicit knowledge can be accessed using controlled processing and declarative representations.

Originally, the concept of implicit learning was coined by Reber (1967) refer to acquiring knowledge regarding complex stimuli without the awareness of the acquired knowledge and associated skill development. Subsequently, Reber (1993) used the following words for implicit learning: “the acquisition of knowledge that takes place largely in the absence of explicit knowledge about what was acquired” (p. 5). It is further defined as an unconscious knowledge whose process is considered as “elementary and ubiquitous”, while explicit learning is intentional where learners are actively engaged in activities, thus making them to access conscious knowledge (Rebuschat & Williams, 2012, p. 829). The importance of implicit knowledge has been proven within the overall learning and skill development of individuals, specifically within social interactions (Lewicki, 1986), intuitive decision-making (Plessner et al., 2008), music perception (Rohrmeier et al., 2011) and language production and understanding (Reber, 1993; Williams, 2009). On the other hand, explicit learning entails applications within different learning-based scenarios requiring learners to complete tasks with prior knowledge of the required goals and objectives that leads to development of conscious knowledge (Rebuschat, 2013). With regards to the distinction between implicit/explicit knowledge

and implicit/explicit learning, Schmidt (1994) contended that both are inter-related yet different concepts, where the term '*learning*' points to the actual processes involved in learning, while the term '*knowledge*' refers to the products at the end of the learning processes. Nevertheless, Ellis (1991) and Hulstijn & Ellis (2005) noted the strong position of explicit and implicit knowledge in SLA. However, the disagreement on the interface issue of these two resulted to further research studies (Ellis et al., 2009)

To gain a better understanding of implicit and explicit knowledge, it is essential to shed light on the existing literature pertaining to the SLA and its contribution towards improvement in linguistic knowledge. DeKeyser (2003) was of the viewpoint that there is a dearth of studies specifically examining the explicit/implicit discussion in relation to SLA. Similarly, according to Manchon (2011), most of the existing research focuses on using feedback to enhance accuracy, but not on the aspect of learning that deals with language acquisition. Williams (2012) believed that error feedback has an impact on the learners' explicit knowledge, but it was still doubtful whether feedback had any relation towards facilitating development of language acquisition skills, referring to implicit knowledge. Similarly, Bitchener (2012) and Polio (2012) emphasized on the importance of investigating the relation between SLA through interaction with implicit knowledge, while depending on the feedback-based strategies for language learning.

When it comes to using feedback, Long (1996) contended that oral feedback dealing with provision of positive evidence to learners (as is the case in recasts) can lead to the development of implicit knowledge. Direct corrective feedback in writing shares some commonalities with recast, along with some distinct and dissimilar aspects as well. However, in view of theories proposed by cognitive-interactionists, one of the basic traits of effective feedback is that it is provided online within a specific time frame (Doughty, 2001). Therefore, written feedback may not have much impact towards initiating development of implicit knowledge within language learners, even in the presence of

positive evidence (Shintani and Ellis, 2013). Consequently, future researches should focus on highlighting the ways in which written corrective feedback is effective towards enhancing the implicit and explicit knowledge of learners. The issues pertaining to explicit/implicit knowledge are critical to a better understanding of SLA and the underlying processes (Shintani and Ellis, 2013).

The primary debate in the context of implicit/explicit began with Krashen's (1985) model. This made clear distinctions between 'learning' (formulation of grammar rules with complete awareness) and 'acquisition' (sub-conscious internalization of rules pertaining to target language's grammar). It provides the idea that these types of knowledge can have no relationship at all. To elaborate more, Krashen (1985), along with Truscott (1998, 1999) and Schwartz (1993) posited that there was no interface or link between explicit and implicit knowledge. Krashen's (1985) theory was met with considerable level of opposition, as it was too simplistic and left many ambiguities in its wake (Ellis et al., 2009). Truscott (1998) associated grammar error correction with explicit knowledge, while also being of the viewpoint that grammar correction has no effect on the "genuine knowledge of language" (p. 120), in reference to the implicit knowledge. Schmidt (1994) deconstructed consciousness into distinct components, namely *intentionality* (intentional vs. unintentional learning), *awareness* (implicit vs. explicit learning), *control* (automatic vs. controlled processing) and *attention* (attended vs. unattended learning). As a result, work by Schmidt (1994) enabled better understanding of Krashen's (1985) work.

Reber's (1993) seminal work was also able to redefine the debate concerning implicit and explicit interface. His studies focused on two types of learning (explicit/implicit) while focusing on two separate groups of learners using artificial languages in two separate settings; one group learned letter strings without feedback and the other was instructed to memorize letter strings of artificial language (Reber, 1993).

The findings of the study challenged the previously outlined notion of disassociation between implicit and explicit learning and knowledge (Reber, 1993). In order to further stress this viewpoint, DeKeyser (2003) believed that the existing gap between implicit and explicit knowledge could be bridged using output practice, instructions and written corrective feedback. Similarly, it is also believed, based on Schmidt's (1994) theory that corrective feedback that enhances explicit knowledge can indirectly facilitate the development of implicit knowledge through 'noticing' (when the learners pay attention to the specific features of the target language input) and 'noticing-the-gap' (when the learners examine the variations between the typically produced output and noticed features of the target language) phenomenon.

Furthermore, Ellis et al. (2009) had the following different perspectives on the issues related to implicit and explicit knowledge with regard to SLA: (i) "explicit knowledge can be converted into implicit knowledge", with the initial assumption of strong interface and in case of learners' practice of explicit and declarative rules, (ii) on account of weak interface between the two knowledge forms, explicit knowledge aids in the acquisition of unconscious type of knowledge through highlighting some aspects of input to the learners, and (iii) when assuming no link between the two knowledge forms, implicit and explicit knowledge are separately formed and placed within the brain, along with completely separate and isolated mechanisms involving knowledge development. The measurement of explicit and implicit knowledge has been viewed as a critical aspect of paramount importance within the existing works. This provide a clear demarcation of the differences between the two types of knowledge which needs to be evaluated. The need for developing accurate measurement tools for explicit and implicit knowledge is also very important. This is because the existing research studies have focused on measuring and operationalizing both types of knowledge in different manners, thus resulting in incomparable findings (Ellis et al., 2009).

2.5. The Measurement of Explicit and Implicit Knowledge

The above section provided definitions and concept of implicit and explicit knowledge. Controversy with regards to the conflicting claims has also been cited with reference to various studies conducted by many researchers (e.g. Rebuschat & Williams, 2012; Schmidt, 1994, Truscott, 1998; Shintani & Ellis, 2013; Hulstijn & Ellis, 2005). The inability to determine how learners engage in processing information in SLA made this controversy sparked, and the difficulty to ascertain the process gathered the researchers to study the issue (Ellis, R., 2005). Doughty (2003) also emphasized the need to design instruments that could measure both explicit and implicit knowledge. As Ellis, Loewen and Erlam (2005) cited, the form of instruction affects the type of knowledge acquired by the learners. Hence, this section examines topics concerning the measurement of these two types of knowledge.

The general criteria, which is being widely accepted by many researches in terms of distinguishing implicit and explicit knowledge for SLA, has been enumerated and used by many research studies: *focus of attention* (whether the designated instrument focuses on learners' attention on meaning or on form), *time* (difference between learners' online or offline processing), *utilization of knowledge regarding metalanguage* (the degree of reliance of instruments on learners' pre-existing knowledge of metalanguage) and *degree of awareness* (the measurement based on either the use of pre-existing and previously specified grammar rules or allowing the learners' to 'feel' their way through and come to their own conclusions) (Krashen, 1985; Rassaei et al., 2012; Shintani and Ellis, 2013). Ellis et al. (2009) proposed a few other features which include *learnability* (the ability to learn language based), *systematicity* (the level of consistency of learners in performing specific set of tasks), and *certainty* (the level of self-confidence of learners to produce linguistic forms that conform to target language grammatical rules).

Ellis (2009) believed that there are no 'pure' measures of examining and measuring acquisition of explicit and implicit knowledge. Thus, different studies have explored varying type of measures for examining and measuring the implicit and explicit knowledge acquisition within language learners. For example, Shintani and Ellis (2013) made use of Error Correction Test (ECT) and Narrative Writing Tasks to measure learners' implicit and explicit knowledge. A study by Rebuschat (2013) also outlined the following different types of methods for knowledge measurement within the context of SLA: (i) *representative studies* (in the form of retrospective verbal recalls and reports), (ii) *subjective measures* (such as subjective thresholds in perception-based experiments), and (iii) *direct and indirect tests*; where direct tests allow learners to make use of their knowledge right away (e.g. generation task) and indirect tests examine performance of learners without instructing them to make use of the existing knowledge (e.g. using serial reaction time task). On the other hand, the different measures employed by Han and Ellis (1998) to examine implicit and explicit knowledge of verb complementation structures include Oral Production Test (OPT), Timed Grammaticality Judgement Test (TGJT), Untimed Grammaticality Judgement Test (UGJT), and Interviews.

The five tests outlined by Ellis et al. (2009) for the purpose of measuring implicit and explicit knowledge include Elicited Oral Imitation Test (EOIT), Oral Narrative Test (ONT), Timed Grammaticality Judgement Test (TGJT), Untimed Grammaticality Judgement Test (UGJT), and Metalinguistic Knowledge Test (MKT). From the five tests outlined above, the first three (EOIT, ONT and TGJT) are used for measurement of implicit knowledge and the following two tests (UGJT and MKT) are employed for the explicit knowledge measurement. Meanwhile, Godfroid et al. (2015) made use of eye movements' data and different variations of Grammatical Judgment Tests (timed, untimed, grammatical and ungrammatical) in written and oral forms. This is to examine the type of knowledge being used in learners' judgment by varying the time pressure and

item grammaticality considerations. Within EOIT, the learners are orally provided with a set of belief statements that include those with and without grammatical sentences containing target structures (Ellis et al., 2009). TGJT requires the learners to check the grammaticality of the sentences. Within OPT and TGJT, learners need to process sentences within a specific time limit. This puts additional strain that cause the utilization of implicit knowledge (Ellis et al., 2009). UGJT is similar to TGJT, but without the additional time constraints and due to this factor, this test relies on the learners' explicit knowledge. Time constraints and grammaticality, both have been established as factors that significantly impact Grammatical Judgement Tests accuracy (Bowles, 2010; Ellis, 2005; Gutierrez, 2013; Zhang, 2015). However, studies have also revealed that time pressure is a greater influencing factor as compared to item grammaticality within tests designed for SLA learners' performance (Ellis, 2005; Ellis & Loewen, 2007; Zhang, 2015). ONT requires the learners to read a story twice and narrate that story within the time limit of three minutes, which are later recorded and transcribed to examine the use of different target structures by learners for measuring implicit knowledge (Ellis et al., 2009). Moreover, MKT is another test for judging the explicit knowledge which requires learners to either solve multiple choice questions containing ungrammatical sentences that need to be explained in terms of types of errors present in each sentence or by allowing learners to identify different grammatical features within provided text (Alderson et al., 1997).

Existing studies have revealed that the specific research area pertaining to analysis of variations between implicit and explicit knowledge requires additional scrutiny (Rebuschat 2013; Godfroid et al., 2015), as there is still much that needs to be understood, specifically regarding implicit knowledge and the interface between implicit and explicit knowledge. Several studies also identified the relationship of corrective feedback to the development of these two types of knowledge. Studies operationalized implicit and

explicit instruction and feedback to measure its effects to different tests biased to force the use of implicit and explicit knowledge among learners. For example, Akakura (2012) used explicit instruction to determine its efficacy in acquiring implicit and explicit knowledge on English articles. Through CALL (Computer Assisted Language Learning) activities, explicit instruction was provided. Participants were assessed in four several tests in different testing time. Although results have been limited for explicit knowledge, the effect on implicit knowledge was analyzed to be delayed as attributed to the time required to process information internally. Such claim contrasted with previous studies claiming deterioration of effects over time (Akakura, 2012; Norris & Ortega, 2000). It is also important to note that the study made use of CALL which could not be available in other classroom settings, as such, it was recommended to make use of classroom-based activities to enhance the validity of the researcher's claims.

In a similar study, explicit instruction was used to determine its effects on the implicit and explicit knowledge of students in relative clauses. Nezakat-Alhossaini, Youhanaee, & Moinzadeh (2014) operationalized explicit instruction in four sessions and students were tested via offline and online metalinguistic knowledge test. The latter aims to measure explicit knowledge, while the former, with speeded time, aims to measure implicit knowledge. Despite differences in proficiency level, results of the experiment showed efficacy of explicit learning in the acquisition of implicit and explicit knowledge on target grammatical features. Results also showed a significant effect during the delayed post-test which implies longer retention of processed information. Moreover, it was recommended to study other English structures to determine whether effects of explicit instruction can provide similar results.

Recasts and metalinguistic corrective feedback have also been used in the study to test their efficacy in acquiring L2 knowledge, specifically implicit and explicit knowledge (Rassaei, Moinzadeh, and Youhannaee, 2012). Using quasi-experimental

designs, learners were tested in three tests in three different times: pre-test, post-test, delayed post-test. Result of the study identified metalinguistic corrective feedback to be effective in the acquisition of both types of knowledge. Although recast also showed significance, the effect was not evident in the delayed post tense implying its short-term efficacy in retaining processed information. The implicit nature of recast was identified as one of the causes why it failed to facilitate learning gains in the delayed post-test. On the contrary, metalinguistic feedback is salient and could be easily identified by the learners, thus it helps learners to identify correct target structure easily. Moreover, the study also calls for reinvestigation of the measures used to investigate implicit and explicit knowledge. While the study used different testing instruments, their validity is required in different contexts.

Nevertheless, corrective feedback has been a subject in research studies and was found to have a significant relationship in shaping implicit and explicit knowledge. As also mentioned in previous discussion, there is no pure measure when it comes to these types of knowledge; however, such can be attained by designing instruction and materials that can bias either type of knowledge. Previous studies have also recommended further investigation of the variables surrounding implicit and explicit knowledge. This includes investigating in other settings, using other forms of instruction and targeting other grammatical structures.

2.6. Disagreements in Corrective Feedback in Second Language Acquisition

The disagreements in corrective feedback typically revolve around the fundamental questions posed by Schmidt (2001) and Swain (1985). As stated previously, Swain's hypothesis claimed that language learners' development of linguistic knowledge is dependent on their ability to notice a gap. When the opportunity arises within written or oral form of SLA, this allows the opportunity to correct their output respectively (Swain & Lapkin, 1995; Swain et al., 1995). Schmidt's (2001) theory stated that learners

are unable to learn the grammatical characteristics of the target language until they take notice of these aspects first, but noticing does not necessarily translate to language acquisition (Venkatagiri & Levis, 2009). Because of these statements, many studies have been conducted to test its validity, and operationalized various feedback strategies to tap the conscious and unconscious process of learning (Ting & Lin 2015, Kassim & Ng, 2014; Gass and Varonis, 1994).

Nevertheless, the various conclusions surrounding corrective feedback (CF) have been circulating in the academic world, specifically in attributing language learning and acquisition to corrective feedback. While many research studies noted the advantages of employing corrective feedback in the SLA classroom, specifically in lessening errors and enhancing accuracy (Al Ajmi & Ahmed Ali Saleh, 2014; Ashwell, 2000; Ferris & Helt, 2000; Ferris, 1999), there are also studies that investigated its short term effects (Van Beuningen, De Jong & Kuiken, 2008) and the negative impact it brings inside the classroom (Truscott, 1996). Truscott (1996; 2004; 2007), the main opponent of CF, perceived that grammar correction practice tend to be detrimental in learning. It decreases motivation of learners to learn and acquire target language structures. Error correction was also claimed to be ineffective, thus needs to be abandoned. Truscott (1996) holds pessimistic views on the capabilities of the teachers to provide consistent and efficient feedback, if otherwise, learners uptake to the kind of feedback received. Krashen (1981) also argues that feedback promotes anxiety among learning which could create negative impact in language learning. Thus, in this regard, the opposing views on corrective feedback need to be considered critically and highlight relevant studies that provide evidences of the difference among various researchers (Diab, 2015; Eslami, 2014; Ferris & Roberts, 2001). Furthermore, although so much research has also been conducted in the field, which confirms the positive effects of corrective feedback, however, many other studies claim that the research designs employed by those studies were not rich (Zohrani

& Ehsani, 2014), and shortcomings in methodology might have caused the differences (Van Beuningen, De Jong & Kuiken, 2008).

Considering the views that CF is ineffective, Kepner (1991) reported the ineffectiveness of CF when his study did not show any significant differences between the one who received CF and the one who received comments. However, the result of his study was criticized by Chandler (2003) to be lacking in validity since the CF received did not allow students to use it in their writing. Chandler (2003) claimed that CF works only if this is understood and used by the learners, thus Kepner's (1991) study was concluded to provide no warrants in claiming that error correction is ineffective.

Due to the methodological shortcoming revealed in other studies, Chandler (2003) attempted to investigate different corrective feedback involving experimental and control groups. Students' attitude towards feedback have also been considered. As a result, he found feedback to be effective, specifically direct corrective feedback helps learners produce accurate revisions of target language structures. On the other hand, learners in the study perceive self-correction as more effective in language learning and acquisition. Contrary to Chandler's claims, Truscott (2004) questioned her judgment since Chandler's control group in the study did not provide written production towards the end of the study, thus effectiveness of error correction in this study could also not be affirmed.

Furthermore, Truscott and Hsu (2008) investigated the widely-accepted notion that revision is a valuable tool in producing refined writing skills and works with respect to both following dimensions. In determining the quality of both content and form, instructors and educators widely believe that a metalinguistic understanding of English grammar structure is essential and of which all learners of second-languages, such as English, must obtain knowledge. Truscott and Hsu (2008) demonstrated that corrections are helpful to students in regard to reducing instances of errors in grammar: "the revised manuscripts of students who received it showed significantly more improvement in

accuracy than those of students who did not receive it" (p. 293). Findings further revealed that correction did improve students' writing skills and lessened the degree of errors in writing activities. However, the improvement in the write up was only visible in the revisions, thus, no effect was found when learners were tasked to write a new narrative. The authors pointed out the inefficacy of feedback, specifically feedback was only able to improve the writing revisions, but was not able to transfer knowledge to new pieces of writing tasks, and therefore "successful error reduction during revision is not a predictor, even a very weak predictor, of learning" (Truscott & Hsu, 2008, p. 299).

Hyland (2003) also conducted a study by observing writers in an English course. By providing feedback on form, revisions of the writers were lessened suggesting the efficacy of the feedback received. This also means that errors can be treated through feedback. Other studies have also claimed the efficacy of correction in producing more accurate texts (Ashwell, 2000; Ferris & Roberts, 2001), and have improved students' accuracy in using the target grammatical features over the course of time (Ferris and Helt, 2000; Lalande, 1982; Robb et al.,1986). These results oppose to the view of Truscott (1996) that corrective feedback is harmful.

Ferris (1999), opposing to the review of Truscott (1996), examined Truscott's arguments and claims in hope to verify the certainty that feedback in classroom provides no evidence in improving students' language and grammar learning. Scrutinizing the claims, Ferris (1999) concluded Trusctott's ideas to be "premature and overly strong" (p. 2). Ferris (1999) also noted the differences of the participants, in terms of background, who participated in his study. This suggests the generalizability issue of the results to other studies in different settings. The inadequacy in evidence pulled down the validity of Truscott's claims; therefore, she urges everyone to provide enough evidence before providing pedagogical decisions. Ferris (1999) is positive to the usefulness of CF and its

continuous usage in the classroom, so as to promote awareness among leaners of their weaknesses.

As a summary, many scholars in SLA perceive the beneficial effect of CF in improving the accuracy of students (Muncie, 2000; Myers, 1997; Zamel, 1983). The primary supporting issue, in this regard, is that CF can and should lessen grammatical errors of the learners, improve fluency, and promote the development of this topic in SLA (Bitchener & Knoch, 2008; Chandler, 2003; Ferris, 2002; Myers, 1997). Qualitative study was also raised to understand and further analyse the role of feedback in SLA (Diab, 2015). Studying one linguistic category rather than comparing to other linguistic categories would be beneficial to monitor the efficacy of error correction and feedback (Al-Jarrah, 2016).

2.7. Direct Corrective Feedback in Second Language Acquisition

Past studies have been conducted to test the efficacy of different types of corrective feedback in written and oral production in SLA. Much of the articles surrounding corrective feedback have either agree or disagree to the usefulness of such in classroom settings. Ellis (2009) proposed different strategies in providing corrective feedback in the written production of the learners. Published in the ELT (English Language Teaching) Journal, the proposed typology consisted of six strategies, and one of those is the traditional direct corrective feedback. Thus, in this section, direct corrective feedback in second-language acquisition will be discussed in terms of its researched benefits and gaps in linking the learners' second language acquisition to current practices and methods in instructional lessons. The aim is to reveal that each study brings forth an aspect of direct corrective feedback that proved the positive and negative impact as a resource for students learning a second language, namely, the English language and English grammar.

Direct corrective feedback is said to be an adequate tool for improving second-language acquisition largely as it provides the simplest form of error correction in the output production of the leaners (Spivey, 2014). It is operationalized by providing the correct form to the students in response to their perceived error production (Ellis, 2009). Daneshvar and Rahimi (2014) describe direct corrective feedback as "the provision of the correct linguistic form or structure by the teacher to the student above the linguistic error" (p. 218). On the other hand, indirect corrective feedback "requires learners to engage in guided learning and problem solving and, therefore, promotes the type of reflection that is more likely to foster long-term acquisition" (Daneshvar & Rahimi, 2014, et. al.). According to Bitchener and Knoch (2010), direct corrective feedback resolves complex grammatical structures that students might have difficulty when learning a particular grammatical feature in the target language. Furthermore, this allows learners to easily recognize incorrect language forms, rather than memorizing error codes (Bitchener & Knoch, 2010).

Eslami (2014) and Spivey (2014) contended the suitability of this feedback specifically to students with lower proficiency level because they have a limited knowledge when it comes to understanding why a particular word is incorrect. This also aids learners to immediately treat the errors and understand the difference between errors and the target correct forms (Spivey, 2014). Reports from research studies also claimed the efficacy of this feedback in promoting long term accuracy among students when compared to indirect corrective feedback (Bitchener, 2012; van Beuningen, De Jong & Kuiken, 2008), however, such claims need further investigation (Bitchener, 2012). Nevertheless, direct corrective feedback can also be applicable to students with higher proficiency level, however, this only functions as fine-tuning tool to help learners treat minor errors that have been overlooked (Spivey, 2014).

Direct corrective feedback was also claimed to reduce errors during the writing revision process among learners in SLA. Ferris (2002) discussed the findings of her study regarding direct and indirect corrective feedback. She reported that the latter improved the revisions of the learners up to 88%, while the former only improved up to 77%. However, it should be noted that over the course of their study, indirect corrective feedback substantially lessened the error frequency of the students as compared to those who were given direct corrective feedback. This could be seen as an issue of longer learning retention in relation to the type of feedback provided to the students.

Comparative effects of direct and indirect correct feedback have also been studied, determining each effectiveness over treating grammatical errors in writing. While these types of feedback are perceived to be effective, some researchers considered the long-term effects that one provides to the learners (Ellis, 2009; Hosseiny, 2014). In a study conducted by Hosseiny (2014), he aimed to improve the writing skills of the Iranian learners through interventions. Control and experimental groups were studied, took tests, and received feedback (direct and indirect). The findings revealed a significant difference between the experimental groups and control group, but not between the two experimental groups. In this manner, the two types of feedback are believed to be effective and provide significant improvement to the performance of the learners, to which, in the contrary, is different from the findings of Fazio (2001) and Truscott & Hsu (2008).

While results favoured both types of corrective feedback, Hosseiny (2014) regarded indirect corrective feedback as an encouragement for learners to actively take part in the repairing and information processing, which leads learners to realize errors, understand, and use them accurately. Direct corrective feedback, on the other hand, does not provide an opportunity for learners to draw out thinking processes on their own, instead rely solely to teachers' provision of correct form, and thus fails to encourage students to perform a pushed output as mentioned by Swain (1985). Ellis (2009) raised

this concern as students' minimal processing of input which may affect securing long-term learning.

In a study conducted by Lalande (1982), he studied two groups by providing direct and indirect corrective feedback respectively. In the course of one semester, students who received indirect corrective feedback provided more accurate writing productions at the end of the semester. Similarly, Lee (1997) reported the findings of his study claiming that students who received indirect corrective feedback performed better in editing than those who did not. The conclusions provided in these reports disagreed to the prior claims that direct CF provides similar effect with indirect corrective feedback, if not better than the latter (e.g. Hosseiny, 2014). Another example takes the study of Bitchener and Knoch (2008). They researched the efficacy of direct and metalinguistic explanation and found its efficacy over the control group. Results between experimental group, however, provide no difference.

In like manner, the result of the study by Ghandi and Maghsoudi (2014) showed the better results of students in learning correct spelling through the aid of indirect corrective feedback. It was further concluded that a mere feedback without learners' engagement results to a failure in improving language accuracy of the students in any EFL/ESL classrooms. One should also consider that participants in this research were only tested in two different times: pre-test and post-test. Thus, this study could not validate past studies in terms of long-term efficacy as evident in the study of Ferris (2002). Nevertheless, Sivaji (2012) also provides a support to the claims of Ghandi and Maghsoudi (2014). Sivaji (2012) encourages learners to be an active part in treating and correcting errors, which is one of the proclaimed evidence in the use of indirect corrective feedback. This also supports the Learner Autonomy theory of Holec (1980) which proposed independent self-engagement in the learning process.

While most of the discussions highlighted two corrective feedback strategies, experts in SLA gradually shifted their attention in separating corrective feedback into two features- focused CF and unfocused CF. Ellis, Sheen, Murakami and Takashima (2008) provided the distinction between these two features. According to them, unfocused refers to the normal activity of teachers where they correct all the errors of the students in their written work. This extensive feedback treats all language errors available in student's writing. On the other hand, focused corrective feedback goes on the term selective correction, that is treating specific errors while ignoring the others. A highly-focused error is said to be treating only one specific error while a less focused error targets a few but restricted grammatical features (Ellis et al., 2008).

In investigating the efficacy of both, Ellis et al. (2008) noted the positive impact of both strategies in improving accuracy in the use of definite and indefinite articles in written works. Although their methodology lacks enough distinction on students' use of articles, the study provided a strong reason for academic researchers to further the investigation and study other grammatical structures (Fazilatfar, Fallah, Hamavandi & Rostamian, 2014; Sheen, Wright & Moldawa, 2009). Furthermore, Ellis et al. (2008) also reported the possibility of unfocused corrective feedback to be not effective, while focused to be effective. Providing a mass of corrections may fail the students to notice their errors and will not be pushed to provide a better output (Ellis et al., 2008). Sheen (2007) stated that an unfocused CF approach promoted cognitive overload that might affect the attentional capacity of the students.

As an offshoot of the study conducted by Ellis et al., (2008), other researchers began to recognize the importance of the distinction in giving corrective feedback. In a study by Sheen, Wright & Moldawa (2009), these two CF distinctions have also been investigated comparing their effectiveness in improving errors on target grammatical forms. With three experimental groups (focused, unfocused, writing practice group) and

one control group, the study showed students who had focused CF overtook the other three groups. The second to perform better was the writing practice group, third was the unfocused CF, and the last was the control group. Moreover, although all groups showed learning gains, it was only the focused CF group outperformed the control group, thus giving us the idea that focused CF is more effective than unfocused CF.

In hoping to arrive with similar results with past studies, Frear (2012) compared the effects of focused CF to unfocused CF. The result showed a significant difference between the experimental groups and the control group in their writing accuracy tasks. On the other hand, when two experimental groups were compared, both did not provide statistical difference. The result of this study is of similar direction with Rouhi & Samiei (2010). In their study, the experimental groups and control group were found to have no statistical difference. Learning gains were evident during the first and second testing, but not evident during the last (delayed) testing. These studies, Frear (2002) and Rouhi & Samiei (2010), support the claims of Ellis et al., (2008) that both strategies provide improvements in improving accuracy of learners in writing, however, their long-term effects still need to be investigated.

Fazilatfar et al. (2014) also operationalized unfocused corrective feedback in comparison to no corrective feedback group. Although results led to the learning gains of students in the experimental group, it cannot be compared to the efficacy of focused CF as it was not part of the comparison in the study. To validate findings of this with direct CF, it is ideal to conduct other investigations that include the two CF over treating errors and improving writing accuracy.

A more recent study was conducted by Frear and Chiu (2015) comparing focused and unfocused indirect corrective feedback. With a quasi-experimental design, participants were tested in three testing times receiving focused CF, unfocused CF, and no corrective feedback treatment procedures. Both experimental group outperformed the

control group in the immediate posttest and delayed posttest. This suggests that both feedback are effective over the course of time. Nevertheless, it was also mentioned that both feedback never lead to metalinguistic understanding, however, push learners to provide more accurate output.

As a summary, this section tackled the differences between direct and indirect corrective feedback, and was further brought to their distinction as focused and unfocused corrective feedback. Relevant studies have highlighted each strengths and weaknesses; however, methodological problems of these studies also need to be considered. Nevertheless, it follows that students of second-language learning must be exposed to classroom opportunities in which the instructor explicitly tries to refrain from exerting complete control of the classroom. The provision of corrective feedback should be in line to the interests of the students to promote metalinguistic understanding and long-term efficacy.

2.8. Metalinguistic Corrective Feedback in Second Language Acquisition (SLA)

This section provides a discussion of the existing research on the efficacy of metalinguistic corrective feedback in SLA. The majority of existing research into metalinguistic feedback examines its use in the formative assessment of oral second language skills (i.e., Motlagh, 2015; Rezaei & Derakhshan, 2011) but very few studies examine the metalinguistic corrective feedback of students' in terms of written work. Research studies reveal that metalinguistic understandings encourage students to reflect on their corrections (Ellis 2013). Its process allows teachers to write 'explicit comments' on the errors that learners made in their writing (Ellis, 2008). Metalinguistic feedback includes any information, feedback or comments by the teachers directed towards the language learners that highlight the linguistic accuracy of learners' utterances without directly providing the corrected linguistic form (Lyster and Ranta, 1997). According to Ellis (2008), explicit comment can be found in two different forms, namely error codes

and direct correction that supply the accurate form. Error codes provide some implicit clues regarding the location and type of error. The indication of an error allows the learners to reflect on the correct solution and evaluate the numerous possibilities of the correct form. This engages learners in a process of metalinguistic thinking about abstract concepts in grammatical systems, particularly in the English language. On the other hand, Chandler (2003) claims that metalinguistic corrective feedback is operationalized by underlining the errors and providing the target form above the word. Teachers point out errors and supply cues or structures regarding the correct forms. However, this feedback may be generic or specific (Rezaei & Derakhshan, 2011). As Ellis (2008) also stated, metalinguistic corrective feedback can also provide learners with metalinguistic explanations of the specific errors made, but this method is less frequently used as it is time-consuming activity when compared with the use of error codes. It also requires instructors to have an adequate metalinguistic knowledge. In operationalizing the feedback, the error code requires the instructor to write the codes in the margin of the paper (Ellis, 2007). It can be anything from ww (wrong word) or art (article), while the teachers number the errors followed by their grammatical description at the end of the text (Ellis, 2008). Nevertheless, focused metalinguistic CF promotes understanding of the errors while unfocused feedback might not be as helpful or beneficial as the former in addressing specific language structures (Ellis, 2009).

In the study by Gholaminia et al. (2014), a side-by-side comparison and analysis was conducted using direct corrective feedback and metalinguistic corrective feedback. This would allow readers to understand their impact on learners' language and error improvement and proficiency within the target language. The results of the study were examined using t-test analysis method and metalinguistic code-correction showed significantly better results when compared to direct corrective feedback in the understandings of the target language grammatical systems (Gholaminia et al., 2014).

This concludes the efficacy of MCF better than DCF when integrated within the learning processes. Furthermore, using MCF in a classroom practice encourages the students to be more involved and dedicated to learning and acquiring the target language, as it places more responsibility on the students to correct their own errors (Gholaminia et al. 2013). Conversely, Rassaei and Moinzadeh (2011) found no significant difference between two groups that were provided metalinguistic corrective feedback and recasts. Further analysis revealed that while both are effective in post and delayed posttest, recast was found to be more effective in the long-term (Rassaei and Moinzadeh, 2011). Nevertheless, metalinguistic feedback plays a critical role in enhancing the second-language learning and acquisition by allowing learners to focus on their linguistic errors (Rassei & Moinzadeh, 2011).

Similarly, in a study investigating corrective feedback by Kazemipour (2014), it was found that when students corrected their own errors, following learning process treatments that encouraged them to do so (teacher underlined errors with no provisions of the correct form), the students were more engaged in the course content. Thus, in this case, metalinguistic corrective feedback took an indirect form and proved to be more effective at raising consciousness among the learners as well as aiding in long-term SLA. Nevertheless, learners were provided with incentives to participate and self-correction was encouraged, rather than receiving correct forms from the instructor. Azizi, Behjat & Sorahi (2014) also utilized various corrective feedback to improve the writing performance of the learners; however, only metalinguistic corrective feedback provided positive impact to the written output of the learners. In another study, Diab (2015) investigated the combination of metalinguistic and direct corrective feedback over metalinguistic feedback alone to report that MCF involved deep internal processing of target structures, which are essential to understanding of the target grammatical concepts. In conclusion, it was reported that the combination of both traditional and MCF were both

effective in immediate and delayed posttest as evident in the fewer lexical errors of the students in the experimental group (Diab, 2015). Concerning the engagement of students in corrective feedback, Han and Hyland (2015) explored how learners engaged, interacted and perceived different forms of corrective feedback techniques. Their findings emphasized the need of the teachers to examine the background of their students in order to facilitate learner's engagement in corrective feedback (Han & Hyland, 2015). However, this study only utilized data from students studying in the same class, while changes in engagement were not studied overtime. Thus, the findings of this study cannot be generalized to students with different background and proficiency level, and other studies in future that are featured in different contexts will help in validating these results (Han & Hyland, 2015).

Many studies delved on the efficacy of corrective feedback using test measurements, while Faqeih (2015) studied the attitudes of the learners towards different feedback types. He operationalized metalinguistic feedback as an intervention and was found to be most preferred by the learners participating in his study. This can be attributed to how the feedback was given explicitly and the level of familiarity of the students with different feedback methods. Moreover, the researcher also claimed that preference must have shifted since traditional method were being used within those institutions for a very long time (Faqeih, 2015). A study in oral proficiency was also conducted using metalinguistic feedback as a treatment. Fahim and Montazeri (2013) incorporated metalinguistic feedback technique in order to examine language learners' improvement within different linguistic aspects such as grammatical range, lexical resource, and overall proficiency and accuracy in oral form. Based on the results, metalinguistic feedback can effectively enhance learner's proficiency level, specifically in the acquisition and accuracy of target grammatical features (Fahim and Montazeri, 2013). It was also emphasized in the study that teachers should encourage error correction as an important

process in language learning and acquisition. They should also use different techniques to ensure good relationship with students is established (Fahim & Montaseri, 2013).

To conclude, this section provided a review of the different evidences proving metalinguistic feedback to be an effective intervention in SLA. While several studies have confirmed the efficacy of direct corrective feedback as provided in the previous section, there is also an adequate amount of evidence that favours the use of indirect feedback such as metalinguistic feedback in classroom practices (Ferris & Helt, 2000; Lalande, 1982; Shintani & Ellis, 2013). During cross-analysis between groups receiving metalinguistic feedback and recast, Ellis et al. (2006) discovered the former group surpassed the latter. Similarly, Sheen (2007) research revealed that when metalinguistic feedback is offered as a complementary method to providing direct written corrective feedback, the feedback becomes more effective in developing ESL learner's writing accuracy as opposed to without the use of metalinguistic discussions. Succeeding studies such as Ellis (2009) as well as Rassaei and Moinzadeh (2011) further proved that metalinguistic feedback outperforms other feedback types such as recasts and clarification requests.

2.9. Theoretical Framework of The Study

The Comprehensible Output and Noticing Hypotheses have been given considerable importance in the research area of Second Language Acquisition (SLA) (Swain, 1985; Schmidt, 2001). The topic of SLA has been surrounded by numerous hypotheses and theories about how individuals learn second languages (Long, 1996; Krashen, 1985). Output and Noticing Hypotheses rely on the different cognitive processes underlying language acquisition, which include noticing aspect of learners (Schmidt, 1990), noticing with metalinguistic understanding (Schmidt, 2001) and pushed output (Swain, 1995). Earlier studies exploring input's contribution to language learning have concluded that it facilitates acquisition of words within the second language, but it has no

effect on the learners' ability to learn syntactic structures associated with the target language (Swain, 1985; Tanaka, 1991; Ellis, 1994; Long, 1996). It is also noted that by solely exposing students to input does not result in sufficient achievements within language learners (Long, 1996; Long & Robinson, 1998; Norris & Ortega, 2000). Similarly, recent studies related to cognitive psychology and second language acquisition have questioned the efficacy of input-based language learning, as not all input efforts translate to learning intake; while at the same time, supporting role of attention and 'noticing' with respect to facilitation of input and assistance in learning and acquisition of second languages (Izumi, 2002).

In view of the Comprehensible Output (CO) Hypothesis, Swain and Lapkin (1985) contended that learning occurs when language learners encounter a gap in their linguistic knowledge of the target language. Consequently, the learners consciously 'notice' the gap and due to this awareness of their linguistic shortcomings, they are given opportunity to modify their output. They can also highlight and learn previously unknown aspects of the target language (Swain, et. al., 1995). In this hypothesis, the learners are able to conduct mental grammatical processing that allows them to quickly learn and acquire knowledge pertaining to the target language (Swain, 1985). Swain (1985) further proposed that the hypothesis has the following functions for the purpose of SLA: (i) it enables learners to be more aware of their deficiencies in their existing language proficiency and knowledge. This enables highlighting their linguistic errors, (ii) role of output as hypothesis testing function. This can be fulfilled in the form of feedback, where learners negotiate to enable transfer of meaning from the speaker to the listener, (iii) output extends the learners' knowledge of the grammatical structures as well as other forms and rules of the language being learnt, thereby fulfilling deeper understanding, (iv) it enables the learners to develop their language learning skills by translating from *comprehension* (semantic use of language) to *syntactic* (allowing learners to pay attention to language forms) use of

language, and (v) promotes fluency and confidence in the use of secondary language. Swain (1995) also believed that learners generally receive concise and clear input, chances to negotiate for significance and opportunities to produce modified output within the context of classroom-based learning and teacher-learner interactions. Thus, it follows that SLA research has mostly involved studies on classroom interaction.

Noticing Hypothesis, on the other hand, claimed that learners for SLA are unable to acquire knowledge regarding the grammatical characteristics of a language unless they are able to consciously 'notice' them first (Venkatagiri & Levis, 2009). In accordance with 'notice the gap' principle, learners will be able to notice the language form of the target language and the linguistic aspects using their own target language skills. With this, it enables learners to 'notice' gaps in their own learning within their interlanguage (Schmidt and Frota, 1986). Schmidt also claimed that noticing alone does not equate to an automated acquisition of language. Instead, Schmidt hypothesized that learners must notice for it is the fundamental beginning point for language acquisition (Venkatagiri, et. al., 2009). Schmidt (2010) emphasized the importance of attention and noticing if learners want to improve their acquisition of second and/or foreign language. Studies have also shown that learners notice 'holes' in their existing linguistic knowledge, which they attempt to fill with the help of either consulting dictionaries, other assistive material or by asking peers and teachers or by making a mental note to pay attention in future relevant input (Kowal & Swain, 1997; Swain, 2000). Moreover, Schmidt (1990) argued that when an item or form constantly occurs, and its saliency is recognized, then it increases the chance of that particular structure to be noticed. Frear and Chiu (2010) also agreed that once feedback is given, it becomes an output which can also increase the likelihood of critical understanding.

It is important to understand that both the selected theories complement each other, as the underlying cognitive processes linked to these theories are similar in nature.

Furthermore, noticing the language form is one of the most important role and function of output (Swain, 2000), which shows that both these theories are intricately linked and inseparable in nature. Therefore, it is essential to consider these theories collectively with regards to their contribution towards second language acquisition.

For corrective feedback, the evidence outlined in this previous section provides a general understanding of the importance of Swain (1985) and Schmidt's (2001) theories within the context of SLA. Although, the highlighted studies in this section only focused on the concept of these hypotheses, it is important to note that these have served as a framework for most of studies in SLA, particularly in corrective feedback (e.g. Frear & Chiu, 2010; Bitchener, Young, & Cameron, 2005; Rahimi & Zhang, 2014). Nevertheless, errors must be corrected through feedback to highlight the input and make necessary changes to help learners understand target language (Lightbown & Spada, 2013)

2.10. Chapter Summary

This literature review sheds light on studies that tested the efficacy of different corrective feedback techniques in the context of SLA. The primary purpose of this literature review is to provide the researchers with understanding of the various existing findings, claims and conclusions of other researchers. Findings from past studies are critical within the scope of this research, as they guide and enable the researcher to direct the focus on the specific aspects that require attention in the forthcoming sections of the research study. Furthermore, an understanding of the gaps in these studies is a valuable tool to highlight research areas of metalinguistic corrective feedback that need to be examined in a thorough manner or require further investigation in future. Some examples of the gaps discovered during this research include: (i) the uncertainty surrounding the impact of external conditions that may or may not favor metalinguistic corrective feedback practices in SLA, (ii) the influence of second-language learners' native tongue on SLA, (iii) the depth of the instructors' understanding of metalinguistic concepts of the

target language (iv) whether or not metalinguistic corrective feedback is more effective in combination with other methods of corrective feedback. There has been a considerable lack of consensus within the research area in terms of providing a conclusive and comprehensive set of evidences that can shed light on the actual efficacy of direct and indirect methods of feedback. Therefore, the current study focused on the existing evidence in a critical manner, while carefully avoiding the pitfalls of bias and subjectivity to ensure that the findings from primary and secondary research can correlate and validate one another.

CHAPTER 3

RESEARCH METHODOLOGY

3.1. Introduction

The previous chapters revealed studies that reviewed and investigated the efficacy of corrective feedback. However, the differential results of those studies did not validate how and why corrective feedback must be used in the classroom. Methodological issues have also been found: the lack of control group (CG), the test for long term effects, the use of other corrective feedback types, and using corrective feedback to enhance grammar accuracy in writing. Thus, this study aims to improve the previous research designs by considering those variables lacking in past research studies.

Nevertheless, this chapter provides an outline of the research design and the procedures it followed to conduct the quasi-experimental study. Instrumentations used in the study are also described considering the different times they were given. Ethical considerations have also been followed which is as an important part in the data collection process. The qualitative part of the study, following thematic analysis, was conducted after the quantitative study to support the statistical results.

The mixed method explanatory design of the study made a strong support to enhance past results and provide a new knowledge in the teaching and learning process in the area of SLA.

3.2. Research Design

A research design is the total plan of a research study that considers data collection and its analysis. Conditions are arranged in a manner that allows the researcher to execute the plan with no or at least minimal complications. It provides a structure on how to collect, analyze and interpret the data using the most appropriate tools available (Creswell, 2012). With the outline being provided, the scientific research becomes

manageable, and variables are easily determined and manipulated. This warrants the credibility of a study, ensures validity and reliability, and helps the researcher to determine whether the research objectives are met, and whether the research questions are fulfilled and answered accurately (Kothari, 2004). Nonetheless, it serves as a blueprint of research study and research problem that determines the appropriate research design (De Vaus, 2001).

In this research study, starting from the problems and objectives, a mixed method design is believed necessary. Mixed method is a combination of two designs, the quantitative and qualitative design (Ary, Jacobs, Sorensen & Razavieh, 2009). Each design aims to contribute to the understanding of a certain phenomenon. Ary et al. (2009) claims the worthiness of a mixed method design would help in the overall understanding of the study rather than by utilizing a single design. Additionally, Creswell (2012) regarded this design as an advanced method and requires a lot of time in the processing, specifically the data collection and analysis. A mixed method design must also have a logical blend, otherwise it will result to a noticeably separate research studies (Yin, 2006).

Following specificness in a scientific study, Ivankova, Creswell & Stick (2006) pointed the important connection of the statement of the research problem of the study in choosing the appropriate mixed method design. Hence, in this study, sequential explanatory mixed method design has been utilized. This design makes use of the quantitative phase first, then followed by the qualitative phase, usually through follow-up interviews. The aim of this design it to seek a sound and in-depth understanding of the statistical results obtained in the quantitative phase (Ivankova et al., 2006).

The figure 3.1 below shows the design of the study from research question to data collection and analysis. It provides all the action involved when conducting this study.

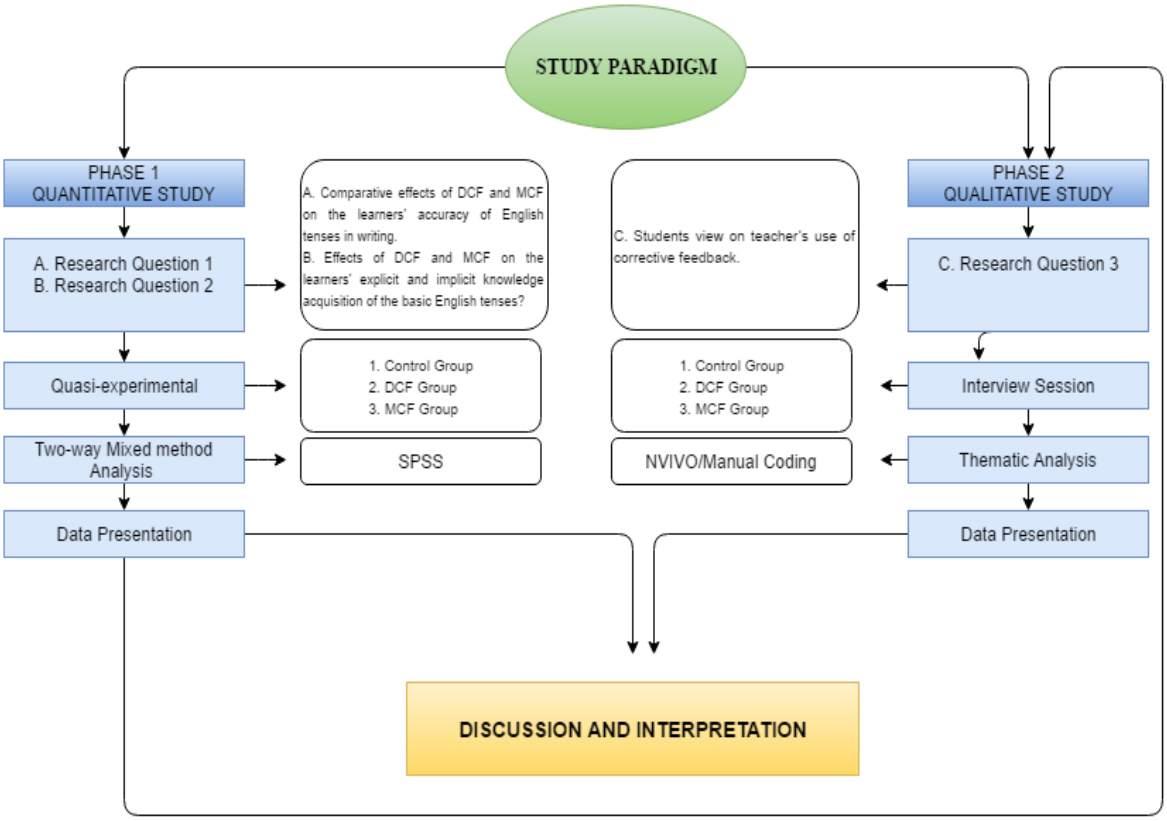


Figure 3.1: Study Design Paradigm

3.3. Ethical Consideration

Saunders, Lewis & Thornhill (2007) associated ethics to moral actions that is practiced when a study is conducted. According to Bryman and Bell (2007), ethics in research has been considerably important that every researcher must be well-informed, otherwise problems may arise, and worst may question the acceptability of the results. Additionally, Creswell (2012) also emphasized the need of ethical considerations before the conduct of the research.

Hence, in this study, the researcher carefully followed the ethical guidelines before and after the conduct of the study.

1. Before the commencement of the study, a letter was sent to the research site asking for permission and the outline of specific dates and activities. Once permission was given, a meeting with selected lecturers was conducted to ask for their

assistance. Informed consent was also collected from the students. Their willingness to participate in the study was asked. Concerned participants were also informed beforehand that they are free to withdraw from their participation in any time they want.

2. When doing the quantitative study, participants were given the chance to ask questions. Instructions on how to answer the test questionnaires were given and any clarification was welcome. Participants were assured that their profile will be kept confidential, and all information is for the sake of research only. No deception or any exaggeration was also given to the concerned participants.
3. When doing the qualitative study, interviewees were shown of the device to be used in recording the interview. The interview process was also explained, and any collected data would only be used for the research study. For confidentiality, participants were assured that their names would be replaced with number case, and that data would be destroyed after the study has been done.
4. Participants have also been informed that any results from their tests and interviews would have no any relationship with their school performance. They were informed that the study was aiming to obtain useful conclusions that would be helpful among language teachers and learners.

3.4. Research Site

The research study was conducted in a language center department of one of the colleges in Selangor. The department was run by an independent education consultancy where access has been permitted. Prior to the access of the site, a letter was sent to the department head asking permission to conduct the study. Research proposal was also given and target schedules for research procedures were provided. The site hosts 85% Malay Felda settlers' children and 15% children from Chinese and Indian race.

Students who enrolled in this programme were required by the research site to undergo a placement test to measure their English proficiency level. The first level starts with the beginner level (preliminary) and ends with advanced English level. As a registered center of City & Guilds UK, the site offers intensive English programme-International Spoken English for Speakers of Other Languages (ISESOL) and International English for Speakers of Other Languages (IESOL). The curriculum is set to the standards of Common European Framework of Reference (CEFR), and learners under this programme are required to acquire specific learning skills to pass and receive international certifications.

3.5. Subject/Participants

The participants in this study were recent graduates from secondary education residing in Malaysia (n = 45). They were grouped according to their placement scores conducted by the research site. Three intact classes were chosen as was also recommended by the center. Because of the research site's placement test, it can be assumed that the all participants have beginner level of English proficiency.

Table 3.1 below presents the groupings and task distribution. The first group served as the control group (CG) while the second and third as the experimental group receiving their respective treatment. All participants were of the same level, confirmed by their placement scores and the site's head of the department. Homogeneity of students could also be assessed based on the pre-tests results of the students. A follow-up interview was also conducted to two class teachers handling the same level to validate the level of proficiency of the students. It was also mentioned through verbal discussion that the participants have difficulty in subject-verb agreement, specifically the basic tenses as required in their CEFR-based curriculum.

The assistance of researchers and teachers from the site enabled the facilitation of the different tests by highlighting the different mistakes in written form and providing relevant feedbacks on the different instruments employed in this research.

Table 3.1: Students' demographics

Background information				
Group		**Group 1**	**Group 2**	**Group 3**
Nationality		Malaysian	Malaysian	Malaysian
Age		18	18	18
Gender	Male	4	4	6
	Female	11	11	9
Years of learning English		7-15 years	11-13 years	2-12 years
Task Distribution				
Facilitator		Lecturer	Researcher/ Lecturer	Researcher/Lecturer
Operationalization		No feedback	Direct Corrective Feedback	Metalinguistic Corrective Feedback with explanation
Number of students/participants		15	15	15

3.6. Selection Criteria

The participants of the study were chosen based on the recommendation of the research site. Three intact classes were provided to make sure that it would not disrupt classes, especially if random sampling is done. This procedure allowed the researcher to easily conduct the study since sampling is readily provided. According to Ary et al. (2009), although error estimation is not possible in this sampling, and that probability sampling is ideal, researchers use the latter method as this is the only choice and available for them. Moreover, Creswell (2012) justifies the use of this by taking the initiative and willingness of the participants to take part in the study. Additionally, age was also

controlled and all possessed beginners' or primary English proficiency based on the college's placement test.

3.7. Validity and Reliability

Blumberg, Cooper and Schindler (2005) defined validity as measuring what we actually want to measure using specific tools or test measurements. On the other hand, reliability refers to how consistent your results are when used in the second time. This simply means generating test results similar to the previous result in similar conditions and procedures (Blumberg et al., 2005). Validity and reliability are considered as vital in conducting a research study. This determines the confidence in your findings and the acceptability of the study in the academic world. Nevertheless, owing to the understanding of the importance of these procedures, the study secured validity and reliability testing before conducting the study. Details of the procedures are discussed below:

3.7.1. Test Validity

Having determined the essentiality of validity in test measurements, face and content validity were conducted prior to pilot testing and actual test administration. Face validity is defined as how the test measures what it seems to measure (McLeod, 2013). It is done by looking through the items without in-depth justifications. When conducting face validity, two lecturers from the college were asked to examine the items in the test. A scale was provided and the tests were rated on its purpose and measurement based on these scale:

1. extremely suitable
2. very suitable
3. adequate
4. inadequate
5. irrelevant/unsuitable

Both lecturers rated the tests as very suitable, number 2 of the scale. This gave the tests a strong face validity. Moreover, to ensure content validity of testing instruments, a rubric was used to assess each items appropriacy to the learning content, curriculum, and learners' proficiency level. Items in the rubric were constructed in reference to Ward and Murray-Ward's (1999) *Assessment in the classroom* book and Groniond's (1982) *Contructing Achievement tests* book.

Two subject matter experts from the site were selected to assess the test items. They were chosen because of their educational background and length of teaching experience. These experts have Master's degree in the related field and have been teaching similar subjects for more than 5 years. By calculating the means of the assessments of TGJT, MKT, and Picture test, result showed a strong agreement by the two raters. This is outlined below:

Table 3.2: Content validity measurement

Test	Rater 1	Rater 2	Average
Metalinguistic Knowledge Test (MKT)	3.75	3.75	3.75
Timed Grammatical Judgment Test (TGJT)	3.78	4	3.89
Picture series test	4	4	4

3.7.2. Test Reliability

A test-retest was conducted in a class who was not involved in the pilot study testing. Pearson correlation was used to determine the reliability of the instruments. For a test to be reliable, *pearson r* must be between .6 to .9 (Creswell, 2012). The closer to 1, the more excellent the reliability is.

Table 3.3: Timed Grammatical Judgment Test

Correlations

		TGJT_Time 1	TGJT_Time 2
TGJT_Time 1	Pearson Correlation	1	**.772****
	Sig. (2-tailed)		.009
	N	10	10
TGJT_Time 2	Pearson Correlation	.772**	1
	Sig. (2-tailed)	.009	
	N	10	10

**. Correlation is significant at the 0.01 level (2-tailed).

There was a positive correlation between TGJT tested in Time 1 and TGJT tested in Time 2 to students in the pilot test, r = .772.

Table 3.4: Metalinguistic Knowledge Test

		MKT_Time 1	MKT_Time 2
MKT_Time 1	Pearson Correlation	1	.737*
	Sig. (2-tailed)		.015
	N	10	10
MKT_Time 2	Pearson Correlation	.737*	1
	Sig. (2-tailed)	.015	
	N	10	10

*. Correlation is significant at the 0.05 level (2-tailed).

There was a positive correlation between MKT tested in Time 1 and MKT tested in Time 2 to students in the pilot test, r = .737.

Table 3.5: Picture Narrative Writing Test

		Picture_Time1	Picture_Time2
Picture_Time1	Pearson Correlation	1	**.716***
	Sig. (2-tailed)		.020
	N	10	10
Picture_Time2	Pearson Correlation	.716*	1
	Sig. (2-tailed)	.020	
	N	10	10

*. Correlation is significant at the 0.05 level (2-tailed).

There was a positive correlation between Picture Test tested in Time 1 and Picture Test tested in Time 2 to students in the pilot test, r = .716.

Judging from the results of test-retest of the testing instruments, we can assume that the tests are reliable. The stability of scores from the first administration to another shows a positive correlation with the generated reliability score of r= .7 in the three tests in two testing time.

3.8. Research procedures

3.8.1. Phase 1: Quantitative Experimental Non-Equivalent Design

This component was conducted in three stages- pre-tests; post-tests; and delayed post-tests.

3.8.1.1. Testing Instrument

Three tests have been given to the participants in reference to Shintani & Ellis's (2013) study. All of the tests conducted focused on the written aspect of corrective feedback strategies. A total of three different instruments were given: Timed Grammaticality Judgement Test (TGJT), Metalinguistic Knowledge Test (MKT) and Picture Narrative Test (PNT). The corrective feedback strategies were employed in the Picture narrative writing test in order to compare and analyse language learning and improvements on the two treatment groups, who were either given direct corrective feedback (DCF) or metalinguistic corrective feedback (MCF). On the other hand, the assessment of learners' implicit and explicit knowledge was gauged using TGJT and

MKT respectively. Each of the three tasks were conducted in a linear sequential fashion, such that written task corresponding to TGJT was tested first, followed by MKT and Picture narrative writing tests for the three groups. In the picture series narrative writing test (PNT), a series of four pictures were given side-by-side, which depicted a single action and its progression throughout the four pictures. Time frame for each instrument was determined by the pilot testing where other group of students timed themselves when answering similar test instruments. The details below provide the measurement tools used in the study.

1. *Timed Grammatical Judgment Test (TGJT)*

Timed Grammatical Judgment Test (TGJT) was administered to measure participants' implicit knowledge on basic English tenses. Similarly, Rassaei, Moinzadeh and Youhannaee (2012) used Timed Grammatical Judgment Test (TGJT) to measure the implicit knowledge of the Persian learners in definite and indefinite articles. According to them, this test can measure implicit knowledge as it allows learners to use their feeling with answers that correspond to their implicit knowledge. This test also forces the students to focus on meaning rather than in form, and that access to metalanguage is impossible, if not limited. Moreover, Godfroid, Loewen, Sehoon, Ji-Hyun, Gass and Ellis (2015) also conducted a study and used Timed Grammatical Judgment Test (TGJT) to measure implicit and explicit knowledge of the performance of native and non-native English speakers. In their study, TGJT targeted language structures which include grammatical and ungrammatical items in timed and untimed manner. Godfroid et al. (2015) stated that time pressure serves as an underlying factor in the test performance of the participants, and that such factor serves as a measurement of implicit and explicit knowledge (Bowles, 2011; Ellis & Loewen, 2007; Zhang, 2014).

With regard to those studies, participants in this study answered a 68-item grammatical judgment test within 8 minutes including the time when they wrote their

personal details. In these items, only 20 was targeting the language structures, and the rest served as distractors and were not included in the analysis. Items in the test were also reshuffled for the different testing periods. The time length was identified through pilot testing where students finished the test within the average time of 5 minutes and 3 minutes for writing their personal details. Some examples of the type of sentences used within TGJT include the following:

1. Martin completes his assignment and print it out the other day.
2. Joseph is running when his mother arrived.
3. Liao works very hard but earns very little every month.
4. Keum is going to buy a computer next week.

Participants were also given instructions before the time was started. During the test, they were not allowed to ask questions for clarifications, and must remain focused to the test. The test reliability was also measured using Cronbach Alpha Test of Internal Consistency through SPSS. The result was r= .8 which represented a good result for internal consistency.

Scores in this test in the different testing periods were tabulated in the SPSS. Each correct answer corresponded to one point against the totals score.

2. *Metalinguistic Knowledge Test (MKT)*

The Metalinguistic Knowledge Test (MKT), with reference to Ellis et al. (2005) was used to measure the explicit knowledge of the participants. MKT affords to measure this type of knowledge as participants are forced to use rules over intuition, and enables participants to access metalanguage and put focus on the form (Ellis, 2004; Ellis, 2005; Shintani & Ellis, 2013). The test is also not time-bounded allowing participants to reflect answers for questions (Ellis, 2005). Nevertheless, MKT draws out the awareness of the participants in the target language structure and allow them to carefully analyse it (Ellis, 2005).

In the operationalization of the test, participants were informed of the instructions. As a standard practice in the target site, participants could use a monolingual English dictionary. This would help them to better understand terminologies that were vague to them. Students were given enough time to answer the test, at least 25 minutes as confirmed during the pilot testing. The test consisted of three parts- A. background information, B. MKT Test, C. Simple grammar test. The metalinguistic knowledge test consisted of 10 incorrect basic tenses in grammar- simple past, past continuous, simple present, present continuous. Within this test, the learners were provided with multiple-choice questions with underlined phrases within highlighted sentences that required learners' intervention (to assess whether the underlined phrase is correct or not, while providing appropriate explanation) using one of the four choices provided to the learners. The other sections and parts of the test were added to make sure the participants would not realize the target structure of the activity. The test was also checked by two professional lecturers who have been teaching at the target site for 2 years to check the contents and test structure. Similar test was also used in the different testing periods; however, items were reshuffled. Additionally, the test used was also used in pilot testing to test for reliability, and was analysed through Cronbach Alpha Test of Internal Consistency in SPSS. The result yielded r=.7 which is considered as acceptable. According to Wells and Wollack (2003), a test should have a high internal consistency when it is closer to 1, and should be not less than 6. Creswell (2012) also considered r= .6 as reasonable to be used in classroom tests and research studies. Nevertheless, the test used in this study has an acceptable internal consistency standard for any testing procedures. An example of the type of questions given in this section has been outlined in the proceeding section:

The school <u>expel</u> him in the class last week.

a. Change 'him' to 'it' because no gender is specified.

b. ‘expel’ should be ‘expelled’ because the noun is plural.

c. Use ‘a’ instead of ‘the’ to show definiteness.

d. The verb should be ‘expelled’ because the action already happened.

In scoring the administered test, each correct answer was rewarded with one point, while the incorrect gained no points. Total scores were entered in SPSS for analysis comparing the effects of the received feedback in different testing periods.

3. Metalinguistic Handout

The metalinguistic handout was used together with the metalinguistic corrective feedback. In the study of Rezazadeh, Tavakoli, & Rasekh (2015), they used this handout as a feedback by providing explicit explanation on the grammatical rules of using English articles. Shintani and Ellis (2013) also claimed the benefits of metalinguistic explanation in the explicit knowledge of the participants. Nevertheless, the current study also made use of Metalinguistic handout; however, it was not being operationalized alone. The handout, unlike other studies, was revised by adding error codes on the right side of its subtitles to relate with the error codes provided in the narrative writing tests. An example of the metalinguistic handout is provided below:

Table 3.6: Sample Metalinguistic Handout

Code: **SP**	**Past Simple** *The action ended in the past and has no real connection with the current time.*
The form is VERB + ED	
Example: ✓ Joshua tried his new shoes last night. ✓ I withdrew my money last Monday.	

4. Picture Narrative Writing test (PNT)

Gutiérrez, Puello & Galvis (2015) adapted the use of picture series to teach and improve students’ writing skill in the sense that it provides contextual elements by bringing realistic concepts in the classroom. Bitchener & Knoch (2010) also claimed that

pictures force the students to write the obligatory target structures because of the actions situations presented in the picture. Nevertheless, the use of picture series as a task has been used by many studies targeting specific grammatical structures (Ellis et al., 2006; Ellis et al, 2008; Rezazadeh, M., Tavakoli, M., & Rasekh, A. E., 2015; Sheen, 2007; 2010; Shintani & Ellis, 2013).

In the current study, there were three different pictures given to the participants for the narrative writing test. The pictures were drawn by the researcher patterned to the test used by Shintani and Ellis (2013) in their study on corrective feedback. The test has also undergone face and content validation through the lecturers at the faculty. Reliability was obtained through pilot testing obtaining r= .7 in Cronbach Alpha Test of Internal Consistency. Time was also determined through the pilot test which had the total average time of 45 minutes. The figure given below is one such example, along with depicted actions taking place within those sequences of pictures used within the picture narration test.

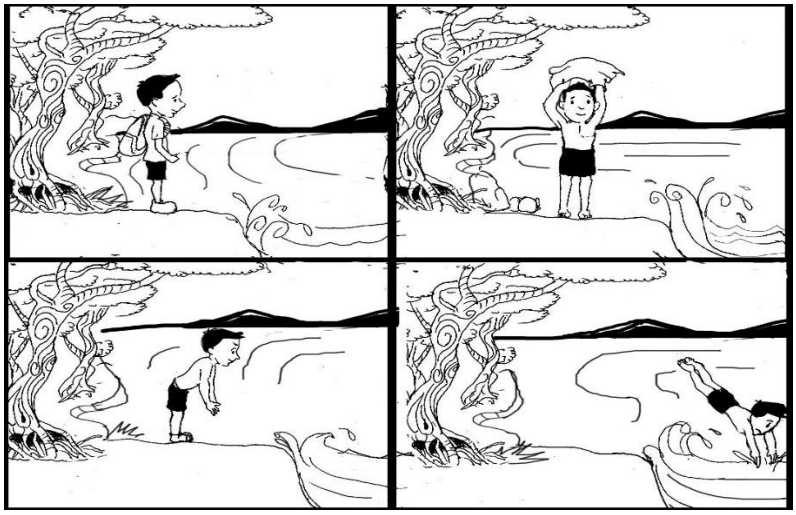


Figure 3.2: Sample picture used within the picture narration test

Participants in this test received individual corrective feedback based on their grouping. Feedback were given after the first and second testing (treatment). Scores in this test were calculated using the obligatory occasion analysis of Pica (1994), which was

also used in earlier studies (e.g., Shintain & Ellis, 2012; Rif'ah, 2012; Rezazadeh et al., 2015). Using this scoring method, researchers are able to trace the target structure development excluding the non-target structure (Rif'ah, 2012). In analysing the test, scores were checked by the researcher and another lecturer at the target site.

$$\frac{\textit{Number of correct English tense}}{\textit{Number of obligatory tense + Overused forms}} \; X \;\; 100 =$$

To determine the accuracy of the participants in the target grammatical forms, scores were encoded into SPSS and were calculated using Two way Mixed-Method Anova. The method was recommended by two statisticians. One is a visiting research fellow at University of Malaya, Academic Development Center while the other is from University of St. La Salle who has lectured statistics and research methods in undergraduate and graduate studies in the Philippines.

3.8.1.2. Data Collection and Treatment set-up/Operationalization

This section outlines the details of the tests given to the three groups. Each group was being assessed using the same type of instruments and timings, but some intricate details were intentionally varied in order to analyse their impact on learners' language learning and acquisition. The control group (CG) was given no feedback treatment, so that the effect of other groups, which had been given different feedback treatment could be compared with the control group (CG) for accurate assessment. While the treatment group 1 was provided with direct corrective feedback (DCF) for the revising results of pre- and post-test results for picture narration test, the treatment group 2 received metalinguistic corrective feedback (MCF) coupled with metalinguistic handout to enhance learners' language usage and acquisition. The complete operationalization details of the testing and data collection process have been outlined in table 3.7 given below (within the table given below, Timed Grammaticality Judgement Test is represented as TGJT, MKT refers to

Metalinguistic Knowledge Test, PNT for Picture Narrative Test, and B represents Break between two tests). The three-week experimental design scheduling was supported by an email enquiry sent by the researcher to two SLA credible researchers: Rod Ellis, SLA book author and research professor from Curtin University, and Ehsan Rassaei, assistant SLA professor at Islamic Azad University:

Table 3.7: Schedule Information for the Different Group of Learners

		Control Group					**Treatment Group 1 (DCF)**					**Treatment Group 2 (MCF)**				
Day		Background Questionnaire														
1	**Time (mins)**	8	30	25	30	45	8	30	25	30	45	8	30	25	30	45
	Pre-test	TGJT	B	MKT	B	PNT	TGJT	B	MKT	B	PNT	TGJT	B	MKT	B	PNT
4		No feedback/Revision of written test					Direct Feedback (Writing)/Revision of written test					Error coding plus handout (writing)/Revision of written test				
	Post-test	TGJT	B	MKT	B	PNT	TGJT	B	MKT	B	PNT	TGJT	B	MKT	B	PNT
5		No feedback/Revision of written test					Direct Feedback (Writing)/Revision of written test					Error coding plus handout (writing)/Revision of written test				
18	Delayed Post-test	TGJT	B	MKT	B	PNT	TGJT	B	MKT	B	PNT	TGJT	B	MKT	B	PNT

Control group (No Treatment). The control group received three tests, namely MKT, TGJT and Picture Narrative Test (PNT). In the first day (Time 1), the groups answered the three tests following the instructions of the teacher. MKT was answered in 25 minutes including the personal background questionnaire, followed by 8-minute allocated for TGJT and 45-minutes for PNT. After each test, participants were given a 30-minute break. No corrective feedback was given to the participants within the control group. The post-tests were conducted on day 4 (Time 2), in which, the teacher gave back the first PNT with its feedback, and participants were asked to rewrite it within 10 minutes. Immediately after the revision, they were given the same tests (TGJT, MKT) but with items reshuffled. The PNT was also changed by replacing the series of picture sets. The delayed post-tests were organized on Day 18 (Time 3), during which, similar tests were

given in reshuffling of queries for the third time for TGJT and MKT, while a new set of pictures were given in PNT to the participants. This was to determine if there was any long-term effect in the activity conducted.

Direct Corrective Feedback Group (Treatment 1). This group (DCF) received similar tests in day 1 (pre-test), 4 (post-test) and 18 (delayed post-test) and received individual focused corrective feedback for their picture series narrative writing. After test was administered, papers were collected and checked during the first day. For the activities in day 4, the group received their feedback from PNT. They were asked to look into their errors for 10 minutes. The feedback highlighted the incorrect and obligatory tenses, and each word was supplied with the correct form on top of each mistake. No feedback was given for TGJT and MKT as those tests were only given if the received feedback from the writing test had an effect to the acquisition of implicit and explicit knowledge of basic tenses of the participant. When feedback was received, their paper was collected again, given a blank sheet of paper and were instructed to rewrite their story. After rewriting, the teacher collected it and they were instructed to get ready for the new tests. TGJT, MKT, and new picture series test were again conducted for immediate testing. Similar procedures were followed with items reshuffled for TGJT and MKT. Feedback for their writing test was given in Day 5. After two weeks (Day 18), delayed post-test was conducted to test long term efficacy of this feedback technique. The final tests followed similar procedures: TGJT, MKT, and Picture Narrative Test (PNT).

Metalinguistic corrective feedback group (Treatment 2). The metalinguistic corrective feedback (MCF) group received similar procedures with direct corrective feedback; however, with different treatment. This group was given MCF and metalinguistic handout as a feedback. The MCF was operationalized by giving error code on top of the incorrect tenses in the participant's' writing task. The participants also received a metalinguistic handout with corresponding error codes that provide the general use of English tenses in

reference to Uchiyama's (2006) English Verb Tenses book. Participants in this group were given 5 minutes to look into the codes and incorrect tenses, and another 5 minutes to check the metalinguistic handout. Once finished checking, papers and handout were collected, and they were asked to rewrite the story. After rewriting, tests (TGJT, MKT, and PNT), items were reshuffled and administered in the similar fashion to the previous groups. The final test, i.e. the delayed post-test was conducted three weeks after the initial pre-test.

3.8.1.3. Target Structure

A host of different researches have analysed the effect of focused direct feedback strategies on target language acquisition (Bitchener & Knoch, 2009; Bitchener & Knoch, 2010a, Bitchener & Knoch, 2010b; Sheen, 2007; Sheen et al., 2009). Similar to previous studies, this research also emphasizes on analysing focused corrective feedback strategies and their impact on language learning and acquisition. A focused corrective feedback allows language teachers to primarily focus on the pre-selected target structures, while neglecting other aspects of writings during participants' assessment of language learning (Ellis et al., 2008). Relevant literature revealed that focused written corrective feedback can increase language learners' level of awareness to the different grammatical structures, while unfocused corrective feedback would tend to increase attentional load and reduce overall awareness of learners to different grammatical structures (Frear & Chiu, 2015; Sheen, 2007). However, research by Ellis et al. (2008) was able to reveal that both focused and unfocused corrective feedback were able to demonstrate effective results in terms of improving the accuracy of definite and indefinite articles within English language learning. The target structure within this research analysed a small set of grammatical issues related to the use of present simple, present continuous, past simple, and past-continuous tenses, while neglecting other errors within the writings.

3.8.1.4. Data Analysis

Test scores for TGJT, MKT and Picture series tests were entered into SPSS and two-way Mixed ANOVA was used to determine the significant differences among the groups. The method of analysis was determined by the visiting lecturer at University of Malaya who specializes in Statistics and by a former Statistics lecturer of the University of St. La Salle in the Philippines. Descriptive and inferential statistics were also calculated for the different groups in three different testing periods.

The two way Mixed Anova was used to determine the effect of the different treatment procedures in the accuracy of using basic tenses in writing, specifically in the pre-test, post test, and delayed post test. This was also used to see whether feedback used has a significant effect on the implicit and explicit knowledge of the participants. Bonferroni's post hoc pair-wise comparison, through SPSS, was also calculated to find where the differences among the groups evident.

During the analysis of the study, the mean scores of all test were considered as the dependent variable, while the different feedback and testing periods were treated as independent variable.

3.8.2. Phase 2: Sequential Explanatory Qualitative Design

3.8.2.1. Interview Instrument

A semi-structured interview using open -ended questionnaire was designed to explore students' response to the received feedback. Semi-structured interview is a formal interview with pre-set questions; however, interviewers can engage in a topical trajectories depending on the needs of the situation (Coheen & Crabtree, 2006). Moreover, according to Okaley (1999), a qualitative interview records not only the standard practices, but are also challenged and put into action. As such, many researchers use semi-structure and/or in-depth interviews. Newton (2010) also stated the purpose of using semi-interview as making use of conversation and questioning to generate insights

and reflections of the themes being investigated. Moreover, semi-structured interview is perfect in exploring complex issues to which this method allows the researcher to probe respondents for more information (Barriball & While, 1994).

When the interview in the current study was conducted, the conversation was free-flowing and natural to gain deeper insight of the responses (Patton, 2004). Since semi-structured interview allows interviewer to adjust and explain complex terminologies, the interviewer/researcher made sure that the question and terminologies used were appropriate to the respondents' proficiency level. According to Barriball & While (1994), some participants do not actively participate in interview sessions due to language barrier, as such the flexibility that this kind of interview gives provides an opportunity for the researcher and participants to mutually understand and cooperate in undertaking the activity. Table 3.8 below provides sample questions given to the participants.

Table 3.8: Sample Open-ended Questionnaire

#	Questions
1.	What did you experience during the last three weeks?
2.	How do you find the corrective feedback you received?
3.	Do you think it was effective? Why?

3.8.2.2. Data Collection

Purposive sampling, specifically total population sampling, was used in selecting the respondents to participate in a 1-3 minute interview. The sampling was done after papers were checked. The process involves examining the entire population who has undergone similar exposure, and the size of the population sharing the same exposure is very small ("Purposive Sampling", 2012). According to Creswell (2012), purposeful sampling allows the researcher to identify the central phenomenon within the population. In this method, researchers chose all the participants to be interviewed since the population is

small, and that each participant shares similar characteristics and treatment exposure. The time length was identified during the actual interview since participants had difficulty expressing themselves due to limited vocabulary. With this interview, the researcher can capture the major variations of the participants' views (Patton, 2004), and can guarantee that the sample possess the specific characteristics that the research wants to use in the study.

Interviews were recorded using a digital recorder and transcribed by a transcriptionist. Transcriptions were assigned with random case numbers, and individuals' names were removed for anonymity. Students were also informed of the procedures. The interview was conducted after the experimental study in one of the vacant classrooms at the college. Before the actual interview, the interviewer assured the interviewee of the privacy of any information gathered, and that data would be only used in the study. Stimulated recall was also used by showing the test answers of the interviewee in the three tests. This would allow them to recall the situation and reflect on their answers to the questions. During the interview, the interviewer asked 10 questions related to their experience when answering the test. Answers were recorded digitally. Since respondents had limited vocabulary, pacing was adjusted and questions were simplified. The interview sessions for all the group lasted for two days.

3.8.2.3. Data Analysis and Procedure

Thematic analysis was performed to explore and analyse the interview data collected from the participants. Braun and Clarke (2006) defined thematic analysis as the process that identifies the theme generated from the interview data. This theme, specifically the one that is related to the objectives which forms a pattern and a meaning, is being analysed and reported in the study.

Six steps were followed in conducting the data analysis. These steps were taken from the study of Braun and Clarke (2006) which was cited by more than 20,000 books and

research articles around the world as determined by google scholar. The first step included the familiarization of the collected data. Since interview answers were transcribed, the researcher has gone through with the transcripts several times before the second step, generating initial coding. The initial coding involved the researcher to get interesting data across the data set and collate them by its relevance. Transcripts were uploaded in a software to code themes and subthemes. However, due to short sentences in the answers, the coding was done manually. The third step conducted in the study involved deeper reflection and loose data interpretation since the researcher tried to look into for potential themes from the coded set of data. The fourth step allowed the researcher for further reflection and interpretation. The researcher reviewed and refined the themes by reading the collated extracts to make sure that those themes were really themes, and to know if other themes might be possible to collate. In the fifth step, names for each theme were confirmed and were given clear definitions. This refers to the essence that each theme provides. Patterns were also cleared for theory formation. Finally, since the researcher is actively positioned in the entire research activity, balance between the extracts and its investigation was made sure. Report was prepared by using carefully selected extracts from the interview. The analytical report was made sure to appropriately relate to the quantitative report of the study, the objective, and the literature review, and that it provided "a concise, coherent, logical, no repetitive, and interesting account of the story the data tell – within and across themes" (Braun & Clarke, 2006, p. 23)

3.9. Chapter Summary

The chapter discusses the design of the study, how the data was collected, analysed and interpreted. Participants have also been described providing their age, proficiency level, and the years they have been learning the target language. A study paradigm was also presented to describe the process from the initial to the final stage. As introduced in the beginning of this chapter, the study used a mixed method explanatory approach to

highlight the statistical results with the interview sessions which were analysed using thematic analysis. Steps in conducting the quantitative study was described following a quasi-experimental pre-test posttest design while qualitative study was done through interview sessions with the participants. Instruments used for each method were described explicitly, while their validity and reliability have been tested through pilot study. Nevertheless, the research procedures were presented systematically in respect to the order in conducting the study.

CHAPTER IV

RESULTS AND DISCUSSION

4.1. Introduction

This section discusses the different results from various quantitative and qualitative analyses that examined the impact of the different types of corrective feedback on learners' knowledge and language learning. The results from each test and instrument were separately identified to outline the statistical and qualitative integrity of data and highlight the overall performance of learners in the different groups in all tests. It begins by providing the data on how Metalinguistic Corrective Feedback (MCF) and Direct Corrective Feedback (DCF) affect the writing performance of the learners. This is followed by how these feedback strategies (DCF and MCF) affect the implicit and explicit knowledge among the learners. Scores in these tests have been encoded into SPSS. The primary statistical tool or method used to check the impact of different feedback strategies on learners' knowledge and acquisition is the two-way mixed-method Analysis of Variance (ANOVA) technique.

The section also detailed the learners' viewpoints expressed during interview sessions. A thematic analysis was conducted to arrive at specific themes from their experience during the experimental stage. Result from this analysis was used to support the statistical result.

4.2. Effects of MCF and DCF on Second Language Learner's (SLL) use of Basic English Tenses in Writing

This section presents the quantitative result of the study that examined the efficacy of corrective feedback in the accuracy of students' use of basic English tenses in writing. Picture Narrative Writing Test (*See Chapter 3*) was used to determine its effects which

was scored using Obligatory Case Analysis *(See Chapter 3).* Scores from this test were encoded to SPSS for two-way Mixed Method Anova calculation.

4.2.1. Picture Narrative Writing Test (PNT)

One of the first important features in analysing the results of the learners is to view the distribution of results obtained by learners in PNT. A boxplot primarily shows data variability outside the minimum and highest quartile. This indicates whether outliers are present in the group or not. Figure 4.1 below outlines the boxplot result for the three different testing periods. The x-axis shows three types of groups and their respective results, while the y-axis represents the marks of learners in the PNT. The first three set of plots correspond to the pre-test (blue bar), post-test (green bar) and delayed post-test (grey bar) on the group. The following three sets of plot in the middle belong to the MCF group, while the final set on the left-hand side belongs to the control group (CG). From the boxplot, it can be clearly seen that the primary distribution of marks for each group in the three tests are closely related, showing no outliers in the graph. In other words, this means that all scores are not numerically distant within each group. When outliers are present, they are removed from the data set since they affect the overall result of the analysis, particularly the mean score.

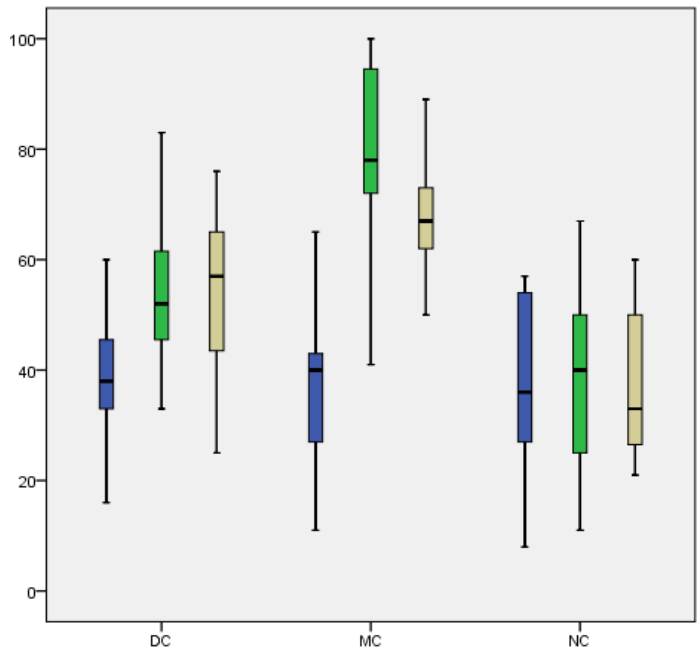


Figure 4.1: Box plot results for three test groups during pre-, post- and delayed post-test

The marginal values of each group in figure 4.2 provides a clearer representation of the overall trend of scores of each group over the course of time. In this figure, the CG group is shown in grey colour, DCF group in blue colour, and MCF group in green.

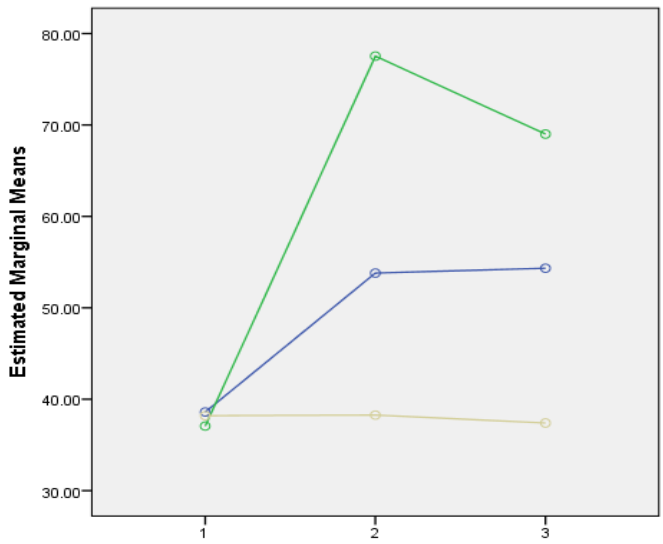


Figure 4.2: Estimated marginal mean values of different groups over the course of three tests

The overall score of the CG group marginally declined to below 40 from pre-test to delayed post-test. On the other hand, learners in the MCF group fared much better outcomes in terms of marginal mean scores for PNT. In comparing the MCF and DCF groups, the former showed considerable improvement in the second test, but the marginal mean results declined steadily after taking delayed post-test. However, even with slight decrease in results after delayed post-test, MCF group outperformed learners in the DCF group. On the other hand, the score of DCF group showed a steady improvement from below 40 to 55.

To determine the normal distribution of scores, table 4.1 outlines the breakdown of analysis from the acquired PNT data. The generated the actual score from the sample was obtained using SPSS, and this was compared to the scores obtained from the normally distributed set of scores having the same mean and standard deviation values. If there are significant differences in the results obtained, it means that the data set is widely spread out. The results are given in the table given below:

Table 4.1: Different tests of normality

		Kolmogorov-Smirnov[a]			**Shapiro-Wilk**		
	Groups	Statistic	df	Sig.	Statistic	Df	Sig.
Pre-Test	DC	.111	15	.200*	.984	15	.990
	MC	.175	15	.200*	.959	15	.669
	NC	.167	15	.200*	.915	15	.161
Immediate Post-Test	DC	.142	15	.200*	.960	15	.687
	MC	.166	15	.200*	.899	15	.091
	NC	.152	15	.200*	.950	15	.526
Delayed Post-Test	DC	.115	15	.200*	.969	15	.848
	MC	.157	15	.200*	.934	15	.308
	NC	.166	15	.200*	.888	15	.062

*. This is a lower bound of the true significance.
a. Lilliefors Significance Correction

When considering the results of the Shapiro-Wilk test given in the table 8 above, all the results of the learners are normally distributed ($p > 0.05$). This means that scores are not widely dispersed. In a similar manner, Levene's test for checking the equality of error variances across the different samples in Table 4.2 below. This provides the examination of variances across the different sample groups to ensure whether the data belong to normal distribution or not.

Table 4.2: Levene's test of equality of error variances

	F	df1	df2	Sig.
Pre-Test	1.706	2	42	.194
Post-Test	.800	2	42	.456
Delayed Post-Test	1.122	2	42	.335

Tests the null hypothesis that the error variance of the dependent variable is equal across groups.
a. Design: Intercept + Groups
Within Subjects Design: Tests

Based on the findings of the Levene's test, data are normally distributed for all the three test samples, $p > 0.05$ - namely pre-, post- and delayed post-test. For the remaining tests, MKT and TGJT, only Levene's test of equality of variance has been used as it generated similar results with Shapiro-Wilk test.

4.2.1.1. Descriptive Statistics of the three groups in PNT

In table 4.3 below, some of the salient features with regards to the descriptive statistics of the three test group. The descriptive information of the PNT shows the general variations of the participants' results for the pre-, post, and delayed post-test. Information for each of the groups includes the number of participants in each group, mean values of score and standard deviation.

Table 4.3: Descriptive statistics for PNT

	Groups	Mean	Std. Deviation	N
Pre Test	DC	38.6000	11.35656	15
	MC	37.0667	15.54471	15
	NC	38.2000	16.18288	15
	Total	37.9556	14.20140	45
Immediate Post-Test	DC	53.8000	12.55957	15
	MC	77.5333	19.05956	15
	NC	38.2667	15.13023	15
	Total	56.5333	22.47787	45
Delayed Post-Test	DC	54.3333	14.86447	15
	MC	69.0000	10.93487	15
	NC	37.4000	13.08107	15
	Total	53.5778	18.25637	45

The descriptive statistics for all the groups reveals that MCF group has the largest improvement in the mean score from pre-test (37.0667) to delayed post-test (69). This is followed by a little improvement reported by DCF (mean value of 38 in pre-test to 54.33 in delayed post-test) and there is no change in the performance of the CG group. The standard deviation, for pre-test results of the CG and MCF groups, showed high levels of variations (15.54 for MC and 16.18 for NC). For the post-test results, the level of variation was higher for MCF (19.06) and CG (15.13). For the delayed post-test, DC and NC reported higher levels of standard deviation (14.86 for DC and 13.08 for NC). Nevertheless, the result implies higher impact of MCF than DCF in terms of accuracy improvement

4.2.1.2. Within-Subject Effects between test groups and scores for PNT

To understand the significant effects of the test in each learner, within subject effect has been analyzed. Table 4.4 below sheds light on the within-subject effects between different tests and the impact of PNT scores on different learner groups. It has been revealed that there is a statistically significant relationship between test scores and groups with $F(4, 84) = 11.397$, $p = 0.000 < 0.05$ and partial $\eta^2 = 0.352$. This demonstrated that the different groups had different levels of performance in terms of answering the test on PNT.

Table 4.4: Test of Within Subject Effects between test groups and scores for PNT

Source		Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Tests	Sphericity Assumed	8968.844	2	4484.422	28.863	.000	.407
	Greenhouse-Geisser	8968.844	1.838	4880.729	28.863	.000	.407
	Huynh-Feldt	8968.844	2.000	4484.422	28.863	.000	.407
	Lower-bound	8968.844	1.000	8968.844	28.863	.000	.407
Tests * Groups	Sphericity Assumed	7082.978	4	1770.744	11.397	.000	.352
	Greenhouse-Geisser	7082.978	3.675	1927.232	11.397	.000	.352
	Huynh-Feldt	7082.978	4.000	1770.744	11.397	.000	.352
	Lower-bound	7082.978	2.000	3541.489	11.397	.000	.352
Error (Tests)	Sphericity Assumed	13050.844	84	155.367			
	Greenhouse-Geisser	13050.844	77.179	169.098			
	Huynh-Feldt	13050.844	84.000	155.367			
	Lower-bound	13050.844	42.000	310.734			

Similarly, table 4.5 revealed that the DCF group has a significant impact on the scores obtained by learners. This has been interpreted in view of the following information: $F(2, 28) = 12.05$, $p = 0.000 < 0.05$ and partial $\eta^2 = 0.463$. The outcome is also similar to the test results of MCF group where ($F(2, 28) = 42.87$, $p = 0.000 < 0.05$ and partial $\eta^2 = 0.754$). On the other hand, a slightly different output was gathered with respect to the test scores obtained by the CG group. This is because the relationship between the CG group and their scores for PNT did not reveal any significant impact on the learners' language uptake ($p = 0.983 > 0.05$).

Table 4.5: Test of Within Subject Effects between different tests for PNT

Source			Type III Sum of Squares	Df	Mean Square	F	Sig.	Partial Eta Squared
DC	Tests	Sphericity Assumed	2394.311	2	1197.156	12.050	.000	.463
		Greenhouse-Geisser	2394.311	1.490	1606.516	12.050	.001	.463
		Huynh-Feldt	2394.311	1.627	1471.433	12.050	.001	.463
		Lower-bound	2394.311	1.000	2394.311	12.050	.004	.463
	Error (Tests)	Sphericity Assumed	2781.689	28	99.346			
		Greenhouse-Geisser	2781.689	20.865	133.317			
		Huynh-Feldt	2781.689	22.781	122.107			
		Lower-bound	2781.689	14.000	198.692			
MC	Tests	Sphericity Assumed	13650.53	2	6825.267	42.87	.00	.754
		Greenhouse-Geisser	13650.53	1.738	7853.322	42.87	.00	.754
		Huynh-Feldt	13650.53	1.963	6953.092	42.87	.00	.754
		Lower-bound	13650.53	1.000	13650.533	42.87	.00	.754
	Error (Tests)	Sphericity Assumed	4457.467	28	159.195			
		Greenhouse-Geisser	4457.467	24.335	183.174			
		Huynh-Feldt	4457.467	27.485	162.177			
		Lower-bound	4457.467	14.000	318.390			
CG (NC)	Tests	Sphericity Assumed	6.978	2	3.489	.017	.983	.001
		Greenhouse-Geisser	6.978	1.780	3.920	.017	.975	.001
		Huynh-Feldt	6.978	2.000	3.489	.017	.983	.001
		Lower-bound	6.978	1.000	6.978	.017	.899	.001
	Error (Tests)	Sphericity Assumed	5811.69	28	207.56			
		Greenhouse-Geisser	5811.69	24.919	233.23			
		Huynh-Feldt	5811.69	28.00	207.56			
		Lower-bound	5811.69	14.000	415.12			

Table 4.6 below demonstrates the results of the between-subject effects of PNT to the different tests conducted at different time periods. From these data, it can be inferred that the pre-test scores do not hold any statistical significance ($F(2, 42) = 0.045$, $p = 0.956 > 0.05$ and partial $\eta^2 = 0.002$).). This supports the idea that the scores of all groups during the pre-test is almost similar.

4.2.1.3. Between-Subject Effects for PNT

The results for the post- and delayed post-tests in Table 4.6 hold considerable statistical significance, for post-test $F(2, 42) = 23.466$, $p = 0.000 < 0.05$ and partial $\eta^2 = 0.528$ and similarly, for delayed post-test, $F(2, 42) = 21.994$, $p = 0.000 < 0.05$ and partial $\eta^2 = 0.512$. In sum, the post- and delayed post-test results for all the groups to PNT have statistical importance, while the scores for pre-tests in relation to PNT do not have any

statistical significance. This implies that after the experimental procedures, scores of the learners in three groups are not already the same.

Table 4.6: Tests of Between-Subject Effects for PNT

Source		Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Pre-Test	Corrected Model	18.978[a]	2	9.489	.045	.956	.002
	Intercept	64828.089	1	64828.089	307.487	.000	.880
	Groups	18.978	2	9.489	.045	.956	.002
	Error	8854.933	42	210.832			
	Total	73702.000	45				
	Corrected Total	8873.911	44				
	Corrected Model	18.978[a]	2	9.489	.045	.956	.002
Post-Test	Corrected Model	11732.133[b]	2	5866.067	23.466	.000	.528
	Intercept	143820.800	1	143820.800	575.334	.000	.932
	Groups	11732.133	2	5866.067	23.466	.000	.528
	Error	10499.067	42	249.978			
	Total	166052.000	45				
	Corrected Total	22231.200	44				
Delayed Post-Test	Corrected Model	7502.044[c]	2	3751.022	21.994	.000	.512
	Intercept	129176.022	1	129176.022	757.426	.000	.947
	Groups	7502.044	2	3751.022	21.994	.000	.512
	Error	7162.933	42	170.546			
	Total	143841.000	45				
	Corrected Total	14664.978	44				

4.2.1.4. Multiple Comparison between different groups

Table 4.7 below shows the multiple comparisons for the different groups who took PNT. Based on the data presented, the pre-test results of the three groups hold no significance. Similarly, the results for DC and MC groups ($p = 0.001 < 0.05$), NC and MC groups ($p = 0.000$) as well as DC and NC groups ($p = 0.027 < 0.05$) for post-test and for delayed post-tests, results for DC and MC groups ($p = 0.01 < 0.05$), MC and NC groups ($p = 0.000 < 0.05$) as well as DC and NC groups ($p = 0.003 < 0.05$) are statistically significant in nature. This implies that the pre-test scores of the groups are closely related, while the immediate and delayed post-test scores are significant, showing variations in scores.

Table 4.7: Multiple Comparison between different groups with pre-, post- and delayed post-tests as dependent variables

(I) Tests		(J) Tests	Mean Difference (I-J)	Std. Error	Sig.[b]	95% Confidence Interval for Difference[b]	
						Lower Bound	Upper Bound
Pre-test	DC	MC	1.5333	5.30197	.955	-11.3478	14.4144
		NC	.4000	5.30197	.997	-12.4811	13.2811
	MC	DC	-1.5333	5.30197	.955	-14.4144	11.3478
		NC	-1.1333	5.30197	.975	-14.0144	11.7478
	NC	DC	-.4000	5.30197	.997	-13.2811	12.4811
		MC	1.1333	5.30197	.975	-11.7478	14.0144
Post-test	DC	MC	-23.7333*	5.77325	.001	-37.7594	-9.7073
		NC	15.5333*	5.77325	.027	1.5073	29.5594
	MC	DC	23.7333*	5.77325	.001	9.7073	37.7594
		NC	39.2667*	5.77325	.000	25.2406	53.2927
	NC	DC	-15.5333*	5.77325	.027	-29.5594	-1.5073
		MC	-39.2667*	5.77325	.000	-53.2927	-25.2406
Delayed Post-test	DC	MC	-14.6667*	4.76859	.010	-26.2519	-3.0814
		NC	16.9333*	4.76859	.003	5.3481	28.5186
	MC	DC	14.6667*	4.76859	.010	3.0814	26.2519
		NC	31.6000*	4.76859	.000	20.0147	43.1853
	NC	DC	-16.9333*	4.76859	.003	-28.5186	-5.3481
		MC	-31.6000*	4.76859	.000	-43.1853	-20.0147

Based on observed means:
The error term for Pre-test is Mean Square (Error) = 210.832
The error term for Post-test is Mean Square (Error) = 249.978.
The error term for Delayed Post-test is Mean Square (Error) = 170.546.
*. The mean difference is significant at the .05 level.
b. Adjustment for multiple comparisons: Bonferroni.

4.2.1.5. Pairwise Comparison of Three Groups

In table 4.8, the results of the pairwise comparison with Bonferroni adjustment between the different types of tests for the three groups have been highlighted. Based on the findings, it can be evaluated that there is no statistical significance in the results of CG group for any of the tests conducted. This can be interpreted that comments had no effects in terms of linguistic intervention and associated improvements in the language learning of the participants in the CG group.

Table 4.8: Pairwise Comparison between pre-, post and delayed post-test for different groups

	(I) Tests	(J) Tests	Mean Difference (I-J)	Std. Error	Sig.[b]	95% Confidence Interval for Difference[b]	
						Lower Bound	Upper Bound
NC	1	2	-.067	5.855	1.000	-15.979	15.846
		3	.800	4.264	1.000	-10.789	12.389
	2	1	.067	5.855	1.000	-15.846	15.979
		3	.867	5.528	1.000	-14.158	15.891
	3	1	-.800	4.264	1.000	-12.389	10.789
		2	-.867	5.528	1.000	-15.891	14.158
DC	1	2	-15.200*	4.114	.007	-26.381	-4.019
		3	-15.733*	4.160	.006	-27.040	-4.426
	2	1	15.200*	4.114	.007	4.019	26.381
		3	-.533	2.346	1.000	-6.909	5.842
	3	1	15.733*	4.160	.006	4.426	27.040
		2	.533	2.346	1.000	-5.842	6.909
MC	1	2	-40.467*	5.415	.000	-55.183	-25.750
		3	-31.933*	3.983	.000	-42.757	-21.109
	2	1	40.467*	5.415	.000	25.750	55.183
		3	8.533	4.300	.202	-3.154	20.221
	3	1	31.933*	3.983	.000	21.109	42.757
		2	-8.533	4.300	.202	-20.221	3.154

Based on estimated marginal means
*. The mean difference is significant at the .05 level.
b. Adjustment for multiple comparisons: Bonferroni.

Conversely, the MCF and DCF groups were able to benefit from their respective feedback to varying degrees. For example, the PNT scores for the pre- and post-tests ($p = 0.007 < 0.05$) as well as pre- and delayed post-tests ($p = 0.006 < 0.05$) had statistical significance for the DCF group. Similarly, for the MCF group participants, the results for pre- and post-tests ($p = 0.000 < 0.05$) as well as pre- and delayed post-tests ($p = 0.000 < 0.05$) also had statistical significance. In comparing DCF and MCF groups, the results of PNT had much more significant impact on the learning outcomes for MCF group ($p = 0.000$), when compared to the effects for the DCF group's language learning ($0.05 > p > 0.007$). These findings revealed that DCF and MCF groups had a positive impact on PNT, with MCF as higher, while no feedback as having no effect on the language learning of the CG group.

4.3. The Effect of MCF and DCF on Second Language Learners' (SLL) use of explicit and implicit knowledge of Basic English Tenses

4.3.1. Metalinguistic Knowledge test (MKT)

Like the previous section, the overall analysis of results of this test has been conducted in the same manner. The boxplot in figure 4.3 below presents the overall variation in the results of MKT obtained by the three groups. In the boxplot, a total of nine different plots have been outlined. The three plots have been devoted for each testing group corresponding to the different timings of the tests conducted. The first set of results at the left-hand side of the plot belongs to the CG group followed by the three plots for DCF group and MCF group at the left-hand side.

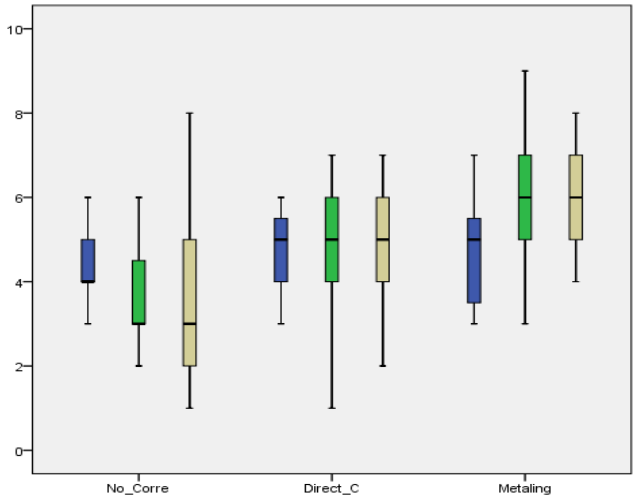


Figure 4.3: Box plot for MKT conducted at three different occasions over the period of three weeks

The box-plot result of all groups showed that there are no outliers in the data set. This means that the scores of each group in the three testing periods are almost homogenous.

In figure 4.4, the grey line corresponds to the marginal value results of the group that received MCF. Findings of CG and DCF groups are outlined in blue and green lines respectively. It is interesting to note that the results between pre- and post-test for CG and DCF groups show a downward trend, which means that the overall performance

decreased for the two groups. However, in the delayed post-test, the DC group was able to recover to the original scores obtained in the pre-test.

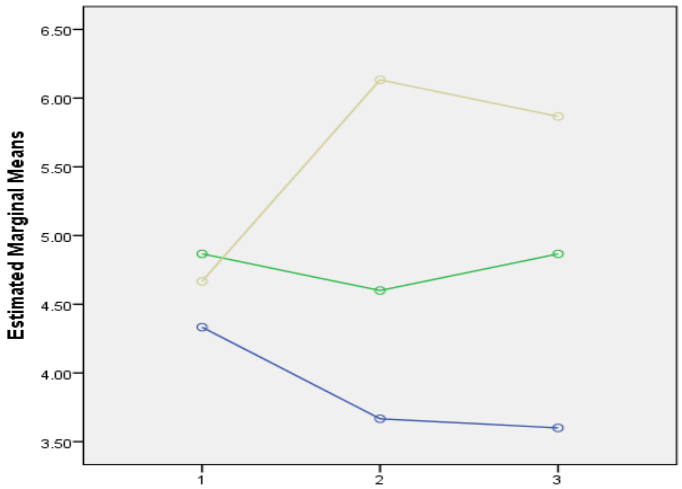


Figure 4.4: Results of the marginal value plots for the three groups

Unlike the CG and DCF groups, the MCF group was able to show considerable improvements throughout the course of the pre-, post and delayed post-test. The marginal mean of values of MCF group increased from above 4.5 to above 6.0 between pre- and post-test. Although the group showed slight reduction in the scores between post- and delayed post-tests, the post-test scores are still relatively higher when compared to the pre-test score. Levene's Test, table 4.9, was also performed for the equality of error variance. The outcome presents the extent to which the obtained test results are normally distributed. Based on the findings, it can be concluded that all MKT tests scored in three testing periods are normally distributed (as $p > 0.05$ for all the three instances). The information is outlined in figure 4.9 below:

Table 4.9: Levene's Test for examining the normal distribution of data obtained for MKT

	F	df1	df2	Sig.
Pre Test	2.295	2	42	.113
Immediate Post Test	.487	2	42	.618
Delayed Post Test	1.932	2	42	.158

Tests the null hypothesis that the error variance of the dependent variable is equal across groups.

a. Design: Intercept + Groups
Within Subjects Design: Tests

4.3.1.1. Descriptive Statistics of the three groups in MKT

The descriptive statistics in table 4.10 provides the overall mean and deviation of results for three groups.

Table 4.10: Details regarding descriptive statistics for the three groups' result for MKT

	Intervention Groups	Mean	Std. Deviation	N
Pre Test	No Corrective Feedback	4.3333	.97590	15
	Direct Corrective Feedback	4.8667	.91548	15
	Metalinguistic Corrective Feedback	4.6667	1.39728	15
	Total	4.6222	1.11373	45
Immediate Post Test	No Corrective Feedback	3.6667	1.34519	15
	Direct Corrective Feedback	4.6000	1.72378	15
	Metalinguistic Corrective Feedback	6.1333	1.40746	15
	Total	4.8000	1.79139	45
Delayed Post Test	No Corrective Feedback	3.6000	2.02837	15
	Direct Corrective Feedback	4.8667	1.45733	15
	Metalinguistic Corrective Feedback	5.8667	1.30201	15
	Total	4.7778	1.84500	45

As presented in table 4.10, the overall mean values of the three groups for pre-test were very close (clustered between 4.3 and 4.7). This suggests that before corrective feedback procedures, all groups were almost having similar test output. However, after receiving corrective feedback, there were diverging trends observed in the three groups. For example, the MCF group was able to increase their performance from a mean value of 4.7 in pre-test to 6.1 in post-test and 5.9 in delayed post-test. The groups' standard deviation also remained constant at 1.4. Similarly, for CG group, the mean values of participants' score decreased from 4.33 (with a standard deviation of 0.976) in pre-test to 3.6 (with a standard deviation of 2.02) in delayed post-test. On the other hand, the DCF group showed an almost constant result throughout the course of the three tests with a mean score of 4.87 for two (pre-test and delayed post-test) out of the total three tests. There is also a slight variation in the standard deviation from 0.91 (pre-test) to 1.72 in post-test and 1.46 in delayed post-test.

4.3.1.2. Within Subject Effects between test groups and scores for MKT

The tables below deal with the actual analysis conducted with respect to the two-way mixed method ANOVA: the within-subject, between-subject effects and multiple comparisons of the different conditions in the test.

Table 4.11: Test of within-subject effects between test scores and different groups

Source		Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Tests	Sphericity Assumed	.844	2	.422	.326	.723	.008
	Greenhouse-Geisser	.844	1.987	.425	.326	.721	.008
	Huynh-Feldt	.844	2.000	.422	.326	.723	.008
	Lower-bound	.844	1.000	.844	.326	.571	.008
Tests * Groups	Sphericity Assumed	23.111	4	5.778	4.464	.003	.175
	Greenhouse-Geisser	23.111	3.974	5.815	4.464	.003	.175
	Huynh-Feldt	23.111	4.000	5.778	4.464	.003	.175
	Lower-bound	23.111	2.000	11.556	4.464	.017	.175
Error (Tests)	Sphericity Assumed	108.711	84	1.294			
	Greenhouse-Geisser	108.711	83.463	1.303			
	Huynh-Feldt	108.711	84.000	1.294			
	Lower-bound	108.711	42.000	2.588			

Table 4.11 highlights the within-subject effects of the different group of participants (MC, NC and DC) and different tests performed by those groups. The test for within-subject involves analysis of the same subjects under different conditions (e.g. for same group type, the effect of variation in the pre-, post and delayed post-test is measured and quantified, and similarly, for same test types, the effect of variations in different groups of participants is examined). As presented in table 18, it can be seen that there was a statistically significant interaction between the different groups (CG, MCF and DCF) and their respective test scores (this can be inferred from the following: $F(4, 84) = 4.464$, $p < .005$, partial $\eta^2 = .175$). In other words, the test scores of all groups had an interaction effect with the treatment procedures following different testing periods. A more detailed analysis of MKT scored for each group is given in Table 4.12.

Table 4.12: Test of Within Subject Effects of three groups for MKT

Source			Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
DC	Tests	Sphericity Assumed	.711	2	.356	.312	.735	.022
		Greenhouse-Geisser	.711	1.895	.375	.312	.723	.022
		Huynh-Feldt	.711	2.000	.356	.312	.735	.022
		Lower-bound	.711	1.000	.711	.312	.586	.022
	Error (Tests)	Sphericity Assumed	31.956	28	1.141			
		Greenhouse-Geisser	31.956	26.525	1.205			
		Huynh-Feldt	31.956	28.000	1.141			
		Lower-bound	31.956	14.000	2.283			
MC	Tests	Sphericity Assumed	18.311	2	9.156	6.149	.006	.305
		Greenhouse-Geisser	18.311	1.971	9.290	6.149	.006	.305
		Huynh-Feldt	18.311	2.000	9.156	6.149	.006	.305
		Lower-bound	18.311	1.000	18.311	6.149	.026	.305
	Error (Tests)	Sphericity Assumed	41.689	28	1.489			
		Greenhouse-Geisser	41.689	27.593	1.511			
		Huynh-Feldt	41.689	28.000	1.489			
		Lower-bound	41.689	14.000	2.978			
CG (NC)	Tests	Sphericity Assumed	4.933	2	2.467	1.970	.158	.123
		Greenhouse-Geisser	4.933	1.799	2.743	1.970	.164	.123
		Huynh-Feldt	4.933	2.000	2.467	1.970	.158	.123
		Lower-bound	4.933	1.000	4.933	1.970	.182	.123
	Error (Tests)	Sphericity Assumed	35.067	28	1.252			
		Greenhouse-Geisser	35.067	25.181	1.393			
		Huynh-Feldt	35.067	28.000	1.252			
		Lower-bound	35.067	14.000	2.505			

Table 4.12 shows the significance of MKT for the three groups who received different experimental procedures. To measure the statistical significance of MKT scores, there is a need to analyze the significance values corresponding to the specific F-ratios. If the value of significance is less than 0.05, this means that a feedback strategy has not played an important role in improving the overall test results of a group. Referring to the above information, it can be inferred that the direct corrective feedback has no significant effect in the MKT scores of the group, as $F(2, 28) = .312$, $p > .05$ (as $p = 0.735$), and partial $\eta^2 = .022$. Similarly, no corrective feedback procedure failed to provide necessary linguistic intervention for the CG group ($F(2, 28) = 1.97$, *sig.* = 0.158 and partial $\eta^2 = 0.123$). On the contrary, the MCF group showed significance as the metalinguistic knowledge test revealed significant impact in the score of MC group, $F(2, 28) = 1.97$, *sig.* = 0.006 and partial $\eta^2 = 0.305$). This can be interpreted that the learners who received

MCF procedure were able to increase their test score performance in different testing conditions.

4.3.1.3. Between-Subject Effects for MKT

Table 4.13 below outlines the effect of between-subject factors on the overall test scores of the language learners.

Table 4.13: Test of between-subject effects with different tests as dependent variables

Source		Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Pre-Test	Corrected Model	2.178[a]	2	1.089	.873	.425	.040
	Intercept	961.422	1	961.422	770.606	.000	.948
	Groups	2.178	2	1.089	.873	.425	.040
	Error	52.400	42	1.248			
	Total	1016.000	45				
	Corrected Total	54.578	44				
	Corrected Model	2.178[a]	2	1.089	.873	.425	.040
Post-Test	Corrected Model	46.533[b]	2	23.267	10.323	.000	.330
	Intercept	1036.800	1	1036.800	459.989	.000	.916
	Groups	46.533	2	23.267	10.323	.000	.330
	Error	94.667	42	2.254			
	Total	1178.000	45				
	Corrected Total	141.200	44				
Delayed Post-Test	Corrected Model	38.711[c]	2	19.356	7.319	.002	.258
	Intercept	1027.222	1	1027.222	388.445	.000	.902
	Groups	38.711	2	19.356	7.319	.002	.258
	Error	111.067	42	2.644			
	Total	1177.000	45				

Based on the data, it can be seen that there was a statistically significant difference in students' test scores between different groups during immediate post-test of the intervention, $F(2, 42) = 10.323$, $p = .000$. Similar result was obtained during the delayed post-test, , F(2, 42) = 10.323, p=002. This shows that for some of the groups, treatment received proved to be effective in improving the overall knowledge acquisition and language learning. As a summary, the MKT results for post- and delayed post tests were statistically significant, while the output of pre-test for MKT did not hold due significance in terms of the statistical analysis conducted above.

4.3.1.4. Multiple comparisons of different groups

We can see from the table 4.14 that there is a statistically significant difference in learners' MKT scores between the CG and MCF group (2.4667±0.54821, $p < .05$) for post-test and delayed post-tests. This is also similar in the MKT post-tests between DC and MC group (1.5333±0.54821, $p = .021$, which is less than 0.05). On the other hand, the test scores of the DCF group were not significantly higher than the CG group (0.9333±0.54821, $p = .216$) in any of the three MKT tests (pre-, post- and delayed post-test). This reveals that DCF group performance did not make any improvement in their tests scores when compared to the CG group

Table 4.14: Multiple comparisons of different groups using different tests as dependent variables

(I) Tests		(J) Tests	Mean Difference (I-J)	Std. Error	Sig.[b]	95% Confidence Interval for Difference[b]	
						Lower Bound	Upper Bound
Pre-test	NC	DC	-.5333	.40786	.399	-1.5242	.4576
		MC	-.3333	.40786	.695	-1.3242	.6576
	DC	NC	.5333	.40786	.399	-.4576	1.5242
		MC	.2000	.40786	.876	-.7909	1.1909
	MC	NC	.3333	.40786	.695	-.6576	1.3242
		DC	-.2000	.40786	.876	-1.1909	.7909
Post-test	NC	DC	-.9333	.54821	.216	-2.2652	.3985
		MC	-2.4667*	.54821	.000	-3.7985	-1.1348
	DC	NC	.9333	.54821	.216	-.3985	2.2652
		MC	-1.5333*	.54821	.021	-2.8652	-.2015
	MC	NC	2.4667*	.54821	.000	1.1348	3.7985
		DC	1.5333*	.54821	.021	.2015	2.8652
Delayed Post-test	NC	DC	-1.2667	.59380	.095	-2.7093	.1760
		MC	-2.2667*	.59380	.001	-3.7093	-.8240
	DC	NC	1.2667	.59380	.095	-.1760	2.7093
		MC	-1.0000	.59380	.223	-2.4426	.4426
	MC	NC	2.2667*	.59380	.001	.8240	3.7093
		DC	1.0000	.59380	.223	-.4426	2.4426

Based on observed means:
The error term for Pre-test is Mean Square (Error) = 1.248
The error term for Post-test is Mean Square (Error) = 2.254.
The error term for Delayed Post-test is Mean Square (Error) = 2.644
*. The mean difference is significant at the .05 level.
b. Adjustment for multiple comparisons: Bonferroni.

4.3.1.5.1. Pairwise Comparison of Three Groups in Three Testing Time

Table 4.15 below highlights the pair-wise comparison between different tests with the test scores as dependent variable. For the DCF group, the MKT scores of the students were not significantly different between pre-test, immediate post-test and delayed post-test (p = 1.000). This showed that DCF had no real impact on improving the implicit knowledge of the DC group. Similar findings were revealed for the CG and MCF groups, except for the pairwise comparison of MCF group's pre- and post-tests with $p < 0.05$ which revealed statistical significance of the test scores. Specifically, only MCF group showed improvement in immediate testing of MKT.

Table 4.15: Pairwise comparison between different tests and groups with test scores as measures

	(I) Tests	(J) Tests	Mean Difference (I-J)	Std. Error	Sig.[b]	95% Confidence Interval for Difference[b]	
						Lower Bound	Upper Bound
NC	1	2	.667	.333	.196	-.239	1.573
		3	.733	.441	.357	-.467	1.933
	2	1	-.667	.333	.196	-1.573	.239
		3	.067	.441	1.000	-1.133	1.267
	3	1	-.733	.441	.357	-1.933	.467
		2	-.067	.441	1.000	-1.267	1.133
DC	1	2	.267	.431	1.000	-.904	1.437
		3	.000	.352	1.000	-.956	.956
	2	1	-.267	.431	1.000	-1.437	.904
		3	-.267	.384	1.000	-1.310	.776
	3	1	.000	.352	1.000	-.956	.956
		2	.267	.384	1.000	-.776	1.310
MC	1	2	-1.467*	.424	.011	-2.619	-.315
		3	-1.200	.470	.069	-2.477	.077
	2	1	1.467*	.424	.011	.315	2.619
		3	.267	.441	1.000	-.933	1.467
	3	1	1.200	.470	.069	-.077	2.477
		2	-.267	.441	1.000	-1.467	.933

Based on estimated marginal means
*. The mean difference is significant at the .05 level.
b. Adjustment for multiple comparisons: Bonferroni.

Based on the overall statistical analysis conducted on the effect of MKT on the different groups' test scores, it can be remarked that there was a statistically significant interaction between the groups and test scores ($F(4, 84) = 4.464$, $p < .0005$, partial $\eta^2 = .175$). Similarly, the results also revealed statistically significant difference in language

learners' MKT scores between the CG and MCF group (2.4667±0.54821, $p < .05$), as well as between the DC and MC group (1.5333±0.54821, $p = .021$). However, the test scores of the DCF were not significantly higher than the CG group in the statistical analyses conducted (M = 0.9333, SE = 0.54821, $p = .216$). Overall, it can be interpreted that CG and DC groups who received different procedures did not show any improvements as compared to the MCF group which presented an increase in test scores during immediate testing.

4.3.2. Timed Grammatical Judgment Test

The boxplot result of TGJT for the three groups have been outlined in figure 4.5. The three sets of results at the left-hand is the DCF group. On the other hand, the results at the right-hand side are CG group, while the three results in the center belong to MCF group. Regarding DCF group result, the overall progression remained very small during the three TGJT testing procedures.

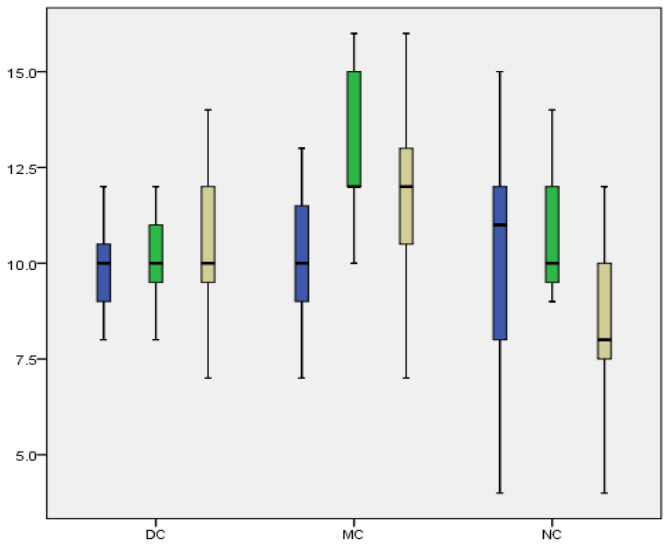


Figure 4.5: Results of the boxplot for TGJT

Basing from the data, figure 4.5 presents no extreme scores in any participants in the groups. Since no outliers were found, data analysis was continued and no participant's score was removed.

To view the overall trend of result, figure 4.6 shows the results of the three groups being treated with varying types of feedback strategies.

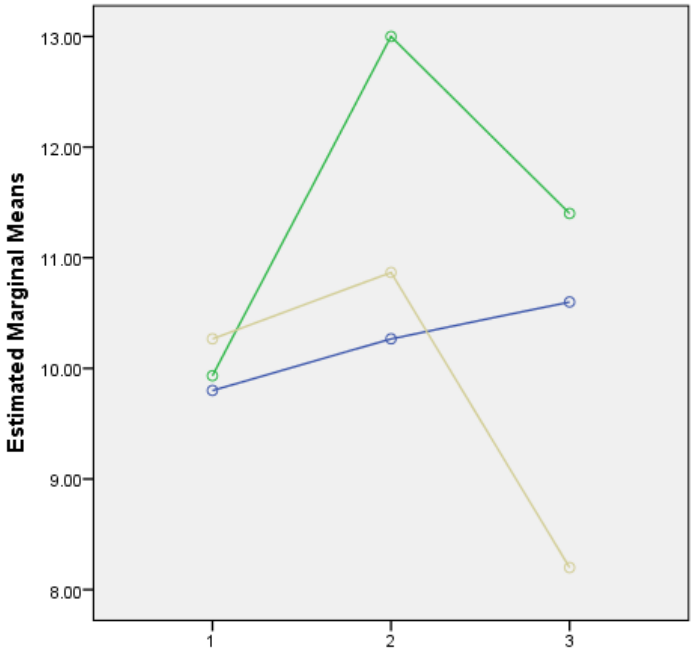


Figure 4.6: Marginal mean values of the test score results for TGJT

In figure 4.6, the green line corresponds to the marginal mean values of score for the MC group, while the blue and grey lines outline scores for DC and CG groups respectively. It can be seen from the above figure that the overall performance of the three groups improved from pre-test to post-test. However, the ability to retain the knowledge learned varied considerably between the three groups. This can be seen when DCF group showed modest yet steady improvements in the three tests. The performance of MCF group experienced the highest improvement in performance, but much of the improvement was short-lived, as the results of the delayed post-test was considerably less when compared to the post-test. It is important to note that the result of MCF group for delayed post-test was slightly higher in comparison with the results obtained by DCF group. On the other hand, although CG group showed improvement in their post-test, the significant reduction of score during the delayed post-test degraded the overall performance of the group.

Levene's test was also conducted to check if variances are equal across groups. Based on the findings, there was homogeneity of variance ($p > .05$) during the post and delayed post tests in all groups after the treatment procedures. This suggests that scores in each are closely related and normally distributed.

Table 4.16: Levene's Test of Equality of Error Variance for TGJT

	F	df1	df2	Sig.
Pre Test	5.152	2	42	.010
Immediate Post Test	2.379	2	42	.105
Delayed Post Test	.713	2	42	.496

Tests the null hypothesis that the error variance of the dependent variable is equal across groups.
a. Design: Intercept + Group
Within Subjects Design: Tests

4.3.2.1. Descriptive Statistics of the three groups in TGJT

Table 4.17 below provides some valuable information about the statistics pertaining to the TGJT performed in three weeks. The overall mean performance of DCF group remains relatively constant throughout the three tests (with mean values ranging between 9.8 and 10.6) and standard deviation ranging between 1.08 and 1.80. In contrast, MCF group's mean score showed considerable increase between 9.93 in pre-test to 13.0 in post-test, but the knowledge acquired was lost in the delayed post-test with mean score at 11.4 and standard deviation ranging between 1.86 and 2.69. On the other hand, the mean scores for the CG group declined with progression from pre-test (M=10.3) to delayed post-test (with mean value of 8.2) with standard deviation ranging between 1.68 (post-test) to 2.86 (pre-test). This suggests that only MCF group showed improvement against the base score when compared with the other groups.

Table 4.17: Descriptive statistics of TGJT for pre-, post, and delayed post-tests

	Group	Mean	Std. Deviation	N
Pre Test	DC	9.8000	1.08233	15
	MC	9.9333	1.86956	15
	NC	10.2667	2.86523	15
	Total	10.0000	2.03381	45
Immediate Post Test	DC	10.2667	1.33452	15
	MC	13.0000	2.03540	15
	NC	10.8667	1.68466	15
	Total	11.3778	2.04816	45
Delayed Post Test	DC	10.6000	1.80476	15
	MC	11.4000	2.69391	15
	NC	8.2000	2.30527	15
	Total	10.0667	2.63197	45

4.3.2.2. Within Subject Effects between test groups and scores for TGJT

In analyzing the within-subject effects, table 4.18 highlights the relationship between individuals in a group and their scores in TGJT. The results revealed statistically significant relationship between groups and their scores, $F(2, 84) = 7.33$, $p = 0.001$ ($p < 0.05$) and partial $\eta^2 = 0.149$. In short, it can be concluded that individuals in each group made changes in their performance over time.

Table 4.18: Test of within-subject effects using test scores as measures

Source		Type III Sum of Squares	Df	Mean Square	F	Sig.	Partial Eta Squared
Tests	Sphericity Assumed	54.326	2	27.163	7.332	.001	.149
	Greenhouse-Geisser	54.326	1.919	28.310	7.332	.001	.149
	Huynh-Feldt	54.326	2.000	27.163	7.332	.001	.149
	Lower-bound	54.326	1.000	54.326	7.332	.010	.149
Tests * Groups	Sphericity Assumed	79.807	4	19.952	5.385	.001	.204
	Greenhouse-Geisser	79.807	3.838	20.795	5.385	.001	.204
	Huynh-Feldt	79.807	4.000	19.952	5.385	.001	.204
	Lower-bound	79.807	2.000	39.904	5.385	.008	.204
Error (Tests)	Sphericity Assumed	311.200	84	3.705			
	Greenhouse-Geisser	311.200	80.595	3.861			
	Huynh-Feldt	311.200	84.000	3.705			
	Lower-bound	311.200	42.000	7.410			

Table 4.19 below outlines the salient features of the within-subject effects for the three test groups. For the DCF group, no statistically significant impact of TGJT could be revealed, $F(2, 28) = 1.596$, $p = .221$, partial $\eta^2 = .102$. Conversely, the data outlined for the MCF group showed a statistically significant results ($F(2, 28) = 9.68$, $p = .001 < 0.05$, and partial $\eta^2 = .409$). This means that MCF procedures had a positive impact on the scores of learners within the MCF group. Similarly, for the CG group, the test score for TGJT showed a statistically significant result, specifically $F(2, 28) = 4.933$, $p = .015 < 0.05$, and partial $\eta^2 = .261$. However, the impact of the treatment on MCF group's performance was greater as compared to its impact on the CG group.

Table 4.19: Test of Within Subject Effects of three groups for TGJT

	Source		Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
DC	Tests	Sphericity Assumed	4.844	2	2.422	1.596	.221	.102
		Greenhouse-Geisser	4.844	1.724	2.810	1.596	.224	.102
		Huynh-Feldt	4.844	1.943	2.493	1.596	.221	.102
		Lower-bound	4.844	1.000	4.844	1.596	.227	.102
	Error (Tests)	Sphericity Assumed	42.489	28	1.517			
		Greenhouse-Geisser	42.489	24.134	1.761			
		Huynh-Feldt	42.489	27.208	1.562			
		Lower-bound	42.489	14.000	3.035			
MC	Tests	Sphericity Assumed	70.578	2	35.289	9.679	.001	.409
		Greenhouse-Geisser	70.578	1.643	42.948	9.679	.002	.409
		Huynh-Feldt	70.578	1.833	38.504	9.679	.001	.409
		Lower-bound	70.578	1.000	70.578	9.679	.008	.409
	Error (Tests)	Sphericity Assumed	102.089	28	3.646			
		Greenhouse-Geisser	102.089	23.007	4.437			
		Huynh-Feldt	102.089	25.662	3.978			
		Lower-bound	102.089	14.000	7.292			
NC	Tests	Sphericity Assumed	58.711	2	29.356	4.933	.015	.261
		Greenhouse-Geisser	58.711	1.814	32.363	4.933	.018	.261
		Huynh-Feldt	58.711	2.000	29.356	4.933	.015	.261
		Lower-bound	58.711	1.000	58.711	4.933	.043	.261
	Error (Tests)	Sphericity Assumed	166.622	28	5.951			
		Greenhouse-Geisser	166.622	25.398	6.560			
		Huynh-Feldt	166.622	28.000	5.951			
		Lower-bound	166.622	14.000	11.902			

4.3.2.3. Between-Subject Effects for TGJT

When it comes to between-subject effects, table 4.20 provides data of the test scores for the different tests taken at different time. The analysis revealed that the pre-test

of all groups does not have any statistically significant effects on the TGJT scores. On the other hand, test scores for post- and delayed post-tests of all groups reveal statistically significant effects. Immediate post-test result showed $F(2, 42) = 10.599$, $p = 0.000$ and partial $\eta^2 = 0.355$. Delayed post-test, nevertheless, outlines significance where $F(2, 42) = 7.884$, $p = 0.001$ and partial $\eta^2 = 0.273$. Data gathered from this table could be interpreted that individual performance of each group has differences between each other after the treatment procedure.

Table 4.20: Test of between subject effects for TGJT

Source		Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Pre-Test	Corrected Model	1.733[a]	2	.867	.202	.818	.010
	Intercept	4500.000	1	4500.000	1048.447	.000	.961
	Groups	1.733	2	.867	.202	.818	.010
	Error	180.267	42	4.292			
	Total	4682.000	45				
	Corrected Total	182.000	44				
	Corrected Model	1.733[a]	2	.867	.202	.818	.010
Post-Test	Corrected Model	61.911[b]	2	30.956	10.599	.000	.335
	Intercept	5825.422	1	5825.422	1994.574	.000	.979
	Groups	61.911	2	30.956	10.599	.000	.335
	Error	122.667	42	2.921			
	Total	6010.000	45				
	Corrected Total	61.911[a]	2	30.956	10.599	.000	.335
Delayed Post-Test	Corrected Model	83.200[c]	2	41.600	7.884	.001	.273
	Intercept	4560.200	1	4560.200	864.298	.000	.954
	Groups	83.200	2	41.600	7.884	.001	.273
	Error	221.600	42	5.276			
	Total	4865.000	45				
	Corrected Total	304.800	44				

4.3.2.4. Multiple-level comparison between different test groups

Table 4.21 below highlights the multiple-level comparison between the different tests and groups. A general overview of the data shows that there was no link between the test scores of different groups over the course of three weeks, except for the delayed

post test between MCF and NCF groups, p= .001. This suggests that MCF and CG groups had high level of differences when tested in the long run.

Table 4.21: Multiple-level comparison between different test groups with test timing as dependent variable

(I) Tests		(J) Tests	Mean Difference (I-J)	Std. Error	Sig.[b]	95% Confidence Interval for Difference[b]	
						Lower Bound	Upper Bound
Pre-test	DC	MC	-.1333	.75649	.983	-1.9712	1.7046
		NC	-.4667	.75649	.812	-2.3046	1.3712
	MC	DC	.1333	.75649	.983	-1.7046	1.9712
		NC	-.3333	.75649	.899	-2.1712	1.5046
	NC	DC	.4667	.75649	.812	-1.3712	2.3046
		MC	.3333	.75649	.899	-1.5046	2.1712
Post-test	DC	MC	-.1333	.75649	.983	-1.9712	1.7046
		NC	-.4667	.75649	.812	-2.3046	1.3712
	MC	DC	.1333	.75649	.983	-1.7046	1.9712
		NC	-.3333	.75649	.899	-2.1712	1.5046
	NC	DC	.4667	.75649	.812	-1.3712	2.3046
		MC	.3333	.75649	.899	-1.5046	2.1712
Delayed Post-test	DC	MC	-.8000	.83874	.610	-2.8377	1.2377
		NC	2.4000*	.83874	.018	.3623	4.4377
	MC	DC	.8000	.83874	.610	-1.2377	2.8377
		NC	3.2000*	.83874	.001	1.1623	5.2377
	NC	DC	-2.4000*	.83874	.018	-4.4377	-.3623
		MC	-3.2000*	.83874	.001	-5.2377	-1.1623

Based on observed means:
The error term for Pre-test is Mean Square (Error) = 4.292
The error term for Post-test is Mean Square (Error) = 2.921.
The error term for Delayed Post-test is Mean Square (Error) = 5.276.
*. The mean difference is significant at the .05 level.
b. Adjustment for multiple comparisons: Bonferroni.

4.3.2.5. Pairwise Comparison of Three Groups in Three Testing Time

Similar to the other tests, pairwise comparison has also been conducted as presented in table 4.22. For the DCF group, the TGJT scores were not statistically significant between pre-test, immediate post-test and delayed post-test, $p > .01$. Similarly, TGJT scores for the CG group did not reveal any statistically significant relationship for the tests conducted, except for the multiple comparison between post- and delayed post-tests, $p = 0.009 < 0.01$. On the other hand, MCF group's performance in relation to the pre- and post-tests for TGJT showed statistically significant results, specifically for pre- and post-tests, as $p = 0.000$. Therefore, it can be concluded that TGJT scores are statistically significant in terms of post and delayed post-tests of CG group and pre- and

post-tests for MCF group. The significance in scores is largely due to the high level of differences of scores between each testing time.

Table 4.22: Pairwise comparison of different groups with test scores as measures

(I) Tests		(J) Tests	Mean Difference (I-J)	Std. Error	Sig.[b]	95% Confidence Interval for Difference[b]	
						Lower Bound	Upper Bound
NC	1	2	-.600	.925	1.000	-3.113	1.913
		3	2.067	.988	.165	-.618	4.751
	2	1	.600	.925	1.000	-1.913	3.113
		3	2.667*	.741	.009	.653	4.681
	3	1	-2.067	.988	.165	-4.751	.618
		2	-2.667*	.741	.009	-4.681	-.653
DC	1	2	-.467	.435	.904	-1.649	.715
		3	-.800	.527	.455	-2.233	.633
	2	1	.467	.435	.904	-.715	1.649
		3	-.333	.374	1.000	-1.349	.682
	3	1	.800	.527	.455	-.633	2.233
		2	.333	.374	1.000	-.682	1.349
MC	1	2	-3.067*	.521	.000	-4.482	-1.652
		3	-1.467	.729	.192	-3.449	.516
	2	1	3.067*	.521	.000	1.652	4.482
		3	1.600	.809	.204	-.600	3.800
	3	1	1.467	.729	.192	-.516	3.449
		2	-1.600	.809	.204	-3.800	.600

Based on estimated marginal means
*. The mean difference is significant at the .05 level.
b. Adjustment for multiple comparisons: Bonferroni.

Based on the information revealed from the wide array of different tables containing varying statistical analyses, it can be claimed that there was a statistically significant difference in students' TGJT scores between the NC and MCF groups (2.1333±0.62402, $p < .05$), and between the DCF and the MCF groups (2.7333±0.62403, $p = .000$). However, the test scores of the DCF group were not significantly higher than the TGJT scores for the CG group (0.6000±0.62403, $p = .605$). Hence, TGJT scores have a positive impact on MCF group, followed by lesser impact on the CG group and very negligible improvements in terms of language learning for the DCF group. It can also be interpreted that those students who received MCF procedures have gained better learning gains as compared to the other groups.

4.4. Second Language Learners' view of the received corrective feedback

An open-ended questionnaire (See Chapter 3) was developed to gain insights on how corrective feedback facilitated students' accuracy in using English basic tenses in writing and how this feedback relates to explicit and implicit knowledge acquisition. The goal was to solicit reactions from the participants regarding the experimental procedures received, specifically identifying the dominant themes that provided concerns on the issue of corrective feedback efficacy.

The information obtained from the interviews for each group's participants was analysed thematically (See chapter 3). The analysis procedure considered the prominent ideas expressed during the conversation, understand and explore them, and generate common themes and their relationship (Thomas & Harden, 2008). Synthesized set of statements were provided to support the generated themes (Willig, 2001). The section begins by presenting the result analysed from the CG group, followed by DCF group, then MCF group.

4.4.1. Control group thematic result

After analysing initial themes, codes and categories developed from the interviews, three major themes were generated and finalized. These were outlined in the table below:

Themes
1. Ambiguity 2. Past Feedback Experience 3. Feedback Seeking

A graphical presentation of themes in Figure 4.7 was also presented to show the percentage of the finalized themes.

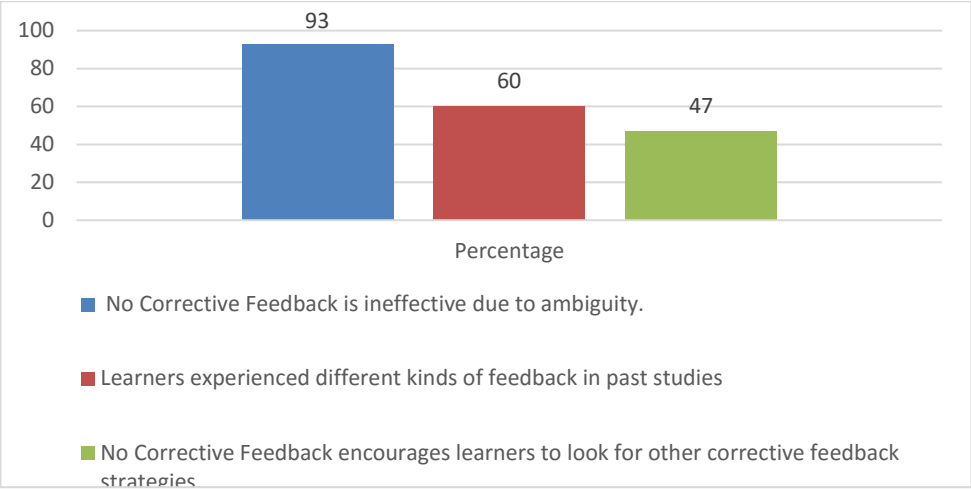


Figure 4.7: Percentage of learners' engagement to different themes

Ambiguity: Ambiguity is a state of uncertainty in meaning and intention ("Ambiguity", 2011). Nearly all the participants (14 out of 15) said that the comments provided were not clear. They did not really know what to do. The fact that no comments nor explanation was provided, these participants had to rely on their instincts or on their own understanding just to correct their errors.

Student C1 - *"No. I don't actually know why my answer is wrong."*

Student C10 - *"No. I don't know. I don't know why it is wrong also."*

Student C14 - *"I cannot understand. Difficult to know. It's not telling me what [why] I'm wrong"* -

As expressed above, due to the vagueness of getting no corrective feedback, some learners claimed that it was not helpful because they could not figure out why their answers were wrong and how to put them right. As there was no comment nor example provided, they could not compare their answer with the correct one.

Student C3- *"I cannot understand what is wrong in my paper."*

Student C15- *"Just like comment but no answer. We don't know what is wrong of where is wrong... Not good…Because we don't know how to answer or what is the correct answer"*

Feedback Experience in the Past: When participants were asked about the feedback received in their previous school, 9 out 15 (60%) expressed the kind of feedback they had received during their secondary education. Several of them stated that in high school, they were provided with the correct answer and a discussion to explain the right and wrong answers.

Student C7- *"Give the correct answer"*

Student C9- *"High School teacher he give…she give[s] the question, and we…we student[s] answer the question. And after that discuss the correct answer."*

Some also mentioned that their teachers just highlight their errors but no correction happens:

Student C3- *"My teacher gives me like that. Just wrong, wrong"*

Feedback Seeking: Almost half (47%) of the participants stated that they want to receive answer keys and explanations to understand the correct and wrong answers. In addition, they expressed that teachers should provide with clues to facilitate learning.

Student C8- *"I think clue [would be helpful] … so I can imagine it [figure out the correct answer]"*

4.4.2. Direct Corrective Feedback

Participants who received direct corrective feedback were asked on their views regarding the kind of feedback provided to them by the teacher.

After analysing initial themes, codes and categories developed from the interviews, four major themes were generated and finalized.

Themes
1. Feedback Experience 2. Critical thinking not encouraged 3. Lacking explanation 4. Other Feedback Preference 5. Error identification efficacy

Figure 4.8 below also provides a clear overview of the themes generated with regards to their percentage of occurrence during the interview sessions.

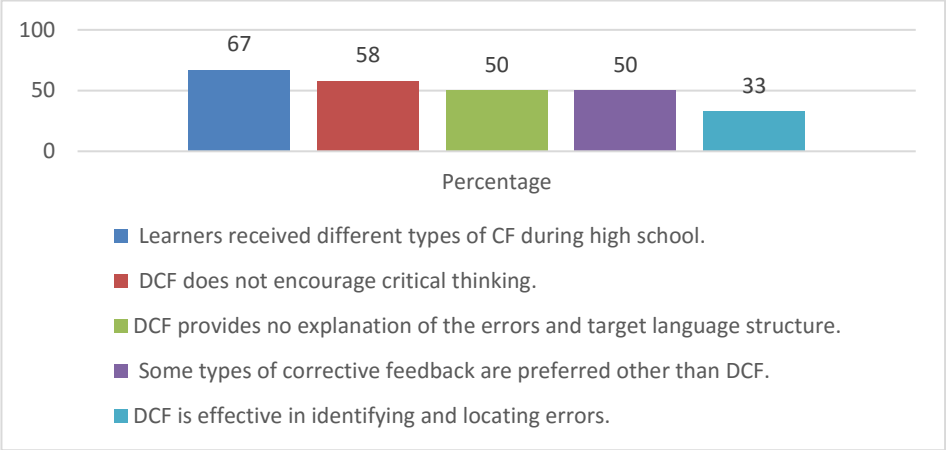


Figure 4.8: Direct Corrective Feedback Themes

Feedback Experience in High School: When participants were asked about the kind of feedback they received in the past, eight out of twelve participants (67%) had some past experiences in high school. Some of the participants stated that they had simply received the correct answer on their test paper from their teachers.

Student DCF3- *"I receive the correct answer."*

Student DCF10- *"He like…he do[es] like this. I just [get the answer] what I cannot [see]…[is] what's wrong with my [answer]…"*

Other participants also stated that they were not given the correct answer and their teacher simply told them that their answers were incorrect and that they were expected to conduct some research.

Student DCF9- *"He just crossed and didn't give the correct answer. And [he] asked us to find [the] answer by your…by myself."*

Student DCF8- *"My teacher only just put [marked] the wrong…put the wrong and I need to find [the] correct answer."*

Critical thinking not encouraged: Seven out of twelve participants (58%) receiving DCF for their incorrect answers felt that this form of feedback encouraged laziness rather than researching or providing critical look and understanding. This was mainly felt because the correct answers were given instead of a discussion or an explanation. Hence,

it can be also inferred that a student might know what the correct answer is but may not know why it is correct. This also suggests rote learning.

Student DCF2- *"Actually, it is helpful but I think It will be too easy like that. I will not think more why."*

Student DCF5 - *"Maybe. But I cannot think more of why the answer is like that. ... Because the teacher did not explain it. Just give the correct answer"*

Student DCF7- *"Not good because the test…the test on avoid…avoid the other students to research their answer. It makes students lazy."*

Lacking explanation: In this regard, some of the participants (50%) viewed DCF to be ineffective. Specifically, they felt that DCF was ineffective since no explanation was provided for the correct target language structure. Aside from that, it can also be implied that it only encouraged the students to memorize correct answers resulting to lesser interaction in error treatment.

Student DCF3 - *"I don't really know what is wrong but I just follow the correct answer that the teacher gave… I cannot think why because the teacher did not explain"*

Student DCF4- *"No [not like it]. Because the teacher did not explain it. So I don't know why it is also wrong."*

Student DCF11- *"Not really [it doesn't help] because the answer is already given... Well I can't…. I mean it was okay but it cannot really help me. Because there is no explanation why. We don't know what really is the reason."*

The term vague in this study needed to be understood carefully in order to figure out just what these participants referred to. Vague here meant that the DCF provided in the shape of a correct answer was not enough for these participants. They could not figure out why it was correct or under what grammatical rule the answer provided was thought correct. Adding an explanation or a discussion would be more beneficial to the students.

Other Feedback Preference: Due to the dissatisfaction taken from the administered DCF, there were students who began to express their preference for another type of corrective feedback. Half of the participants (6 out of 12 participants) wanted corrective

feedback that could explain why their answer was wrong, and feedback which could encourage critical thinking.

Student DCF1- *"But I need some clues so I can find it by myself later. Maybe I can easily forget if the teacher will just give me answer."*

Student DCF8- *"May[be] give me a clue…Because I need to know how to…where is the right answer on a…on my…on my [paper]…"*

Student DCF9 - *"I think no because I just saw the correct answer, but I don't know why, why it should be the answer."*

Error Identification Efficacy: On the other hand, 4 out of 12 participants (33%) also expressed that the DCF provided on this test promoted learning as it helped participants figure out what the correct answer was. These four respondents felt that having the correct answer would help them correct their mistakes easily. However, some of these respondents also agreed that in the long run, it would be better for them to have codes as it would help them think critically.

Student DCF2- *"Actually, it is helpful but I think It will be too easy like that. I will not think more why."*

"If the teacher gives like this [the correct answer], I can think easily."

Student DCF8- *"Yes. It's helpful for me to know how…what my wrong [is my mistake]. Yes, it will help me to think more about all…about my grammar."*

4.4.3. Metalinguistic Corrective Feedback

Participants in this group received metalinguistic corrective feedback with a handout. Similar to the other groups, this group received three tests tested in different times.

When this group was asked about the kind of feedback they received, three major themes emerged.

Themes
1. Effective and critical thinking promoted 2. Past Feedback Experience 3. Handout effectiveness

The themes were generated through their frequency of the codes generated from the interviewees. This could be best viewed by Figure 4.9 below.

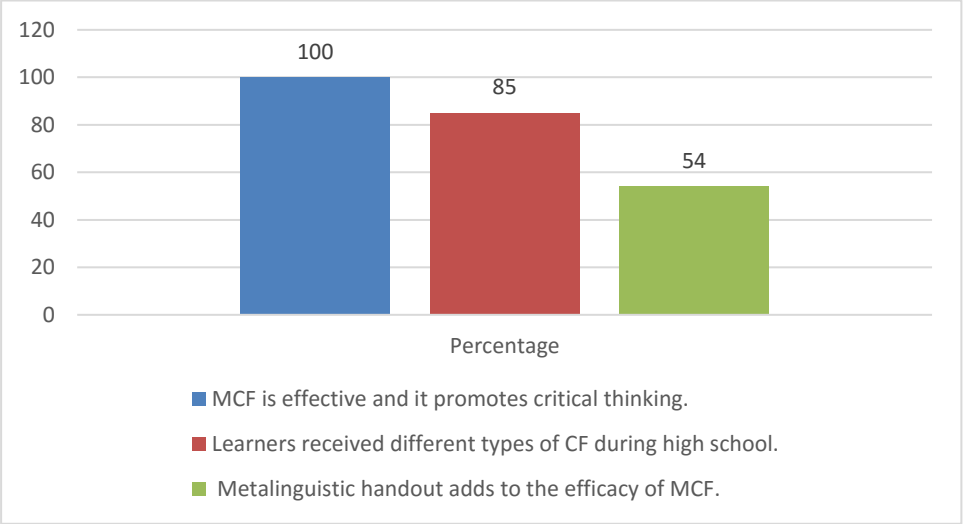


Figure 4.9: Metalinguistic and Handout Corrective Feedback Themes

Effective and critical thinking promoted: All participants expressed that the MCF was effective. Students said that they could evaluate their errors and correct them independently. Some participants also expressed their interest to this kind of feedback.

Student MCF3 - *"I think the same. The handout is helpful with the codes too."*

Student MCF4 - "*I think it's very helpful because we will independence [independently] do [the task]*

Student MCF5- *"Yes [helpful], because of the clues. I also check it with the handout. I use the clues [to figure out the correct answers"*

Student MCF6 - *"For me, it is helpful because I…he can make me try to find the answer. It makes me interested to see the codes."*

One student also mentioned the suitability of this feedback to those with higher level of proficiency. However, one also countered it by acknowledging handout to assist those who were in lower levels.

Student MCF4 - "I think it's helpful for the intelligent people. We need to understand the code."

Student MCF3 - "*It was helpful for beginners if the code is coupled with handout"*

Past Feedback Experience: With regards to this test, participants were also asked about the kind of feedback they received during their high school. A large proportion of them claimed that the feedback they were receiving now was more in detail and helpful than what they received in the past. This type of feedback helps them think through how and where they went wrong.

Student MCF7 - *"Different. My previous teacher just give [gave] me a code."*

Student MCF8 - *'The code no because if when I [was in] high school the teacher just say [said] my answer is correct…is uncorrect [incorrect]."*

Student MCF10 - *"No. Sometimes he (teacher) will ask me to…to see my friend's paper who you get high mark[s]."*

Handout effectiveness: While all agrees to the effectiveness of MCF, some of the participants also recognized the benefits of using the handout. They could refer to the note and figure out their errors. The handout also served them with details on what target structures needed to be learnt and on why such particular structure is incorrect. For them, this made it easier to understand the rules by matching the error codes to the rule codes.

Student MCF2- *"I can refer to the note that you give so that I know the…how to find the true answer."*

Student MCF3- "*It was helpful for beginners if the code is coupled with handout"*

Student MCF3- "*For, for first sentence, I think the mistake, but the much more grammar mistake,I don't know how to correct. I use the handout the teacher gives [gave]."*

4.5. Analysis and Discussion of Quantitative and Qualitative Result

The research study serves three purposes: (i) to identify the comparative effects of the two corrective feedbacks on learners' accuracy of basic English tenses in narrative writing; (ii) to identify the effects of corrective feedback on learners' explicit and implicit knowledge of basic English tenses; and (iii) to identify learners' views on the corrective feedback received. Considering these purposes, the study made use of statistical and thematic analyses to provide answers to the questions formulated out of the study

purposes. The combination of the results would validate the findings through further support of the literature review.

The previous section has provided a comprehensive analysis of the data collected. It also examined the numerical data and the qualitative thematic analysis for the interviews. In this section, the insights highlighted in the previous section will be reiterated and summarized in terms of their implications for the different groups (MCF, CG and DCF groups) and their interaction with the different types of tests conducted, namely the TGJT, MKT and PNT (See Chapter 3). Specifically, the findings of the statistical analyses noted in PNT would be interpreted to answer the first research question on the efficacy of the employed corrective feedback in improving participants' accuracy in using basic English tenses. Subsequently, it would also discuss and interpret how these feedback strategies influence the implicit and explicit knowledge of the participants in these tenses. The generated themes from the qualitative study would also be presented to support the interpretation of the statistical results. This makes the findings more credible and valid. The following table outlines some of the descriptive features of the different groups and the associated statistical analyses conducted:

Table 4.23: Descriptive Statistics for all of the tests and groups involved in this study

Tests	**Group**		**Pre-test**		**Post-test**		**Delayed Post-test**	
		N	**M (%)**	**SD**	**M (%)**	**SD**	**M (%)**	**SD**
PNT	**DC**	15	38.6	11.4	53.8	12.6	54.3	14.9
	MC	15	37.1	15.5	77.5	19.1	69.0	10.9
	CG	15	38.2	16.1	38.3	14.2	37.4	22.5
MKT	**DC**	15	48.7	9.15	46.0	17.2	48.7	14.6
	MC	15	46.7	13.4	61.3	14.1	58.7	13.0
	CG	15	43.3	9.75	36.7	13.5	36.0	20.2
TGJT	**DC**	15	75.4	8.3	79	10.2	81.5	13.8
	MC	15	76.1	14.4	100	15.7	87.7	20.7
	CG	15	78.9	22.0	83.6	12.9	63.1	20.2

As shown in table 4.23, the overall mean and standard deviation values are given as percentages, as different tests employed different scales for evaluating learners' ability to acquire learning and knowledge. Mean scores of all groups in different tests showed close scores among each other. Specifically, the highest mean score as shown in PNT was from DCF group, M=38.6 while the lowest score was from MCF group, M=37.1. The MKT mean scores also showed that the DCF group has the highest, M=48.7 while the lowest score was in CG Group, M=43.3. The table also shows that CG group scored the highest in TGJT with M=78.9 while the DCF group had the lowest, M= 75.4. From this data, it can be implied that these learners were in the same level of proficiency which supports the need of homogeneity prior to the conduct of experimental study.

It is also necessary to examine the data in a graphical form which demonstrates the progression in the improvement of language learning among the three groups in the three different tests. Figure 4.10 shows all the important results of the mean performance of the learners in the three different groups. It highlighted an in-depth information that also illustrates inter- and intra-group comparisons.

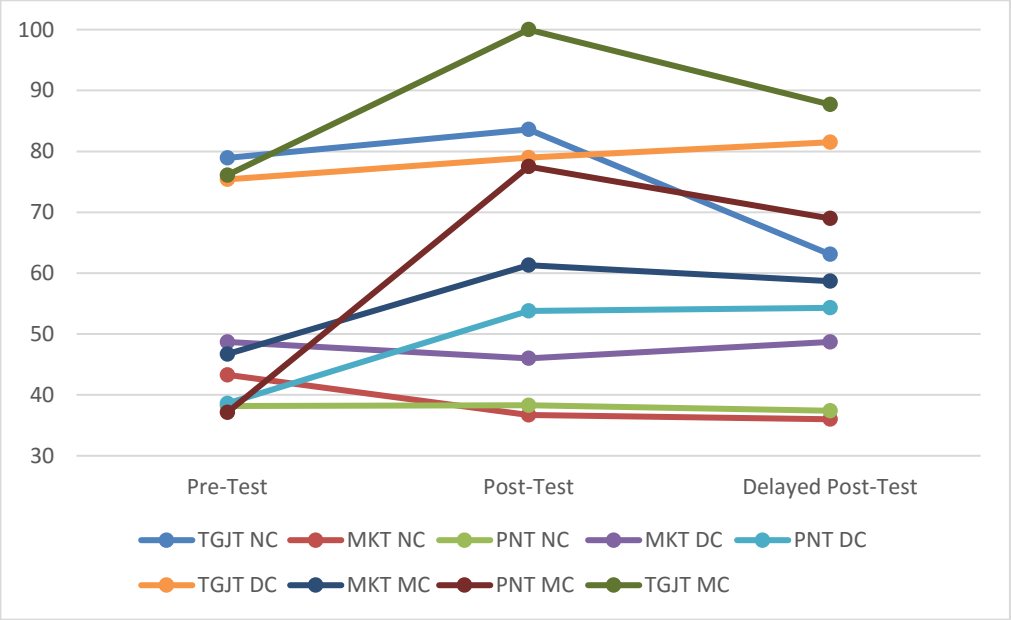


Figure 4.10: Mean Performance Of The Three Groups Over The Period Of Time

4.5.1. Discussing the effects of corrective feedback on PNT scores

The **first purpose** of the study was to answer the question of the possibility of linking corrective feedback in enhancing students' accuracy of using basic English tenses in writing. It is important to review that the purpose of Picture Series Narrative Test (PNT) (See Chapter 3) was to determine the impact of direct and metalinguistic (with handout) corrective feedback on students' accuracy of using the specified target grammatical feature. The TGJT and MKT were employed to analyze the implicit and explicit knowledge acquisition in second language acquisition (SLA).

Figure 4.11 below provides the mean scores of the three groups in three different times. The figure would show how scores increase and decrease during the testing, and give us an insight on how treatment procedures affected the overall performance.

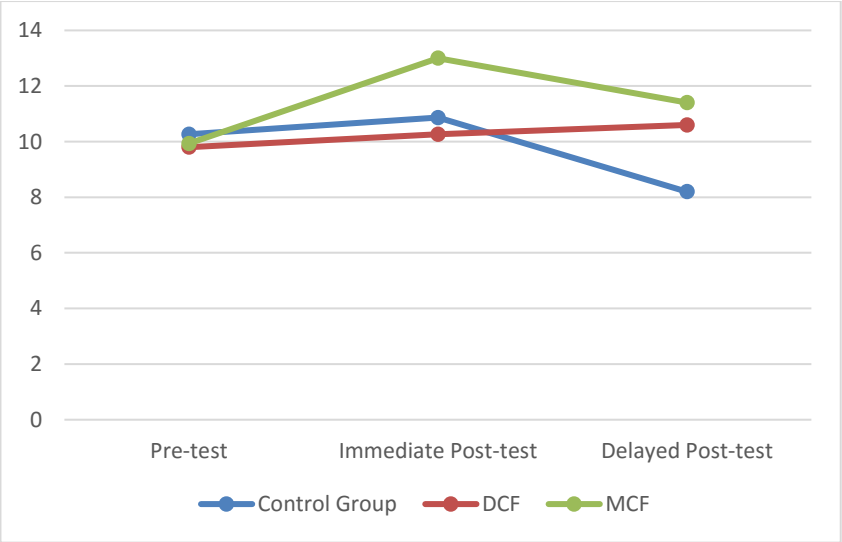


Figure 4.11 : Mean scores of the Picture Narrative Writing test

With respect to the efficacy and comparison of the MCF and DCF in PNT, it can be observed that the test scores of **MCF group performed higher than the DCF group**. As evident in Figure 14, by a statistically significant margin, it implies that the metalinguistic feedback is much more effective in improving the accuracy of Malaysian students in using English tenses in their writing. The Control group (CG) did not show any significant differences across their test scores in the three testing points. However, the DCF

group test scores showed significant differences in these time points: between pre-test and immediate post-test, p=.007, between pre-test and delayed post-test, p=.006. However, the DCF group's scores from immediate to delayed post-test did not show any significant difference, p= 1.000. With the MCF group, test scores showed a significant difference from pre-test to post-test, p= .000, and from pre-test to delayed post-test, p= .000. Similar with the DCF group's tests scores, result from immediate post-test to delayed post-test showed no significant differences, p= .202. From these results, it can be assumed that both treatments were effective in improving the accuracy of students, and were able to retain its effects after the three weeks. It should also be noted that learners in the MCF group performed higher in immediate post-test as compared to delayed post-test. This claim is further supported when between group comparison was analysed. The immediate post-test scores of DCF and MCF groups showed statistically significant difference (23.77325±5.77325), $p < .001$ as well as in delayed post-test scores (14.6667±4.76859), $p < .010$ (See figure 4.7). The result generated in this study has also been confirmed by many of the existing studies in Second Language Acquisition (SLA) and Second Language Learning (SLL) (e.g. Bitchener, 2008; Ellis, 2008; Bitchener and Knoch, 2010b; Khodie & Sardari, 2015; Rezaei and Derakhshan, 2011; Rezazadeh et al., 2015). However, the performance of the DCF group was notably better than the CG group (there was a positive statistical significance of results for post-and delayed post-tests between DCF and CG groups' scores), but the overall scores remained much more conservative than the MCF group. The DCF group revealed higher performance than the CG group, which is substantiated by previous studies (e.g. Ellis et al., 2008; Daneshvar & Rahimi, 2014).

Nevertheless, with these findings, it can be said that **both direct corrective feedback and metalinguistic corrective feedback (with handout) contributed to the immediate learning gains** of MCF and DCF groups in using English tenses accurately

in their writing. While both feedback were effective, the learning gains of the MCF group were higher in both post and delayed post-test. This corresponded to the findings of Lalande (1982) and Ebadi (2014) where it was shown that error coded feedback like MCF increased the accuracy of the learners when compared to those who received direct corrective feedback. As Ferris (1997) also noted, indirect error coded feedback provides less grammatical errors in writing tasks, and that self-correction is beneficial in promoting accuracy of students in using correct usage of grammatical items in writing (Baleghizadeh & Dadashi, 2011).

The time factor has also a significant effect on the efficacy of the different feedback techniques, as scores for pre-tests were considered statistically insignificant, while the scores for post and delayed post-tests were considered statistically significant (See table 4.7). Based on this insight, it can be inferred that conducting follow- up tests with the time duration less than or equal to one day is not feasible, as there will be not much difference in performance revealed. On the other hand, when time factor in terms of weeks can provide a better insight into the overall uptake of linguistic learning for language learners. Since the results obtained by the three different groups for pre-test were insignificant, results from the post- and delayed post-test are crucial for understanding the effect of time on the overall scores of the different groups. For example, all the groups for different tests types were unable to maintain their scores between post- and delayed post-tests, as the time duration of two weeks reduced the overall language learning and its retention rate for these groups. However, inter-tests gaps of more than a week allowed for a better ability to determine the efficacy of different feedback techniques. For example, the inter-test time gap for study by Bitchener et al. (2005) was 4 weeks by measuring efficacy of different feedback techniques at four different time frames, while studies of Rassaei et al. (2012) and Shintani & Ellis (2013) measured long-term efficacy after two weeks.

4.5.2. Discussing the effects of corrective feedback on MKT and TGJT scores

The **second purpose** of the study was to draw a link between corrective feedback and the type of knowledge that the students engaged in when performing a task. Explicit knowledge, as defined in the first chapter, is a conscious knowledge responsible for learning. On the other hand, implicit knowledge is a procedural knowledge which mostly relies into intuition. Based on the findings of previous studies (e.g. Polio, 2012; Shintani & Ellis, 2013), there is a little information regarding the ways in which language learning is facilitated with the help of either implicit knowledge, explicit knowledge or a combination of both. However, existing studies have revealed that written corrective feedback techniques could have more impact on the explicit knowledge, rather than implicit knowledge (Polio, 2012; Bitchener, 2012). Therefore, it is very difficult to pin-point the interaction of language learning with these types of knowledge, specifically in the context of SLA. However, with prior assumptions regarding the interplay of implicit knowledge acquisition in (Timed Grammatical Judgment Test (TGJT) and explicit knowledge acquisition in Metalinguistic Knowledge Test (MKT), the overall results of the test scores can shed light on the implicit and explicit knowledge acquired during the tests conducted. With the use of DCF and MCF (with metalinguistic handout), both of the types knowledge in this study were measured using the MKT and TGJT assessments respectively.

Results revealed after performing the two-way mixed Anova revealed that the MCF group outclassed the DCF group and CG group in the two tests. Specifically, using pairwise comparison with Bonferroni adjustment, only the **MCF group's test scores in MKT were statistically significant different between pre-test, immediate post-test and delayed post-test**. The results in this study coincided with various study (e.g. Azizi, Behjat & Sorahi, 2014; Gholaminia et al., 2014; Diab, 2015) whose findings stated MCF

is more effective when compared to other forms of corrective feedback. Figure 4.12 below provides an overview of the scores of each group when explicit knowledge was measured.

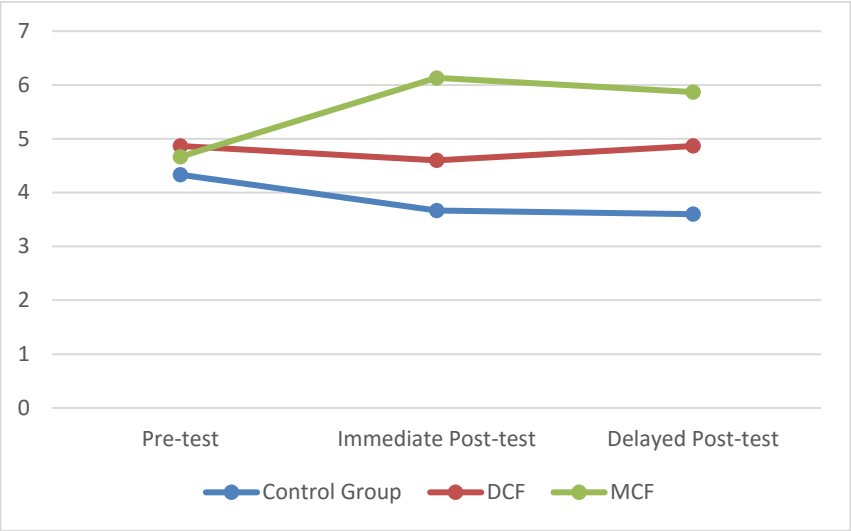


Figure 4.12: Mean scores of the metalinguistic knowledge test

In figure 4.11, MKT scores of the MCF Group revealed that there was an increase in the performance between pre-test and immediate post-test. This shows a significant difference of p=.011; however when pre-test was compared to delayed post-test, no significant difference was found, p=.069. This is because there is a slight decrease in the delayed post-test when compared with the immediate post-test, p= 1.000. On the other hand, the CG and DCF groups have not shown any significant differences in the respective testing time points.

On the other hand, the scores **in TGJT for the test of implicit knowledge**, using pairwise comparison with Bonferroni adjustment, **revealed no significant difference in the three testing time points of the DC group**, F (2,28)=1.596, p= .221. However, significant differences were found between the pre-test and immediate post-test of MCF group, p=.000, and between immediate and delayed post-test of Control Group, p=.009.

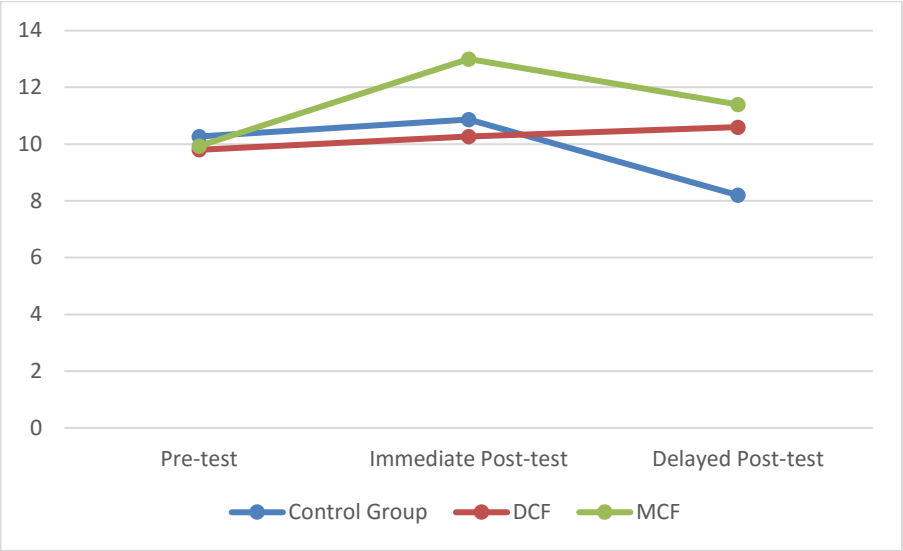


Figure 4.13: Mean scores of the Timed Grammatical Judgment test

Figure 4.12 shows that the significant difference found in the MCF Group was due to the increase in TGJT scores when compared to pre-test. While for CG group, it was due to the decrease in score in the delayed post-test. This suggests that the treatment received by MCF Group has a high significant effect during the immediate post-test, and only significant, not very high, when tested for the long term. It can be assumed that the treatment effect decreases as the time goes on.

Considering the scores of TGJT and MKT above, it is evident in the line graph that **MCF group showed superiority to all the other groups** in the immediate post-tests with a considerable margin. This is followed by the DCF group, and the CG group with least improvement. The findings suggest that metalinguistic corrective feedback (MCF) with metalinguistic handout is successful in enhancing explicit and implicit knowledge. However, its inconsistency to further improve the knowledge after three weeks suggests its decreasing effects. While DCF also showed positive immediate effects for implicit knowledge, statistical analyses revealed that the results obtained by the group are not statistically significant in nature.

Despite the performance of the MCF group, it was still impossible to observe a generalizable positive or linear positive trend of improvement for any of the three group

of learners. This phenomenon has been witnessed in earlier studies (Ellis, 1994; Lightbown and Spada, 1999; Bitchener et al., 2005).

4.5.2.1. The Timed Grammatical Judgment Test (TGJT) and the three groups

One interesting issue with regards to the TGJT was the lack of ability to retain improvements in performance as highlighted by the CG and MCF groups between post- and delayed post-test. On the other hand, the DCF group was consistent in their performance throughout the three tests. The lack of retention of linguistic knowledge by MCF group can be attributed to the short-term learning effects of the MCF, which were lost after a significant amount of time had passed. In addition, some of the improvements in language learning were not sustainable in nature. Researches have shown that the effects of implicit knowledge are long-term and more durable (Shintani and Ellis, 2013). The concern on improvements in implicit learning and knowledge of MCF group after taking has been confirmed (Ellis et al., 2009), as this group outperformed all the others in test scores between the MCF and other groups. However, the durability of the effects of implicit knowledge within the MCF group could not be confirmed in the present context. Another issue is regarding the level of influence of explicit and implicit knowledge within TGJT. This is because it has been generally used for testing improvements in implicit knowledge (Ellis et al., 2009), but studies (e.g. Bialystok, 1979, 1982) have shown that written TGJT is responsible for increase in explicit knowledge and oral TGJT is linked with increase in implicit knowledge (Bitchener, 2012; Polio, 2012; Godfroid et al., 2015). This raises the question on the level of involvement of implicit and explicit knowledge within the written TGJT, which brings to the forefront the importance of investigating the link between explicit and implicit knowledge. This can be formulated for future studies to focus on examining written error feedback techniques and the level of improvements on implicit and explicit knowledge (Polio, 2012). Therefore, basing only from the findings of the study, it can be concluded from the TGJT results that **metalinguistic**

corrective feedback has positive immediate effects to the implicit knowledge of the MCF group; however, despite they outperformed the other groups, there was significant reduction in test results between post- and delayed post-tests. Conversely, the **improvements in implicit knowledge for the DCF group were consistent, but statistical analyses have revealed that the results obtained by the DCF group are not statistically significant in nature**. This finding is consistent with Rezazadeh, Tavakoli, & Rasekh (2015) who measured implicit knowledge after using DCF in students' writing task. In their study, the DCF showed a positive impact to the immediate post-test of TGJT, however its significance when compared to the post-test result did not show any statistical significance. Thus, the efficacy of the DCF when it comes to long term efficacy remains questionable in this study supporting the claims of Rezazadeh et al., (2015) but opposes the findings of Bitchener & Knoch (2010) and Van Beuningen et al. (2008). Some studies have also presented the inefficacy of DCF in improving implicit and explicit knowledge (Shintani & Ellis, 2013). Therefore, it is uncertain how DCF can promote explicit knowledge when it comes to the durability of its effects.

4.5.2.2. The Metalinguistic Knowledge Test (MKT) and the three groups

While implicit test (TGJT) above was discussed, it is also necessary to provide equal discussion when it comes to the test for explicit knowledge. The MKT has been linked with improvements in explicit knowledge of language learners (Ellis et al., 2009). Therefore, the results obtained by this test in this study will shed light on the effects of different corrective feedback strategies on the overall explicit knowledge development in the context of SLA.

Within-subject effects revealed that the test scores of the CG and DCF groups were found to have no significance. The Bonferroni multiple comparisons showed that MCF and DCF groups as well as the MCF and CG groups' results for test scores of post- and delayed post-tests were statistically significant. Thus, **the results showed that MCF**

group showed highest performance on MKT, along with relatively consistent results for the post- and delayed post-tests (even though there was slight reduction in results for delayed post-tests). The performance of the DCF group was modest overall, with no improvements registered from pre-test to post-and delayed post-tests. However, there was no statistical significance of the results for the DCF group in general as well as the DCF and CG groups, even though the DCF group had higher performance than the CG group. This fact tackles on the efficacy of direct corrective feedback in terms of its contributions towards improving language learners' explicit and implicit knowledge (Truscott,1999, 2004, 2007, 2010). Furthermore, **the result provided by the MC group also never claims the durability of effects** as there was a slight reduction in the delayed post-test result which can be seen from figure 4.11. This outcome is similar to the test for implicit knowledge of the same group. Although the MCF group's score is relatively higher than the other groups, the immediate and delayed post-test did not reach statistical significance. The results imply that metalinguistic corrective feedback is effective in enhancing explicit knowledge, but its long term efficacy cannot be determined further in this study. Also, the MCF efficacy corresponded to the claims of prior studies that learners who are involved in treating errors perform much better than those who only rely to the answers provided by the teachers (Lalande, 1982; Sivajo, 2012; Ghandi & Maghsoudi, 2014; Hosseiny, 2014; Bitchener and Knoch, 2008; Ellis, 2009).

Nevertheless, the overall result of the statistical analysis for TGJT and MKT, as also discussed in the beginning (See section 4.3), recognizes the **positive immediate effects of metalinguistic corrective feedback in enhancing implicit and explicit knowledge of the learners**. Although the DCF showed a slight improvement towards the delayed post-test for TGJT, the result was still not significant. A similar outcome was found in their test for explicit knowledge. The DCF's overall result implies direct corrective feedback inefficacy in providing both short and long term effects. Furthermore,

the non-significant result between immediate and delayed post-tests for the MCF group determines the MCF and metalinguistic handout efficacy to be decreasing as time passes. The statistical result of this study could also be supported by how learners viewed the experimental study experienced. Such topic will be elaborated when students' view will be discussed below.

4.5.3. Discussing participants' views of the experienced experimental procedure

The **third purpose** of this study was to provide interesting details to why one corrective feedback performed better than the other. In aiming to understand the participant's views on the received feedback, interviews and series of thematic analyses were conducted. As discussed above, the CG group did not provide any significant effects to all given tests (MKT, TGJT, PNT). Ninety-three (93) percent of the participants in this group claimed that they did not understand when the teacher did not give any corrections to their paper. To them, learning has never taken place. **The ambiguity that CG group experienced played a major contributing factor to their performance**. Excerpts below strengthen this idea:

Student C10 -	*"No. I don't know. I don't know why it is wrong also."*
Student NC14- I'm	*"I cannot understand. Difficult to know. It's not telling me what [why] wrong"*

The vagueness of this process contributed to the confusion of the participants on what was supposed to be corrected on their paper. The identified problem of this process directly links to how participants in this group performed at their tasks. The quantitative result for this group has not been significant since no improvement was seen in any of the tests in different testing time. On the other hand, the differential effects of DCF and MCF were evident in the PNT immediate post-test. However, between effects showed MCF to be more effective than DCF (See section 4.3). The evidence of MCF superiority was also evident in the tests for implicit and explicit knowledge where MC group outperformed DCF and CG group (See section 4.23.

To discuss in detail, the themes generated within these two groups (DCF and MCF) speak of the efficacy of both. The direct and metalinguistic (with handout) corrective feedback, as perceived by DCF and MCF groups, have been helpful in reducing errors and becoming accurate in using basic English tenses in writing. **DCF group thought that direct corrective feedback helped them to identify, locate, and correct their errors as indicated in the themes generated during the interviews**. This finding is similar to the claims of Bitchener and Knoch (2010) that DCF allows learners to easily recognize incorrect language forms. This is also much better than memorizing error codes as these learners have limited knowledge of the target grammatical structures (Eslami, 2014; Spivey, 2014). Below in an excerpt on how DCF is effective to few participants:

Student DCF8- *"I receive[d] a lot of strong grammar. … Yes. It's helpful for me to know how…what my wrong [is my mistake]."*

However, although the DCF group has been successful, the statistical result speaks of the superiority of MCF group in terms of the learning gains of the learners. The DCF efficacy was also tested for its long-term effects; however, statistical result between post-test and delayed post-test is insignificant. This is opposite to the claims of Van Beuningen, De Jong & Kuiken (2012) who found DCF to have significant long term effects as compared to the other groups. As there are also a number of learners in DCF group who claimed the benefits of DCF in PNT, as expressed above, these learners also believed that it does not promote critical thinking. As cited by Ellis (2009), students who received DCF require minimal processing that could affect metalinguistic understanding. In particular, **some learners in the DCF group believed that DCF makes them lazy to think and that it makes the task too easy for them, thus requiring them not to think critically.** The excerpts below provide evidence that learners who received DCF were not encouraged to think critically.

Student DCF2 - *"Actually, it is helpful but I think It will be too easy like that. I will not think more why."*

Student DCF7 - *"Not good because the test…the test on avoid…avoid the other students to research their answer. It makes students lazy."*

Nevertheless, Swain (1985) also believes that learners need an opportunity to be involved in drawing out conclusions through active thinking processes for them to learn, which direct corrective feedback failed to do in this study. Ellis (2009) and Sivaji (2012) also supports the stance that effective acquisition occurs when learners are encouraged to notice their errors and actively participate in treating them. In comparison with MC Group, learners have achieved much learning gains as compared in DC group in terms of writing accuracy using the English tenses. **One of the major themes speak of MCF efficacy and its positive influence in promoting critical thinking**. Participants believed that it made them understand the error and the appropriate rules to apply. Also, the statistical result of MC group outperformed DC and NC groups in post-test and delayed post-test. When learners in this group were interviewed, all of them claimed that coding was helpful coupled with the metalinguistic handout. **The handout has been used to correct their errors with the accompanying codes which makes MCF more effective**.

Student MCF2- *"I can refer to the note that you give so that I know the…how to find the true answer."*

Student MCF6- *"For me, it is helpful because I…he can make me try to find the answer. It makes me interested to see the codes."*

With this, they are able to speed up their understanding of the target structure. As being said, the coding process makes the corrective feedback interesting to them. According to them, **the MCF and handout provided them an opportunity to work on their own and become independent learner**. Similarly, Ellis (2013) also said that learners who achieve metalinguistic understanding reflect on their answers and they try to correct their errors. Diab (2015) also claimed the deep internal processing involved when metalinguistic corrective feedback is provided to the learners.

Additionally, Faqeih (2015) claimed that learners preferred MCF because it offers a new method of correcting errors as compared to the traditional DCF as these learners "used to be instructed via traditional teaching methods in most language classes in Saudi Arabia" (p. 670). This is evident in the sample excerpts written above where **learners thought of the coding process in MCF to be new and interesting**. However, when codes are given, it is also a must to ensure that these learners familiarize themselves to the codes in order for this process to work (Faqeih, 2015). Fahim & Montaseri (2013) also claimed the efficacy of MCF in improving lexical source and grammatical accuracy of students, so as the study of Rassaei and Moinzadeh (2011) where MCF outperformed other forms of feedback during the immediate post-test. Nevertheless, both corrective feedback improved the accuracy of learners in using English tenses. This supports the claim that corrective feedback is effective (i.e. Ferris, 1999; Hyland, 2003, Ashwell, 2000; Ferris & Roberts, 2001), and opposes the views that corrective feedback is useless (i.e. Kepner, 1991; Truscott, 1996; Truscott and Hsu, 2008).

Schmidt's (2001) Noticing Hypothesis also speaks of awareness as an important factor in language acquisition. According to him, learners need to notice the errors and understand them in order to facilitate improvement. With this notion, the **MCF and metalinguistic handout can be assumed to possess the level of awareness necessary to identify and treat errors**. Swain and Lapkin (1985) also introduced awareness as a necessity to notice linguistic gaps, which consequently provides them an opportunity to fill those gaps by learning previously unknown language structure. Swain (1985) believes that with this process, learners are able to conduct mental grammatical processing that allows them to quickly learn and acquire knowledge pertaining to the target language. Nevertheless, the result of this study corresponded to the prior claims regarding the efficacy of metalinguistic corrective feedback (i.e. Diab, 2015; Gholaminia, I., Gholaminia, A. & Marzban, 2014; Kazemipour, 2014), and its effectiveness to encourage

learning independence through self-engagement (Holec, 1980). Moreover, this can also lead to further investigation regarding operationalizing metalinguistic corrective feedback in combination with coded metalinguistic handout. To date, no studies has been found to use the combination of these two.

4.6. Chapter Summary

This chapter has presented the result of the investigation that highlighted the relationship of corrective feedback to the written accuracy of the students as well as in the improvement of their explicit and implicit knowledge. The statistical result proved the overall efficacy of metalinguistic corrective feedback plus handout over the traditional direct corrective feedback across the three tests in different times. Such outcome is substantiated by the generated themes of the qualitative study. Learners who received metalinguistic corrective feedback enjoyed the benefits of the feedback, while at the same time were encouraged to make use of critical thinking when treating errors. This positive theme was not yield from the DCF group. Learners in this group viewed direct corrective feedback as helpful; however, it promotes rote learning and does not provoke critical thinking.

CHAPTER V

SUMMARY AND CONCLUSION, IMPLICATIONS AND RECOMMENDATIONS

5.1. Introduction

This chapter summarizes the key findings of the research study, its implications and limitations. Recommendations for further study has also been provided outlining future actions including validations of the current findings.

5.2. Summary and Conclusion

The quantitative non-equivalent experimental design of the study aimed to investigate the efficacy of corrective feedback in the writing accuracy of using Basic English Tenses of the ESL learners. It also aimed to identify corrective feedback's relationship to the improvement of the learner's implicit and explicit knowledge. Through the two-way mixed method ANOVA, metalinguistic corrective feedback (MCF) was found to be successful in improving the writing accuracy of the ESL learners with a positive impact to the development of implicit and explicit knowledge of the students. Direct corrective feedback (DCF) has also been successful in improving the writing accuracy; however, the increase in score of those who received MCF is much higher when compared to the DCF. The statistical result also showed that DCF failed to improve the implicit and explicit knowledge of the students; hence, this questions its efficacy in enhancing either type of knowledge.

The thematic analysis from the interview transcripts also supported and validated the statistical result. Students who received MCF claimed that the feedback helped them to think and be critical. On the other hand, DCF just allowed the students to easily locate errors and did not allow the students be an active participant in the treatment process. Thus, DCF students claimed that it did not promote critical thinking.

Both quantitative and qualitative findings also agree that MCF has been much more helpful in improving the performance of the students. Despite the claims that error coding can be difficult (Sheen, 2007), the study has supplemented the students with metalinguistic handout which made it easier for the students to understand the codes. In other words, error coding has no problem with the suitability as long as the symbols are introduced and explained to the students given an appropriate amount of time.

The result of this study is also consistent to the findings of previous studies (Diab, 2015; Eslami, 2014; Ferris & Helt, 2000; Ferris & Roberts, 2001) that affirmed the efficacy of corrective feedback and opposed to the findings (i.e. Kepner, 1991; Semke, 1984; Sheppard, 1992; Trsucott, 1996, 1999, 2004, 2007, 2010) that it creates a negative impact to language acquisition. It further implies that corrective feedback does not impede language acquisition, rather it enhances the process towards successful learning. Sukasame et. Al (2014) also believed that pointing errors is never a problem, rather a means to improve one's performance.

Nevertheless, this study enlightened the issues regarding the effects of corrective feedback to improving the accuracy of students in writing, and its impact to the explicit and implicit knowledge. This contributes to the body of research allowing a deeper reflection surrounding corrective feedback. However, while this study achieved its objectives, it never puts a stop to the on-going controversy and debate on various corrective feedback variables. More studies are required to explore the many unanswered questions.

5.3. Pedagogical Implications

Several implications were identified towards the completion of the study. It covers the appropriacy of feedback of teachers in the classroom and how can this feedback make learners participate in the error treatment process.

When giving corrective feedback, teachers need to identify the type of feedback suitable to their level. It is necessary that a feedback's complexity can be simplified by using another means such as a handout to let students cope with its difficulty. Sheen (2007) mentioned that metalinguistic corrective feedback is suitable to students with higher proficiency level; however, in this study, the researcher made use of the feedback and was coupled with a metalinguistic handout to allow low proficient learners grasp its complexity. A clear instruction is also helpful to create a general understanding of the tasks among the participants.

Focused corrective feedback was also a good strategy to limit the confusion that students face when correcting their errors. Ellis (2009) claimed that focused corrective feedback facilitates the better acquisition of specific language structures. While unfocused corrective feedback is also helpful, focused corrective feedback limits the amount of affective factors that Krashen (1985) believes to be harmful in learning a language. It lessens the anxiety that the students may feel when receiving error treatments. Krashen (1985) believes that when the anxiety level of students is high, learning does not take place. Receiving a lot of errors might trigger emotional and psychological disturbance to the students which could limit the amount of comprehensible input they take in. Thus, when teachers are correcting students' errors, emotional and psychological factors must be considered.

Following the result that the operationalized corrective feedback did not show a consistent score up to the delayed post-test, it does not merit the consideration that it serves as a positive reinforcement for the students to increase their performance beyond from their current level. The corrective feedback pushed the students to perform better from what they only know and what they believe they can only do. This only proves that the immediate provision of corrective feedback enhances the performance of the students

and creates a positive support to the understanding of the students regarding the target language.

Teachers also need to consider the proficiency level of the students when giving corrective feedback. In this study, it was assumed that the students had homogenous proficiency level based from the research site's grouping through their placement result. Diagnosing students' level helps teachers to design and choose the appropriate feedback to the students' oral or written work. Sheen (2007) also believed that choosing the appropriate feedback to the students' proficiency level would be much more effective and beneficial. Hence, teachers who opt to provide contemporary corrective feedback should have a general understanding of the type of feedback before its implementation.

Apart from the need of teachers to understand feedback, the findings of this study can also serve as a driving force for both teachers and researchers who claimed that feedback is useless to modify their belief and conduct an in-depth investigation over the course of time (Freeman, 1992).

Schmidt's (1995) Noticing and Swain's (1985) Output Hypothesis served as the theoretical framework of the study. The reflection towards students' involvement in the treatment process merits the consideration of how students provide attention and action to the kinds of error they have committed. Both hypotheses believe that it is necessary for the students to notice their errors so they can assess by themselves their current understanding of the target language. By becoming aware, they can reformulate their understanding through the feedback they receive from their peers and teachers. Once they achieve this, reflection follows which allows them to internalize their linguistic knowledge and achieve metalinguistic understanding.

While it can be said that self-involvement in the treatment process is evident in the metalinguistic corrective feedback, there is a tendency that the students who received direct corrective feedback would be teacher reliant and passive. In most classroom

settings, students tend to be passive and rely only on the teachers' comments and do nothing about it. As most of the students during the interview said that this kind of feedback does not promote critical thinking, it can be assumed that students do not desire to act further since the answer is already given. This can be changed by designing the kind of feedback that involves the students in correcting the errors. Since Schmidt's (1995) Noticing and Swain's (1985) Output Hypothesis mentioned the need of students' engagement, teachers can encourage the students to do self and peer evaluation. Metalinguistic corrective feedback was able to do this as most of the students claimed that they could check and correct their own mistakes using the feedback and handout. Nevertheless, teachers also need to be reflective of their practices. There is a need to constantly examine the methodologies and adopt new ideas for innovation. Only by then teachers can achieve a better understanding of their students' needs, and eventually improve the practices that benefit both teacher and students.

5.4. Limitation of the Study

Due to the established classes of the research site, random selection of participants in the groups was not permitted. Although students were considered to have homogenous level of proficiency, a random assignment would be ideal to improve the result of the study. Apart from that, all the participants were local Malay students with a limited number of fifteen per group. A large number per group and an extension to the different races in Malaysia would provide a more valid result which could also result to an insightful comparison of the differential effects of the feedback to the major races in Malaysia.

Furthermore, only the metalinguistic corrective feedback group was given a handout. Giving handouts to the other groups might give us a similar result from the metalinguistic corrective feedback group. The study also only focused on students with beginner level proficiency. This limits the generalizability of the results with in this

particular level; hence extending it to higher levels in another target language structure would allow us a new insight regarding the efficacy of the operationalized feedback.

Nevertheless, the findings revealed at the end of this research has only been confined to beginner level English learners for the specific sociocultural context of Malaysia. The study may not be generalized to students outside this age range, nationality, or proficiency level. Therefore, extrapolation of the findings to another country, region or educational level of learners cannot be possible. At the same time, evidence has shown that different types of feedback methods have fared differently under varying circumstances. Consequently, extension of results obtained to other types of feedback methods is also not advisable or feasible. Due to the cross-sectional design being employed, longitudinal studies may be necessary to test the treatment efficacy over time. Furthermore, the study result may be influenced by students' attrition during the testing and inaccurate self-reporting during the presentation. Since quasi-experimental non-equivalent control group design with convenience sampling procedure was used, the study was also limited to controlling the history, maturation, testing, instrumentation, selection and mortality. The study recognizes the interaction of selection and maturation as possible sources of internal invalidity, with interaction of testing and treatments as threats to external validity. In this manner, some of the limitations of the existing study is that it is confined to analysing improvements in Beginner English level students' basic tenses of the English language.

5.5. Suggestions for future research

While it is true that the efficacy of corrective has been debated by many researchers, practically started by Truscott (1996) and was argued by Ferris (1999) presenting own studies, it seems that the direction now heads to what kind of corrective feedback is most effective. It is an undeniable fact that the demand for the efficacy of methods, specifically corrective feedback, in SLA classroom has been increasingly popular to support and

validate past studies. However, it also undeniable that there is still a room to improve by just considering the variables in a study. This includes the variations we can do in providing the corrective feedback, the teaching styles, the attitude of the students towards the feedback, and the perception of the teachers when using feedback.

With regard to the proficiency levels of the students, a future research study may focus on identifying the relationship of the type of corrective feedback and how it benefits the level of proficiency of the students. Utilizing two experimental groups with different proficiency levels will be good to identify how a feedback can affect a group with low and high proficiency levels. Both groups should be given a similar treatment process, and the differences between and within each group should be analysed.

Further research can also consider studying how excessive feedback can affect the motivation of the learners. Regardless of the type of feedback, an excessive feedback may have a different effect to different learners. Otherwise, it can also be compared to how less a feedback should be. A comparison between these two may shed light to the amount of feedback should be given to the learners.

A longitudinal study with a large number of participants offers a potential to further explore the efficacy of direct and metalinguistic corrective feedback. This will provide a more valid result considering the length of time and its long-term effects when it comes to SLA. Aside from that, a study on students' feedback preference could also provide a better understanding of how to improve SLA. A student-driven corrective feedback may offer a more interesting result when compared to the commonly used corrective feedback in the classroom.

A more investigation must also be devoted to examining the tools to measure implicit and explicit knowledge of the learners. While it was claimed that there was no pure measures, it is still possible to design a test that could limit one's access to either type of

knowledge. A researcher may also look into the time factor and how it affects the use of explicit and implicit knowledge. An accurate measurement tool would help teachers and researchers design appropriate methods to improve better language acquisition process.

Nevertheless, one should note that a perfect research design is impossible to construct; however, even a slight variation in the process could make a difference.

REFERENCES

Abdul Karim , H, (2016). Unemployed because they can't speak English. *New York Strait Times*. Retrieved from https://www.nst.com.my/news/2016/02/125529/unemployed-because-they-cant-speak-english.

Abdul Karim, H. (2016). Unemployed because they cant's speak English. *New Straits Times*. Retrieved from https://owl.english.purdue.edu/owl/resource/560/10/.

Abdullah, A. T. H. (2013). Error analysis on the use of the simple tense and the simple past tense in writing essays among TESL college students. *International Journal of Education and Research, 1*(12).

Akakura, M. (2012). Evaluating the effectiveness of explicit instruction on implicit and explicit L2 knowledge. *Language Teaching Research, 16*, 1, 9–37.

Al Ajmi, A. A. S. (2015) The Effect of Written Corrective Feedback on Omani Students' Accuracy in the Use of English Prepositions. *Advances in Language and Literary Studies, 6,* 61-71.

Alderson, J.C., Steel, D. & Clapham, C. (1997). Metalinguistic knowledge, language aptitude and language proficiency. *Language Teaching Research, 2,* 93-121.

Aljaafreh, A., & Lantolf, J. P. (1994). Negative feedback as regulation and second language learning in the zone of proximal development. *The Modern Language Journal, 78(4),* 465-483.

Al-Jarrah, R. (2016). A suggested model of corrective feedback provision. *Ampersand, 3,* 98-107.

Allwood, J. (1992). *Feedback in second language acquisition*. Gothenburg papers in Theoretical Linguistics.

Ary, D., Jacobs, L. C., Sorensen, C., & Razavieh, A. (2009). *Introduction to Research in Education.* Wadsworth: Cengage Learning.

Ashwell, T. (2000). Patterns of teacher response to student writing in a multiple-draft composition classroom: Is content feedback followed by form feedback the best method? *Journal of Second Language Writing, 9*, 227–258.

Azizi, M., Behjat, F., & Sorahi, M. (2014). Effect of metalinguistic teacher corrective feedback on writing performance of Iranian EFL learners. *International Journal of Language and Linguistics, 2 (6),* 54-63.

Azrinda, M. H. (2013). "Indirect Corrective Feedback: A Tool to Students' Self-Editing of ESL Writing (Thesis). Faculty of Eduation, UTM.

Bakar, K. Hamid, N. Z., Mat Awal, N. & Jalaluddin, N. H. (2007). *First Language Influence on Second Language Performance: A Study of Common English Grammatical Errors among Rural Secondary School Students.* Faculty of Social Sciences and Humanities. Universiti Kebangsaan Malaysia.

Baleghizadeh, S., & Dadashi, M. (2011).The Effect of Direct and Indirect Corrective Feedback on Students'SpellingErrors. *PROFILE, 13(1),* 129-137.

Barriball, K. and While, A. (1994). Collecting data using a semi-structured interview: a discussion paper. *Journal of Advanced Nursing, 19 (2),* 328-335.

Bitchener, J. (2012). A reflection on 'the language learning potential' of written CF. *Journal of Second Language Writing,* 21(4), 348-363.

Bitchener, J., & Knoch, U. (2009). The relative effectiveness of different types of direct written corrective feedback. *System, 37(2),* 322-329.

Bitchener, J., & Knoch, U. (2010). Raising the linguistic accuracy level of advanced L2 writers with written corrective feedback. *Journal of Second Language Writing, 19*(4), 207–217.

Bitchener, J., Young, S. Cameron, D. (2005). The effect of different types of corrective feedback on ESL student writing. *Journal of Second Language Writing, 12,* 267-296.

Blumberg, B., Cooper, D.R. & Schindler, P.S. (2005). *Business Research Method.* Maidenhead: McGraw Hill Education.

Bowles, M. (2011). Measuring implicit and explicit linguistic knowledge: What can heritage language learners contribute*? Studies in Second Language Acquisition, 33,* 247-271.

Braun, V. & Clarke, V. (2006) Using thematic analysis in psychology. *Qualitative Research in Psychology, 3,* 77-101.

Brown, H. D. (1994). *Teaching by Principles: Interactive language teaching methodology*. New York: Prentice Hall Regents.

Brown, H. D. (2004). Principles of Language Learning and Teaching. Englewood Cliffs: Prentice-Hall.

Bryman, A. & Bell, E. (2007). Research Designs. In: *Business Research Methods*. New York. Oxford University Press.

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

Chomsky, N. (1965). *Aspects of the Theory of Syntax.* Cambridge, Mass.: MIT Press.

Chou, T.-L. & Chan-Lin, L.-J. (2015). Autonomous ESL Learning: "Read & Reflect in English". In *Proceedings of 6th World Conference on Educational Sciences*, 191, 357-360.

Cohen, D. & Crabtree, B. (2006). *Qualitative research guidelines project.* Robert Wood Johnson Foundation. Retrieved November 1, 2011, from http://www.qualres.org/index.html.

Collins, L. (2007). L1 Differences and L2 similarities: teaching verb tenses in English. *ELT Journal, 61(4),* 295-303.

Common European Framework of Referenc. (2017). In *Erasmus University Rotterdam*. Retrieved from https://www.eur.nl/english/ltc/cefr_levels/.

Control group. (2015). In *Dictionary*. Retrieved from http://www.dictionary.com/browse/control-group?s=t

Corder, S. P. (1967).The significance of learners' errors. *IRAL*, 5, 161-170.

Cowan, R. (2008). *The Teacher's Grammar of English*. Cambridge: Cambridge University Press.

Creswell, J. (2013). *Educational Research: Planning, Conducting, and Evaluating Quantitative and Qualitative Research* (4th ed.). Boston, MA: Pearson Education, Inc.

Daneshvar, E., & Rahimi, A. (2014). Written corrective feedback and teaching grammar. *Procedia, 136*, 217-221.

De Vaus, D. (2001): *Research Design in Social Research.* London: Sage.

DeKeyser, R. M. (2003). Implicit and explicit learning. In C. J. Doughty & M. H. Long (Eds), *The Handbook of Second Language Acquisition* (pp. 313-348). Oxford: Blackwell.

Diab, N. M. (2015). Effectiveness of written corrective feedback: Does type of error and type of correction matter? *Assessing Writing*, *24*, 16-34.

Dienes Z., Perner, J. (1999). A theory of implicit and explicit knowledge. *Behavioral and Brain Sciences, 22,* 735–808.

Doughty, C. (2001). Cognitive underpinnings of focus on form. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 206-257). Cambridge: Cambridge University Press.

Doughty, C. (2003). Instructed SLA: Constraints, compensation, and enhancement. In C. Doughty & M. H. Long (Eds.), *The handbook of second language acquisition* (pp. 256-310). Oxford, UK: Blackwell Publishing.

Ebadi, E. (2014). The Effect of Focused Meta-linguistic Written Corrective Feedback on Iranian Intermediate EFL Learners' Essay Writing Ability. *Journal of Language Teaching and Research, 5*(4), 878-883.

Egi, T. (2010). Uptkae, modified output and learner perceptions of recasts: learner responses as language awareness. *The Modern Language Studies, 6*, 85-92.

El Tatawy, M. (2002). Corrective feedback in second language acquisition. *Working papers in TESOL and Applied Linguistics, 2(2),* 1-19.

Ellis, R. (1990*). Instructed Second Language Acquisition*. UK: Basil Black Well.

Ellis, R. (1991). Grammaticality judgments and second language acquisition. *Studies in Second Language Acquisition, 13*(2), 161-186.

Ellis, R. (1994). *The study of second language acquisition*. Oxford: Oxford University Press.

Ellis, R. (2004). The definition and measurement of explicit knowledge. *Language Learning, 54,* 227-275.

Ellis, R. (2005). Measuring implicit and explicit knowledge of a second language. *Studies in Second Language Acquisition, 27*(2), 141-172.

Ellis, R. (2009). A typology of written corrective feedback types. *ELT journal, 63(2),* 97-107.

Ellis, R. (2012). *Language teaching research and language pedagogy.* Oxford: Wiley Blackwell.

Ellis, R. (2012). *Language teaching research and language pedagogy*. Oxford: Wiley Blackwell.

Ellis, R., & Loewen, S. (2007). Confirming the operational definitions of explicit and implicit knowledge in Ellis (2005): Responding to Isemonger. *Studies in Second Language Acquisition, 29,* 119-126.

Ellis, R., Loewen, S., & Erlam, R. (2006). Implicit and explicit corrective feedback and the acquisition of L2 grammar. *Studies in Second Language Acquisition, 28,* 339-368.

Ellis, R., Loewen, S., Elder, C., Erlam, R., Philp, J. & Reinders, H. (2009). *Implicit and explicit knowledge in a second language learning, testing and teaching*. Clevedon: Multilingual Matters.

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

Eslami, E. (2014). The Effects of Direct and Indirect Corrective Feedback Techniques on EFL Students' Writing. *Procedia-Social and Behavioral Sciences, 98,* 445-452.

Fahim, M., &Montazeri, M. (2013).The impact of metalinguistic corrective feedback on EFL learners' levels of lexical resource and grammatical range and accuracy in their oral proficiency. *International Research Journal of Applied and Basic Sciences, 4(7*), 1776-1782.

Faqeih, H. (2015). Learners' Attitudes Towards Corrective Feedback. *Procedia - Social and Behavioral Sciences, 192,* 664 – 671.

Farid, S., & Samad, A. A. (2012). Effects of different kind of direct feedback on students' writing. *Procedia-Social and Behavioral Sciences, 66,* 232-239. doi:10.1016/j.sbspro.2012.11.265.

Fazilatfar, A. M., Fallah, N., Hamavandi, M., & Rostamian, M. (2014). The effect of unfocused written corrective feedback on syntactic and lexical complexity of L2 writing. *Procedia - Social and Behavioral Sciences, 98,* 482–488. doi:10.1016/j.sbspro.2014.03.443.

Fazio, L. (2001). The effect of corrections and commentaries on the journal writing accuracy of minority- andmajority-language students. *Journal of Second LanguageWriting, 10,* 235-249.

Ferris, D. (1999). The case for grammar correction in L2 writing classes: A response to Truscott (1996). *Journal of second language writing, 8*(1), 1-11.

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

Ferris, D. R. (2011). *Treatment of error in second language student writing* (2nd ed.). Ann Arbor, MI: University of Michigan Press.

Ferris, D. R., & Hedgcock, J. S. (1998). *Teaching ESL composition: Purpose, process and practice*. Mahwah, NJ: Lawrence Erlbaum.

Ferris, D. R., & Helt, M. (2000). Was Truscott right? New evidence on the effects of error correction in L2 writing classes. *Paper presented at the American Association of Applied Linguistics Conference*, March 11–14, 2000, Vancouver, BC.

Ferris, D., & 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.

Frear, D. (2012). The effect of focused and unfocused direct written corrective feedback on a new piece of writing . *Issues and Trends 3*, 57-72.

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

Gass, S. (2003). Input and interaction. In: C. Doughty & M. H. Hong (eds.). *The handbook of second language acquisition*. Oxford, UK: Basil Blackwell.

Gass, S. M., & Varonis, E. M. (1994). Input, interaction, and second language production. *Studies in Second Language Acquisition, 16,* 283–302.

Ghabool, N., Mariadass, M. E. & Kashef, S. H. (2012). Investigating Malaysian ESL students' writing problems on conventions, punctuation, and language use at secondary school level. *Journal of Studies in Education, 2,* 3, 130-143.

Ghandi, M., & Maghsoudi, M. (2014). The Effect of Direct and Indirect Corrective Feedback on Iranian EFL Learners' Spelling Errors. *English Language Teaching, 7(8),* 53. doi: 10.5539/elt.v7n8p53

Gholaminia, I., Gholaminia, A., & Marzban, A. (2014). An Investigation of Meta-linguistic Corrective Feedback in Writing Performance. *Procedia - Social And Behavioral Sciences, 116*, 316-320.

Gitsaki, C. & Althobaiti, N. (2010). ESL Teachers' Use of Corrective Feedback and Its Effect on Learners' Uptake. *The Journal of Asia TEFL,* 7(1), 197-219.

Godfroid, A., Loewen, S., Jung, S., Park, J., Gass, S. Ellis, R. (2015). Timed and Untimed Grammaticality Judgments Measure Distinct Types of Knowledge: Evidence from Eye-Movement Patterns. *Studies in Second Language Acquisition*, 37 (2).

Groniond, N.E. (1982). *Constructing achievement tests*. Englewood Cliffs, NJ: Prentice-Hall.

Gutiérrez, K. G. C., Puello, M. N. & Galvis, L. A. P. (2015). Using pictures series technique to enhance narrative writing among ninth grade students at Institución Educativa Simón Araujo. *English Language Teaching 8(5*). http://dx.doi.org/ 10.5539/elt.v8n5p45. http://dx.doi.org/10.5539/elt.v8n5p45

Han, Y. & Hyland, F. (2015). Exploring learner engagement with written corrective feedback in a Chinese tertiary EFL classroom. *Journal Of Second Language Writing, 30,* 31-44.

Harmer. J. (2001).The Practice of English language Teaching.3rd Edition. *Longman Handbooks for Language Teachers.* Addison Wesley Publishing Company.

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

Hijjo, N. F. M. (2013). A morphosyntactic analysis on Malaysian secondary school students' essay writing in English class. *International Journal of Humanities and Social Science 3, 11*, 286-291.

Holec, H. (1980). Autonomy and foreign language learning. *Nancy: Centre de Recherches et d"Applications Pedagogiques en Langues*. Council of Europe.

Hosseiny, M. (2014). The Role of Direct and Indirect Written Corrective Feedback in Improving Iranian EFL Students' Writing Skill. *Procedia-Social and Behavioral Sciences, 98*, 668-674.

Hulstijn, J. (2003). Incidental and intentional learning. In C. J. Doughty & M. H. Long (Eds*) The Handbook of Second Language Acquisition* (pp. 349-381). Oxford: Blackwell.

Hulstijn, J. H., & Ellis, R. (2005). Implicit and explicit second-language learning. *Thematic issue of Studies in Second Language Acquisition, 27(*2).

Hyland, F. (2003). Focusing on form: student engagement with teacher feedback. *System 31,2*, 217–230.

Ivankova, N. V., Creswell, J. W., & Stick, S. L. (2006). Using mixed methods sequential explanatory design: From theory to practice. *Field Methods, 18(1*), 3-20.

Izumi, S. (2002). "Output, input enhancement, and the noticing hypothesis. An experimental study on ESL relativization". *SSLA, 4*, 541–577.

Jimenez, J. (2013*). Corrective Feedback*. Corso di perfexionamento per 1'insegnamento di discipline non linguistiche (DNL) in lingua staniera secondo la metodologia CLIL.

Kassim, A., & Ng, L. L. (2014). Investigating the efficacy of focused and unfocused corrective feedback on the accurate use of prepositions in written work. *English Language Teaching, 7*, 119-130.

Kazemipour, S. (2014). Comparing the outcomes of two types of corrective feedback on EFL classes' final exam. *Procedia - Social and Behavioral Sciences, 98,* 876–881. doi:10.1016/j.sbspro.2014.03.495.

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

Khodie & Sardari (2015). The Effect of Metalinguistic Corrective Feedback on Students' Writing Performance. *International Journal of Educational Investigations, 2 (4).*

Kim, H. & Mathes, G. (2001). Explicit vs. implicit corrective feedback. *The Korea TESOL Journal, 4,* 1-15.

Kirkgöz, Y. (2010). An analysis of Written Errors of Turkish Adult Learners of English. World Conference on Educational Sciences February, *Procedia Social and Behavioral Sciences, 2,* 4352-4358.

Kothari, C. R. (2004). *Research Methodology: Methods and Techniques*. India: New Age International.

Kowal, M. and Swain, M. (1997) From semantic to syntactic processing: How can we promote metalinguistic awareness in the French immersion classroom? In R.K. Johnson and M. Swain (eds) *Immersion Education: International Perspectives* (pp. 284-309). Cambridge: Cambridge University Press.

Krashen, S. (1985). *The Input Hypothesis: Issues and Implications*. London: Longman.

Krashen, S. 1981. *Second language acquisition and second language learning*. Oxford: Pergamon Press.

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

Lee, E. J. (2013). Corrective feedback preferences and learner repair among advanced ESL students. *System, 41,* 217-230.

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

Lewicki, P. (1986a). *Nonconscious social information processing*. San Diego, CA: Academic Press.

Li, S. F. (2010). The effectiveness of corrective feedback in SLA: A meta-analysis. *Language Learning*, 60, p. 309-365.

Lightbown, P. M., and Spada, N. (2013). *How Languages Are Learned.* (4th ed.) Oxford: Oxford University Press.

Loewen, S. & Philp, J. (2006). Recasts in the adult English L2 classroom: Characteristics, explicitness and effectiveness. *Modern Language Journal*, 90, 536-556.

Loewen, S. (2004). Uptake in incidental focus on form in meaning-focused ESL lessons. *Language Learning*, 54, 153-188.

Loftie M., Salleh, N. & Kadir, M. (2015). Linguistic Knowledge Types and Past-Time Inflectional Output of Malay Users of English. *Asiatic, 15* (9).

Long, M. (1980). *Input, Interaction and Second Language Acquisition* (Doctoral Dissertation). University of California, Los Angeles.

Long, M. (1983). Native speaker/non-native-speaker conversation and the negotiation of comprehensible input. *Applied Linguistics*, *4*, 126-141.

Long, M. (1996). The role of the linguistic environment in second language acquisition. In W. Ritchie & T. Bhatia (eds.). *Handbook of research on second language acquisition*. New York, NY: Academic Press.

Long, M., & Robinson, P. (1998). Focus on form: Theory, research and practice. In C. Doughty, & J. Williams (Eds.), *Focus on form in classroom second language acquisition* (pp. 15–41). Cambridge, England: Cambridge University Press.

Lyster, R. & Ranta, L. (1997). Corrective feedback and learner uptake. *Studies in Second Language Acquisition*, 20, 37-66.

Lyster, R. & Saito, K. (2010). Oral feedback in classroom SLA. *Studies in Second Language Acquisition*, 32, p. 265-302.

Lyster, R. (1994). The effect of functional-analytic teaching on aspects of French immersion students‘ sociolinguistic competence. *Applied Linguistics*, 15, 263–287.

Lyster, R. (1998a). Negotiation of form, recasts, and explicit correction in relation to error types and learner repair in immersion classrooms. *Language Learning*, 48, p. 183-218.

Lyster, R. (1998b). Recasts, repetition, and ambiguity in L2 classroom discourse. *Studies in Second Language Acquisition*, 20, 51-81.

Lyster, R. (2004). Different effects of prompts and effects in form-focused instruction. *Studies in Second Language Acquisition, 26,* 399-432.

Mackey, A. & Oliver, R. (2002). Interactional feedback and children’s L2 development. *System, 30,* 459–477.

Mackey, A. (1999). Input, interaction and second language development. *Studies in Second Language Acquisition*, 21, p. 557-587.

Mackey, A., & Goo, J. (2007). Interactional research in SLA: A meta-analysis and research synthesis. In: A. Mackey (ed.). *Conversational interaction in second language acquisition: A collection of empirical studies*. Oxford: Oxford University Press.

Mackey, A., Gass, S. & McDonough, K. (2000). How do learners perceive international feedback? *Studies in Second Language Acquisition, 22,* 471-497.

Mackey, A., Oliver, R. & Leeman, J. (2003). Interactional input and the incorporation of feedback: an exploration of ns-nns and nns-nns adult and child dyads. *Language Learning*, 53, 35-66.

Manchon, R. M. (2011a). Writing to learn the language: Issues in theory and research. In R. M. Mancho´n (Ed.), *Learning-to-write and writing-tolearn in an additional language* (pp. 61–82). Amsterdam: John Benjamins.

Manokaran, Ramalingam, & Adriana .(2013). A Corpus-Based Study on the Use of Past Tense Auxiliary ‘Be’ in Argumentative Essays of Malaysian ESL Learners. *English Language Teaching, 6 (10),* 111-119.

Mariko, A. (2007). Grammatical errors across proficiency levels in L2 spoken and written English. *The economic journal of Takasaki city, 49*, 117-129.

Maros, M., Tan, K. H., & Khazriyati, S. (2007). Interference in learning English: Grammatical errors in English essay writing among rural Malay secondary school students in Malaysia. *Journal e-Bangi, 2*(2), 1-15.

McDonough, K. (2005). Identifying the impact of negative feedback and learners' response on ESL question development. *Studies in Second Language Acquisition, 27*, 27-103.

McLeod, S. A. (2013). What is Validity?. Retrieved from www.simplypsychology.org/validity.html

Motlagh, L. N. (2015). Irinan EFL Teachers' Preferences For Corrective Feedback Types, implicit Vs Explicit. *Procedia - Social and Behavioral Sciences Procedia - Social and Behavioral Sciences, 192*(2), 364-370.

Mourssi, A. M. (2013). The efficacy of error analysis on second language learners' written accuracy: an empirical study in the context of Arab learners of English. *Educational Research, 4*, 249-256.

Muncie, J. (2000). Using written teacher feedback in EFL composition classes. *ELT Journal, 54(1)*, 47-53.

Musa, N., Lie, K., & Azman, H. (2012). Exploring English Language Learning and Teaching In Malaysia. *Journal Of Language Studies, 12*(1).

Myers, J. (1997). "Second language writing and research: The writing process and error analysis in student texts". *Teaching English as a Second or Foreign Language,6 (2*), 361-386.

Naidu, S. (2015). English a failing subject for Malaysia despite national push. Channel Nes Asia. Retrieved from http://www.channelnewsasia.com/news/asiapacific/english-a-failing-subject/2277498.html

Newton, N. (2010*). The use of semi-structured interviews. Exploring Qualitative Methods.* Retrieved from https://www.academia.edu/1561689/The_use_of_semi-structured_interviews_in_qualitative_research_strengths_and_weaknesses.

Nezakat-Alhossaini, Youhanaee, M. & Moinzadeh, A. (2014). Impact of explicit instruction on EFL learners' implicit and explicit knowledge: A case of English relative clauses. *Journal of Language and Linguistic Studies, 10(2),* 183-199.

Noor, F. (1985). A study of verb tense problems found in the writings of Malay speakers. *Retrospective Theses and Dissertations*. IOWA State University.

Norris, J., & Ortega, L. (2000). Effectiveness of L2 instruction: A research synthesis and quantitative meta-analysis. *Language Learning, 50*, 417-528.

Oakley, A. (1981). Interviewing women: A contradiction in terms? In Roberts H. (ed.) *Doing Feminist Research.* London: Routledge and Kegan Paul.

Oliver, R. & Mackey, A. (2003). Interactional context and feedback in child ESL classroom. *Modern Language Journal*, *87*, 519-533.

Paradis J. (2004). The relevance of specific language impairment in understanding the role of transfer in second language acquisition *Applied Psycholinguistics*. 25: 67-82. DOI: 10.1017/S0142716404001043

Paradis J., Nicoladis E, Crago M, Genesee F (2011). Bilingual children's acquisition of the past tense: a usage-based approach. *Journal of Child Language*. 38: 554-78. PMID 20738891 DOI: 10.1017/S0305000910000218

Patton, M. Q. (2004). *Qualitative research & evaluation methods*. Thousand Oaks, Calif. [u.a.]: Sage.

Pica, T. (1987). Second language acquisition, social interaction and the classroom. *Applied Linguistics*, 8, 3-21.

Plessner, H., Betsch, C., & Betsch, T. (2008*). Intuition in judgement and decision making.* New York: Lawrence Erlbaum.

Polanyi, M. (1966). *The Tacit Dimension.* Doubleday, Garden City, NY.

Polio, C. (2012). The relevance of second language acquisition theory to the written error correction debate. *Journal of second language writing, 21*(4), 375–389.

Rahman, M. S. & Ali, M. (2015). Problems in Mastering English Tense and Aspect and the Role of the Practitioners. *Journal Of Humanities And Social Science, 20,* 131-135.

Raimes, A. (1991). Errors: Window into the mind. *College ESL, 1(2),* 55-64.

Rassaei, E. & Moinzadeh, A. (2011). Investigating the effects of Three Types of Corrective Feedback on the Acquisition of English wh-questions forms by Iranian EFL Learners. *English language teaching, 4 (2*), 97-106.

Rassaei, E., Moinzadeh, A., & Youhannaee, M. (2012). Effects of Recasts and Metalinguistic Corrective Feedback on the Acquisition of Implicit and Explicit L2 Knowledge. *The Journal of Language Teaching and Learning, 2*(1), 59-75.

Reber, A. S. (1967). Implicit learning of artificial grammars. *Journal of Verbal Learning and Verbal Behavior, 6,* 855–863.

Reber, A. S. (1993). *Implicit learning and tacit knowledge: An essay on the cognitive unconscious*. Oxford: Oxford University Press.

Rebuschat, P. & Williams, J. N. (2012). Implicit and explicit knowledge in second language acquisition. *Applied Psycholinguistics, 33(4),* 829-856.

Rebuschat, P. (2013) Measuring implicit and explicit knowledge in second language research. *Language Learning,* 298-302.

Rezaei, S., & Derakhshan, A. (2011). Investigating Recast and Metalinguistic Feedback in Task-based Grammar Instruction. *JLTR Journal of Language Teaching and Research, 2*(3), 655.

Rezazadeh, M., Tavakoli, M., & Eslami Rasekh, A. (2015). The Effects of direct corrective feedback and metalinguistic explanation on EFL learners' implicit and explicit knowledge of English definite and indefinite articles. *Journal of English Language Teaching and Learning, 16*, 113-146.

Rifʿah, F. N. (2012). *An Obligatory Occasion Analysis of the Use of Past Regular and Past Irregular Forms in Recount Texts Produced by VIII Grade Student at SMPN 1 Bojonegoro. Thesis. English Department*. Faculty of Letters. State University of Malang.

Rohollahzadeh Ebadi, M., Mohd Saad, M. R., & Abedalaziz, N. (2014a). Corrective Feedback and Second Language Acquisition: Differential Contributions of Implicit and Explicit Knowledge. *The Malaysian Online Journal of Educational Science, 2(2)*, 10-19.

Rohrmeier, M., Rebuschat, P., & Cross, I. (2011). Incidental and online learning of melodic structure. *Consciousness and Cognition, 20*, 214–222.

Rouhi, A. and Samiei, M. (2010). The Effects of Focused and Unfocused indirect Written Corrective Feedback on Grammatical Accuracy of Iranian EFL Learner. *Social Science, 5(6)*, 481-485.

Russel, J. & Spada, N. (2006). The effectiveness of corrective feedback for second language acquisition: A meta-analysis of the research. In: J. Norris & L. Ortega (eds.). *Synthesizing research on language learning and teaching*. Amsterdam: Benjamins.

Salima, R. (2012). Measure of eliminating EFL students' errors in writing. *Procedia*, 318-327. http://dx.doi.org/10.1016/j.sbspro.2012.11.416

Sanz, C. (2003). Computer delivered implicit vs. explicit feedback in processing instruction. In: B. VanPatten (ed.). *Processing instruction: Theory, research and commentary*. Mahwah, NJ: Erlbaum.

Sanz, C., & Leow, R. P. (2011) (Eds.). Implicit and explicit language learning: *Conditions, processes, and knowledge in SLA and Bilingualism*. Washington, DC: Georgetown University Press.

Sardareh, S. A., & Saad, M. R. M. (2012). A Sociocultural Perspective on Assessment for Learning: A Case of a Malaysian Primary School ESL Context. *Procedia-Social and Behavioural Sciences*, 66, 343-353.

Saunders, M. P., Lewis & Thornhill. (2007). *Research Methods for Business Students*. Harlow, England: Pearson.

Schacter, D. L. (1987). Implicit memory: History and current status. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 13*, 501-518

Schmidt, R. (2001). Attention. In P. Robinson, *Cognition and Second Language Instruction* (1st ed., pp. 3-32). Cambridge: Cambridge University Press.

Schmidt, R. W. (1995). Consciousness and foreign language learning: A tutorial on the role of attention and awareness in learning. In R. Schmidt (Ed.), *Attention and awareness in foreign language learning* (pp. 1-63). Honolulu: University of Hawaii Press.

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

Sheen, Y. (2004). Corrective feedback and learner uptake in communicative classrooms across instructional settings. *Language Teaching Research, 8*, 263-300.

Sheen, Y. (2007). The Effect of Focused Written Corrective Feedback and Language Aptitude on ESL Learners' Acquisition of Articles. *TESOL Quarterly, 41*(2), 255-283.

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

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

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, 22*, 286-306.

Sivaji, K. (2012). The Effect of Direct and Indirect Correction Feedback on the Grammatical Accuracy of ESL Writing of Undergraduates. *Journal of Humanities & Social Sciences*, 8, 78-94.

Soleimani, E., Ismail, K. & Mustaffa, R. (2014). The Acceptance of Mobile Assisted Language Learning (MALL) among Post Graduate ESL Students in UKM. In *Proceedings of International Conference on Knowledge-Innovation-Excellence: Synergy in Language Research and Practice*, 118, 457-462.

Spivey, K. (2014). Written corrective feedback in ELS : Strategies, approaches, influences and factors. *Theses and Dissertations*. http://utdr.utoledo.edu/theses-dissertations/1703/

Stapa, S. H. & Izahar, M.M. (2010). Analysis of errors in subject-verb agreement among Malaysian ESL learners. *3L The Southeast Asian Journal of English Language Studies*,16(1).

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*, *2*, 262-282.

Sukasame, N., Kantho, S., & Narrot, P. (2014). A Study of Errors in Learning English Grammatical Structures on Tenses of MatthayomSuksa 4 Students of the Demonstration School, KhonKaen University. *Procedia - Social and Behavioral Sciences, 116*(21), 1934-1939. doi: 10.1016/j.sbspro.2014.01.498

Swain, M. (1985) Communicative competence: Some roles of comprehensible input and comprehensible output in its development. In S. Gass and C. Madden (eds) *Input in Second Language Acquisition*. Rowley, Mass.: Newbury House.

Swain, M. (2005). The output hypothesis: theory and research. In E. Hinkel (Ed.) *Handbook of research in second language learning and teaching (pp. 471-484)*. New Jersey: Lawrence Erlbaum Associates.

Swain, M. and Lapkin, S. (1995) Problems in output and the cognitive processes they generate: A step towards second language learning. *Applied Linguistics*, 16: 371-91.

Tanaka, Y. (1991). Comprehension and L2 acquisition: *The role of interaction. Unpublished paper*. Tokyo: Temple University Japan.

Ting, W. & Lin, J. (2015). Studies on Written Corrective Feedback: Theoretical Perspectives, Empirical Evidence, and Future Directions. *English Language Teaching, 8,* 110-120.

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

Truscott, J. (1996). The case against grammar correction in L2 writing classes. *Language Learning, 46*(02), 327-369.

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 for grammar correction in L2 writing classes”: A response to Ferris. *Journal of Second Language Writing, 8,* 111-122.

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

Truscott, J. (2007). The effect of error correction on learners' ability to write accurately. *Journal of Second Language Learning*, 16(4), 255-272.

Truscott, J. (2010). Some thoughts on Anthony Bruton's critique of the correction debate. *System, 38(2*), 329-335.

Uchiyama, K. (2006). *English Verb Tenses: An Informal but Extensive Reference for ESL Students, the Good Folks Who Teach Them, the Idly Curious*. Retrieved from https://docs.google.com/file/d/0B-e6qHPbxSdNSEZ5NXRYNmpGdmc/edit

Unin, N. & Bearing, P. (2016). Brainstorming as a Way to Approach Student-centered Learning in the ESL Classroom. In *Proceedings of 6th International Research Symposium in Service Management, 224*, 605-612.

Van Beuningen, C., De Jong, N., & Kuiken, F. (2012). Evidence on the effectiveness of comprehensive error correction in second language writing. *Language Learning, 62 (1),* 1-41.

Varonis, E. & Gass, S. (1985). Non-native/non-native conversations: a model for negotiation of meaning. *Applied Linguistics*, *6*, 71-90.

Velayutham, N. M. (2013). *The Effectiveness of Controlled Written Metalinguistic Feedback in Improving Writing*. Universiti Teknologi Malaysia.

Venkatagiri, H.S. & Levis, J.M. (2009). Phonological awareness and speech comprehensibility: *An exploratory study. Language awareness,* 16(4), 263-277.

Ward, A. W., & Murray-Ward, M. (1999). Assessment in the classroom. London: Wadsworth.

Wells and Wollack, 2003 Wells, C.S., & Wollack, J.A. (2003). *An Instructor's Guide to Understanding Test Reliability, Testing and evaluation services*. Madison: university of Wisconsin.

Williams, J. N. (2009). Implicit learning in second language acquisition. In William C. Ritchie & Tej K. Bhatia (Eds.), *The New Handbook of Second Language Acquisition*, 319-353. Bingley: Emerald Press.

Zainal, Z. (1990). *Contrastive analysis: the problems of L1 interference faced by UTM students when learning English. English Language Acquisition*, 3, 40-49. UTM.

Zamel, V. (1983). The composition processes of advanced ESL students: Six case Studies. *TESOL Quarterly, 17(2*), 165-178.

Zhang, L. & Rahimi, M. (2014). EFL learners' anxiety level and their beliefs about corrective feedback in oral communication classes. *System*, *42*, 429-439. http://dx.doi.org/10.1016/j.system.2014.01.012

Zhang, R. (2015): Measuring university-level L2 learners' implicit and explicit linguistic knowledge. Studies in Second Language Acquisition 37, 457-486.

Zohrani, K. & Ehsani, F. (2014).The role of ımplicit & explicit corrective feedback in persianspeaking efl learners' awareness of and accuracy in english grammar. *Procedia - Social and Behavipral Sciences, 98*,2018-2024.

APPENDICES

Appendix A: Permission and Consultation Letters

UNIVERSITI MALAYA
KUALA LUMPUR

April 25, 2016

Madam Sherly Sharif
Head of the Department/Chief-Executive-Officer
Yayasan Felda English Language Center/ SCYE Global Education
Kolej Yayasan Felda
Kelana Jaya, Selangor
Malaysia

Dear Madam Sherly Sharif,

I, Gringo Jusa, a Master in English as Second Language student at University of Malaya, is writing to request permission from your good office to allow me to conduct a research study at your college from May to June 2016.

The tentative title of my study is "The Acquisition of Implicit and Explicit Knowledge on Basic English Tenses through Direct and Metalinguistic Corrective Feedback". This involves classroom intervention processes, observations, and interviews. Additional help, not compulsory, from the faculty members for the implementation of the study procedures will be greatly appreciated.

Your approval to this study will be greatly appreciated. Rest assured that all information derived from the college will be treated with utmost confidentiality.

Thank you very much.

Yours sincerely,

Gringo M. Jusa

RECEIVED
28 APR 2016

SHERLY BINTI SHARIF
CHIEF EXECUTIVE OFFICER
SCYE GLOBAL EDUCATION SDN. BHD.
(896717-A)
No. 6A, Tingkat 1, Jalan Tun Mohd Fuad 2,
Taman Tun Dr. Ismail,
60000 Kuala Lumpur

SIGNED
DATE
Received

Rod Ellis May 5

to me

from: **Rod Ellis** <r.ellis@auckland.ac.nz>
to: GM Jusa <gringojusa@gmail.com>
date: Thu, May 5, 2016 at 4:03 AM
subject: RE: Corrective Feedback study inquiry
mailed-by: auckland.ac.nz
signed-by: auckland.ac.nz
encryption: Standard (TLS) Learn more
: Important mainly because it was sent directly to you.

Rod Ellis
Add to circles
Show details

1. Pica's formula - this takes account of over use of the target feature as well as suppliance in obligatory occasions.
2. The purpose of the delayed test is to ascertain whether any learning is durable. 2 weeks is a fairly standard time for a delayed test.
3. Repeated measures ANOVA is preferred when there are multiple groups and times to avoid Type 1 errors.

Ehsan Rassaei May 7

to Manijeh, me

People (2)
ehsanrassae
Add to circles
Show

Dear Gringo,
Thank you for inquiring.
Regarding your first question, the timed nature of the timed grammaticality judgment test does not allow learners to draw on their explicit knowledge and therefore they have to resort to their implicit knowledge to answer each item on the spot (see Godfroid, A., Loewen, S., Jung, S., Park, J. H., Gass, S., & Ellis, R. (2015). Timed and untimed grammaticality judgments measure distinct types of knowledge. *Studies in Second Language Acquisition*, *37*(2), 269-297.

As to your second question, the post-test should normally follow immediately after the treatment (i.e., in the same day or the following day). Likewise, delayed post-test should normally be administered at least two weeks after the post-test because it is supposed to measure learning gains after an extended period of time. With regard to your third question, metalinguistic feedback is an explicit feedback type while recasts are implicit. In that study, we were interested to examine the effects of implicit (recasts) and explicit (metalinguistic feedback) CF on L2 development. Respecting your next question, in order to understand the advantage of mixed between-within group ANOVA over one-way ANOVA some knowledge of statistics is needed. But, to put it simply, mixed ANOVA is preferable when the researcher investigates the effects of an independent variable (i.e., treatment) over time on a dependent variable. Here, the treatment is considered as the between group independent variable while time is considered as the within group independent variable and therefore the researcher is able to examine the effect of (a) the treatment and (b) time on the dependent variable. Regarding your last question, CF studies can fall within both L2 learning and L2 acquisition frameworks.

...

Ehsan Rassaei,
Assistant Professor,
IAU, Shiraz Branch,
Shiraz, Iran,

Appendix B: Proposal Defence Approval

UNIVERSITY OF MALAYA

UM.T/374/35

2 June 2016

Mr. Gringo Jusa (TGB140041)
Faculty of Languages and Linguistics

Dear Mr. Gringo

PRESENTATION OF RESEARCH PROPOSAL – GRINGO JUSA (TGB140041)

The panel of assessors is pleased to inform you that your research proposal presented on May 26, 2016 has been **accepted** subject to you attending to the comments and suggestions made by the panel **(see attached form).**

Please take note that the comments and suggestions by the panel are not exhaustive, and thus, you are expected to work with your supervisor to ensure that your research meets the standards and requirements set by the university.

As a guide to ensure that you will graduate on time, you are advised to present your Candidature Defence by the end of Semester I, 2016/2017 Academic Session.

We wish you all the best in your research efforts.

Thank you.

Yours sincerely

DR. SURINDERPAL KAUR
Deputy Dean (Postgraduate Studies)

cc: Assoc. Prof. Dr. Kuang Ching Hei (Supervisor)

ESA/nz

FACULTY OF LANGUAGES AND LINGUISTICS
University of Malaya, 50603 Kuala Lumpur, Malaysia

Dean's Office: (603) 7967 3177 • Fax: (603) 7957 9707 • E-mail: dekan_fbl@um.edu.my
Research and Development's Office: (603) 7967 3160 • E-mail: tdpp_fbl@um.edu.my
Postgraduate's Office: (603) 7967 3003/3133/3144 • Fax: (603) 7957 9707 • E-mail: ddfll_pgrad@um.edu.my
Undergraduate's Office: (603) 7967 3002/3029 • Fax: (603) 7967 3155 • E-mail: tdekan_fbl@um.edu.my
Department's Office: (603) 7967 3032/3083/3085/3158/3158 • Fax: (603) 7957 9707
Media Unit: (603) 7967 3085/3056 • Fax: (603) 7957 9707
Multimedia Planning Unit: (603) 7967 3162 • E-mail: multimedia_fbl@um.edu.my
Language Unit: (603) 7967 3161 • E-mail: ketua_unitbahasa@um.edu.my

Appendix C: Candidature Defence Approval

UNIVERSITY OF MALAYA

UM.T/374/36

17 January 2017

Mr. Gringo Jusa (TGB140041)
Faculty of Languages and Linguistics

Dear Mr. Gringo

CANDIDATURE DEFENCE BY MASTER'S DEGREE CANDIDATE – GRINGO JUSA (TGB140041)

The panel of assessors is pleased to inform you that your candidature defence presented on January 12, 2017 has been **approved.**

Please take note that the comments and suggestions by the panel are not exhaustive, and thus, you are expected to work with your supervisor to ensure that your research meets the standards and requirements set by the university **(see attached form).**

You are reminded to submit the **'Three Months' Notice'** via myum.um.edu.my as soon as you receive this approval letter to facilitate the approval of your title and appointment of your examiners.

We wish you all the best in your research efforts, and the successful completion of your programme of study.

Thank you.

Yours sincerely

DR. SURINDERPAL KAUR
Deputy Dean (Postgraduate Studies)

cc: Assoc. Prof. Dr. Kuang Ching Hei (Supervisor)

SK/ESA/nzkk

FACULTY OF LANGUAGES AND LINGUISTICS
University of Malaya, 50603 Kuala Lumpur, Malaysia

Dean's Office: (603) 7967 3177 • Fax: (603) 7957 9707 • E-mail: dekan_fbl@um.edu.my
Postgraduate Office: (603) 7967 3003/3133/3144 • Fax: (603) 7957 9707 • E-mail: ddfl_pgrad@um.edu.my
Undergraduate Office: (603) 7967 3002/3029 • Fax: (603) 7967 3155 • E-mail:dekan_fbl@um.edu.my
Department Office: (603) 7967 3032/3063/3065/3156/3158 • Fax: (603) 7957 9707
Media Unit: (603) 7967 3085/3056 • Fax: (6030 7957 9707 • E-mail: husnik@um.edu.my
Multimedia Planning Unit: (603) 7967 3162 • E-mail: multimedia_fbl@um.edu.my
Language Unit: (603) 7967 3161 • E-mail: ketua_unitbahasa@um.edu.my

Appendix D: Test reliability

Timed Grammatical Judgment Test

Correlations

		TGJT_Time1	TGJT_Time2
TGJT_Time1	Pearson Correlation	1	**.772****
	Sig. (2-tailed)		.009
	N	10	10
TGJT_Time2	Pearson Correlation	.772**	1
	Sig. (2-tailed)	.009	
	N	10	10

**. Correlation is significant at the 0.01 level (2-tailed).

There was a positive correlation between TGJT tested in Time 1 and TGJT tested in Time 2 to students in the pilot test, r = .772.

Metalinguistic Knowledge Test

		MKT_Time 1	MKT_Time 2
MKT_Time1	Pearson Correlation	1	.737*
	Sig. (2-tailed)		.015
	N	10	10
MKT_Time2	Pearson Correlation	.737*	1
	Sig. (2-tailed)	.015	
	N	10	10

*. Correlation is significant at the 0.05 level (2-tailed).

There was a positive correlation between MKT tested in Time 1 and MKT tested in Time 2 to students in the pilot test, r = .737.

Picture Narrative Writing Test

		Picture_Time1	Picture_Time2
Picture_Time1	Pearson Correlation	1	**.716***
	Sig. (2-tailed)		.020
	N	10	10
Picture_Time2	Pearson Correlation	.716*	1
	Sig. (2-tailed)	.020	
	N	10	10

*. Correlation is significant at the 0.05 level (2-tailed).

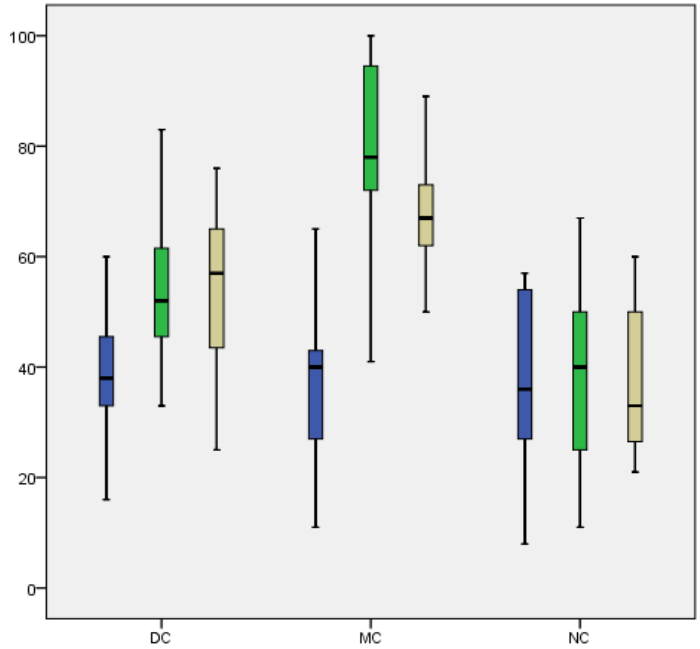
There was a positive correlation between Picture Test tested in Time 1 and Picture Test tested in Time 2 to students in the pilot test, r = .716.

‘

Appendix E: Two-Way Mixed Method

a. Picture Narrative Writing Test

- Boxplot Analysis



Box plot results for three test groups during pre-, post- and delayed post-test

- Shpiro-Wilk Test

Different tests of normality

	Groups	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
		Statistic	df	Sig.	Statistic	Df	Sig.
Pre-Test	DC	.111	15	.200*	.984	15	.990
	MC	.175	15	.200*	.959	15	.669
	NC	.167	15	.200*	.915	15	.161
Immediate Post-Test	DC	.142	15	.200*	.960	15	.687
	MC	.166	15	.200*	.899	15	.091
	NC	.152	15	.200*	.950	15	.526
Delayed Post-Test	DC	.115	15	.200*	.969	15	.848
	MC	.157	15	.200*	.934	15	.308
	NC	.166	15	.200*	.888	15	.062

*. This is a lower bound of the true significance.

b. Lilliefors Significance Correction

- Levene's Test

Levene's test of equality of error variances

	F	df1	df2	Sig.
Pre-Test	1.706	2	42	.194
Post-Test	.800	2	42	.456
Delayed Post-Test	1.122	2	42	.335

Tests the null hypothesis that the error variance of the dependent variable is equal across groups.
a. Design: Intercept + Groups
Within Subjects Design: Tests

- Overall Within Subject Effects

Test of Within Subject Effects between different tests for PNT

	Source		Type III Sum of Squares	Df	Mean Square	F	Sig.	Partial Eta Squared
DC	Tests	Sphericity Assumed	2394.311	2	1197.156	12.050	.000	.463
		Greenhouse-Geisser	2394.311	1.490	1606.516	12.050	.001	.463
		Huynh-Feldt	2394.311	1.627	1471.433	12.050	.001	.463
		Lower-bound	2394.311	1.000	2394.311	12.050	.004	.463
	Error (Tests)	Sphericity Assumed	2781.689	28	99.346			
		Greenhouse-Geisser	2781.689	20.865	133.317			
		Huynh-Feldt	2781.689	22.781	122.107			
		Lower-bound	2781.689	14.000	198.692			
MC	Tests	Sphericity Assumed	13650.53	2	6825.267	42.87	.00	.754
		Greenhouse-Geisser	13650.53	1.738	7853.322	42.87	.00	.754
		Huynh-Feldt	13650.53	1.963	6953.092	42.87	.00	.754
		Lower-bound	13650.53	1.000	13650.533	42.87	.00	.754
	Error (Tests)	Sphericity Assumed	4457.467	28	159.195			
		Greenhouse-Geisser	4457.467	24.335	183.174			
		Huynh-Feldt	4457.467	27.485	162.177			
		Lower-bound	4457.467	14.000	318.390			
NC	Tests	Sphericity Assumed	6.978	2	3.489	.017	.983	.001
		Greenhouse-Geisser	6.978	1.780	3.920	.017	.975	.001
		Huynh-Feldt	6.978	2.000	3.489	.017	.983	.001
		Lower-bound	6.978	1.000	6.978	.017	.899	.001
	Error (Tests)	Sphericity Assumed	5811.69	28	207.56			
		Greenhouse-Geisser	5811.69	24.919	233.23			
		Huynh-Feldt	5811.69	28.00	207.56			
		Lower-bound	5811.69	14.000	415.12			

- Multiple Comparison

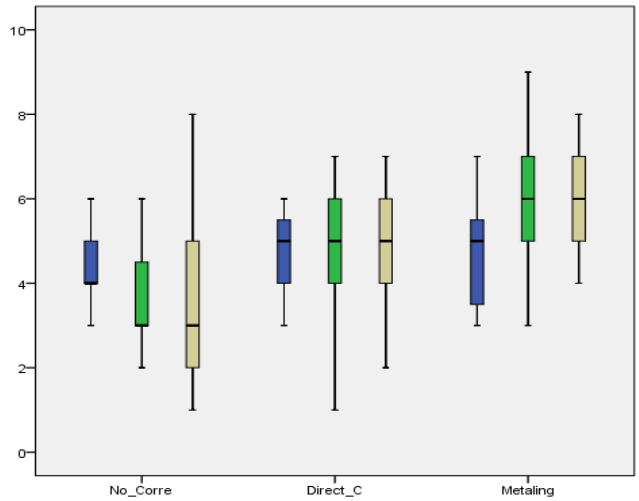
Multiple Comparison between different groups with pre-, post- and delayed post-tests as dependent variables

(I) Tests		(J) Tests	Mean Difference (I-J)	Std. Error	Sig.[b]	95% Confidence Interval for Difference[b]	
						Lower Bound	Upper Bound
Pre-test	DC	MC	1.5333	5.30197	.955	-11.3478	14.4144
		NC	.4000	5.30197	.997	-12.4811	13.2811
	MC	DC	-1.5333	5.30197	.955	-14.4144	11.3478
		NC	-1.1333	5.30197	.975	-14.0144	11.7478
	NC	DC	-.4000	5.30197	.997	-13.2811	12.4811
		MC	1.1333	5.30197	.975	-11.7478	14.0144
Post-test	DC	MC	-23.7333*	5.77325	.001	-37.7594	-9.7073
		NC	15.5333*	5.77325	.027	1.5073	29.5594
	MC	DC	23.7333*	5.77325	.001	9.7073	37.7594
		NC	39.2667*	5.77325	.000	25.2406	53.2927
	NC	DC	-15.5333*	5.77325	.027	-29.5594	-1.5073
		MC	-39.2667*	5.77325	.000	-53.2927	-25.2406
Delayed Post-test	DC	MC	-14.6667*	4.76859	.010	-26.2519	-3.0814
		NC	16.9333*	4.76859	.003	5.3481	28.5186
	MC	DC	14.6667*	4.76859	.010	3.0814	26.2519
		NC	31.6000*	4.76859	.000	20.0147	43.1853
	NC	DC	-16.9333*	4.76859	.003	-28.5186	-5.3481
		MC	-31.6000*	4.76859	.000	-43.1853	-20.0147

Based on observed means:
The error term for Pre-test is Mean Square (Error) = 210.832
The error term for Post-test is Mean Square (Error) = 249.978.
The error term for Delayed Post-test is Mean Square (Error) = 170.546.
*. The mean difference is significant at the .05 level.
b. Adjustment for multiple comparisons: Bonferroni.

b. Metalinguistic Knowledge Test

- Boxplot Analysis



Box plot for MKT conducted at three different occasions over the period of three weeks

- Levene's Test

Levene's Test for examining the normal distribution of data obtained for MKT

	F	df1	df2	Sig.
Pre Test	2.295	2	42	.113
Immediate Post Test	.487	2	42	.618
Delayed Post Test	1.932	2	42	.158

Tests the null hypothesis that the error variance of the dependent variable is equal across groups.

a. Design: Intercept + Groups
Within Subjects Design: Tests

- Overall Within Subject Effects

- Test of Within Subject Effects of three groups for MKT

	Source		Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
DC	Tests	Sphericity Assumed	.711	2	.356	.312	.735	.022
		Greenhouse-Geisser	.711	1.895	.375	.312	.723	.022
		Huynh-Feldt	.711	2.000	.356	.312	.735	.022
		Lower-bound	.711	1.000	.711	.312	.586	.022
	Error (Tests)	Sphericity Assumed	31.956	28	1.141			
		Greenhouse-Geisser	31.956	26.525	1.205			
		Huynh-Feldt	31.956	28.000	1.141			
		Lower-bound	31.956	14.000	2.283			
MC	Tests	Sphericity Assumed	18.311	2	9.156	6.149	.006	.305
		Greenhouse-Geisser	18.311	1.971	9.290	6.149	.006	.305
		Huynh-Feldt	18.311	2.000	9.156	6.149	.006	.305
		Lower-bound	18.311	1.000	18.311	6.149	.026	.305
	Error (Tests)	Sphericity Assumed	41.689	28	1.489			
		Greenhouse-Geisser	41.689	27.593	1.511			
		Huynh-Feldt	41.689	28.000	1.489			
		Lower-bound	41.689	14.000	2.978			
NC	Tests	Sphericity Assumed	4.933	2	2.467	1.970	.158	.123
		Greenhouse-Geisser	4.933	1.799	2.743	1.970	.164	.123
		Huynh-Feldt	4.933	2.000	2.467	1.970	.158	.123
		Lower-bound	4.933	1.000	4.933	1.970	.182	.123
	Error (Tests)	Sphericity Assumed	35.067	28	1.252			
		Greenhouse-Geisser	35.067	25.181	1.393			
		Huynh-Feldt	35.067	28.000	1.252			
		Lower-bound	35.067	14.000	2.505			

- Multiple Comparison

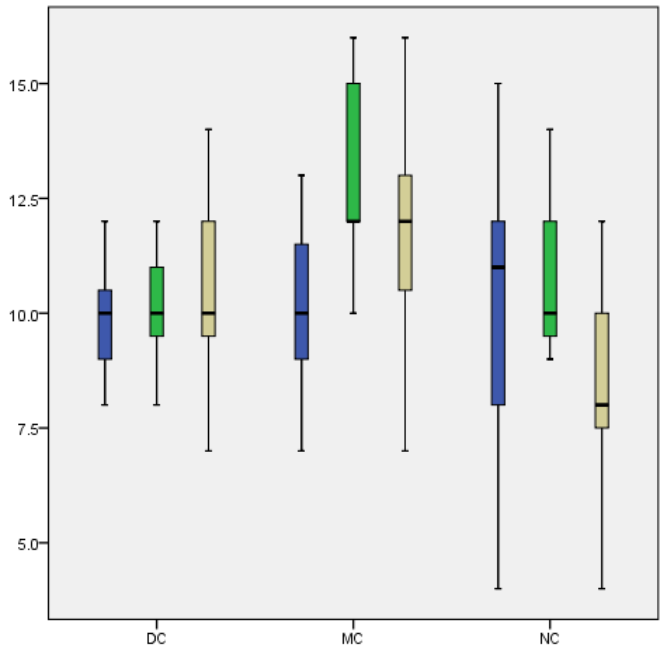
Multiple comparisons of different groups using different tests as dependent variables

(I) Tests		(J) Tests	Mean Difference (I-J)	Std. Error	Sig.[b]	95% Confidence Interval for Difference[b]	
						Lower Bound	Upper Bound
Pre-test	NC	DC	-.5333	.40786	.399	-1.5242	.4576
		MC	-.3333	.40786	.695	-1.3242	.6576
	DC	NC	.5333	.40786	.399	-.4576	1.5242
		MC	.2000	.40786	.876	-.7909	1.1909
	MC	NC	.3333	.40786	.695	-.6576	1.3242
		DC	-.2000	.40786	.876	-1.1909	.7909
Post-test	NC	DC	-.9333	.54821	.216	-2.2652	.3985
		MC	-2.4667*	.54821	.000	-3.7985	-1.1348
	DC	NC	.9333	.54821	.216	-.3985	2.2652
		MC	-1.5333*	.54821	.021	-2.8652	-.2015
	MC	NC	2.4667*	.54821	.000	1.1348	3.7985
		DC	1.5333*	.54821	.021	.2015	2.8652
Delayed Post-test	NC	DC	-1.2667	.59380	.095	-2.7093	.1760
		MC	-2.2667*	.59380	.001	-3.7093	-.8240
	DC	NC	1.2667	.59380	.095	-.1760	2.7093
		MC	-1.0000	.59380	.223	-2.4426	.4426
	MC	NC	2.2667*	.59380	.001	.8240	3.7093
		DC	1.0000	.59380	.223	-.4426	2.4426

Based on observed means:
The error term for Pre-test is Mean Square (Error) = 1.248
The error term for Post-test is Mean Square (Error) = 2.254.
The error term for Delayed Post-test is Mean Square (Error) = 2.644
*. The mean difference is significant at the .05 level.
b. Adjustment for multiple comparisons: Bonferroni.

c. Timed Grammatical Judgment Test

- Boxplot Analysis



Results of the boxplot for TGJT

- Levene's Test

Levene's Test of Equality of Error Variance for TGJT

	F	df1	df2	Sig.
Pre Test	5.152	2	42	.010
Immediate Post Test	2.379	2	42	.105
Delayed Post Test	.713	2	42	.496

Tests the null hypothesis that the error variance of the dependent variable is equal across groups.
a. Design: Intercept + Group
Within Subjects Design: Tests

- Overall Within Subject Effects

- Test of Within Subject Effects of three groups for TGJT

Source			Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
DC	Tests	Sphericity Assumed	4.844	2	2.422	1.596	.221	.102
		Greenhouse-Geisser	4.844	1.724	2.810	1.596	.224	.102
		Huynh-Feldt	4.844	1.943	2.493	1.596	.221	.102
		Lower-bound	4.844	1.000	4.844	1.596	.227	.102
	Error (Tests)	Sphericity Assumed	42.489	28	1.517			
		Greenhouse-Geisser	42.489	24.134	1.761			
		Huynh-Feldt	42.489	27.208	1.562			
		Lower-bound	42.489	14.000	3.035			
MC	Tests	Sphericity Assumed	70.578	2	35.289	9.679	.001	.409
		Greenhouse-Geisser	70.578	1.643	42.948	9.679	.002	.409
		Huynh-Feldt	70.578	1.833	38.504	9.679	.001	.409
		Lower-bound	70.578	1.000	70.578	9.679	.008	.409
	Error (Tests)	Sphericity Assumed	102.089	28	3.646			
		Greenhouse-Geisser	102.089	23.007	4.437			
		Huynh-Feldt	102.089	25.662	3.978			
		Lower-bound	102.089	14.000	7.292			
NC	Tests	Sphericity Assumed	58.711	2	29.356	4.933	.015	.261
		Greenhouse-Geisser	58.711	1.814	32.363	4.933	.018	.261
		Huynh-Feldt	58.711	2.000	29.356	4.933	.015	.261
		Lower-bound	58.711	1.000	58.711	4.933	.043	.261
	Error (Tests)	Sphericity Assumed	166.622	28	5.951			
		Greenhouse-Geisser	166.622	25.398	6.560			
		Huynh-Feldt	166.622	28.000	5.951			
		Lower-bound	166.622	14.000	11.902			

- Multiple Comparison

Multiple-level comparison between different test groups with test timing as dependent variable

(I) Tests	(J) Tests	Mean Difference (I-J)	Std. Error	Sig.[b]	95% Confidence Interval for Difference[b]	
					Lower Bound	Upper Bound
DC	MC	-.1333	.75649	.983	-1.9712	1.7046
	NC	-.4667	.75649	.812	-2.3046	1.3712
MC	DC	.1333	.75649	.983	-1.7046	1.9712
	NC	-.3333	.75649	.899	-2.1712	1.5046

Pre-test	NC	DC	.4667	.75649	.812	-1.3712	2.3046
		MC	.3333	.75649	.899	-1.5046	2.1712
Post-test	DC	MC	-.1333	.75649	.983	-1.9712	1.7046
		NC	-.4667	.75649	.812	-2.3046	1.3712
	MC	DC	.1333	.75649	.983	-1.7046	1.9712
		NC	-.3333	.75649	.899	-2.1712	1.5046
	NC	DC	.4667	.75649	.812	-1.3712	2.3046
		MC	.3333	.75649	.899	-1.5046	2.1712
Delayed Post-test	DC	MC	-.8000	.83874	.610	-2.8377	1.2377
		NC	2.4000*	.83874	.018	.3623	4.4377
	MC	DC	.8000	.83874	.610	-1.2377	2.8377
		NC	3.2000*	.83874	.001	1.1623	5.2377
	NC	DC	-2.4000*	.83874	.018	-4.4377	-.3623
		MC	-3.2000*	.83874	.001	-5.2377	-1.1623

Based on observed means:
The error term for Pre-test is Mean Square (Error) = 4.292
The error term for Post-test is Mean Square (Error) = 2.921.
The error term for Delayed Post-test is Mean Square (Error) = 5.276.
*. The mean difference is significant at the .05 level.
b. Adjustment for multiple comparisons: Bonferroni.

Appendix F: Sample Tests

a) Sample Metalinguistic Knowledge Test (Adapted from Ellis, 2005)

Name: _____ **Address:** _____

Level: _____ **Age:** _____

Birthday: _____

I. Background Information

1. Which country do you come from? _____
2. What is your mother tongue (i.e. the language first acquired)? _____ (If your answer to Question 2 is English, go to Section 6.)
3. How old were you when you started to learn English? _____ years old
4. How many years have you been learning English (including the years at school in New Zealand)? _____ years
5. Altogether, how many years have you spent living in a country where English is widely spoken (including New Zealand)?)? _____ years
6. What other languages have you studied?

Language	***Length of time I have studied it***

II. Metalinguistic Knowledge Test

In this part of the test there are 20 sentences. The part of the sentence containing the error is underlined. For each sentence, choose which statement best explains the error. Circle a, b, c or d to indicate your choice.

Questions

1. Martin lost <u>his friend book</u>.

a. We need possessive ‘s’ to show that the friend owns the book.
b. You cannot have two nouns next to one another in a sentence.
c. The verb refers to a personal object, so must have an apostrophe.
d. Insert ‘of’ before book to show that it belongs to the friend.

2. I <u>was born</u> in Malaysia.

a. Use “were” because “I” is a special pronoun.
b. Use “was” because the sentence states general truth.
c. We should use “to” to clearly indicate the place in the sentence.
d. Use the past tense of verb ‘be’ because it already happened.

b) Sample Timed Grammatical Judgment Test (Adapted from Ellis, 2005; Rassaei et al., 2012)

Name: _____ **Address:** _____

Level: _____ **Age:** _____

Birthday: _____

TIMED GRAMMATICAL JUDGMENT TEST

Identify whether the sentences below are grammatical or not. Tick the appropriate box. Indicate whether you use "feel" or "rule" when answering. You have 8 minutes to answer it.

R/F	YES	NO	ITEMS
			1. I haven't seen him for a long time.
			2. Liao says he wants buying a car next week.
			3. Martin completed his assignment and print it out the other day.
			4. We will leave tomorrow, isn't it?
			5. He has played soccer very well since 2010.
			6. Did Keiko completed her homework?

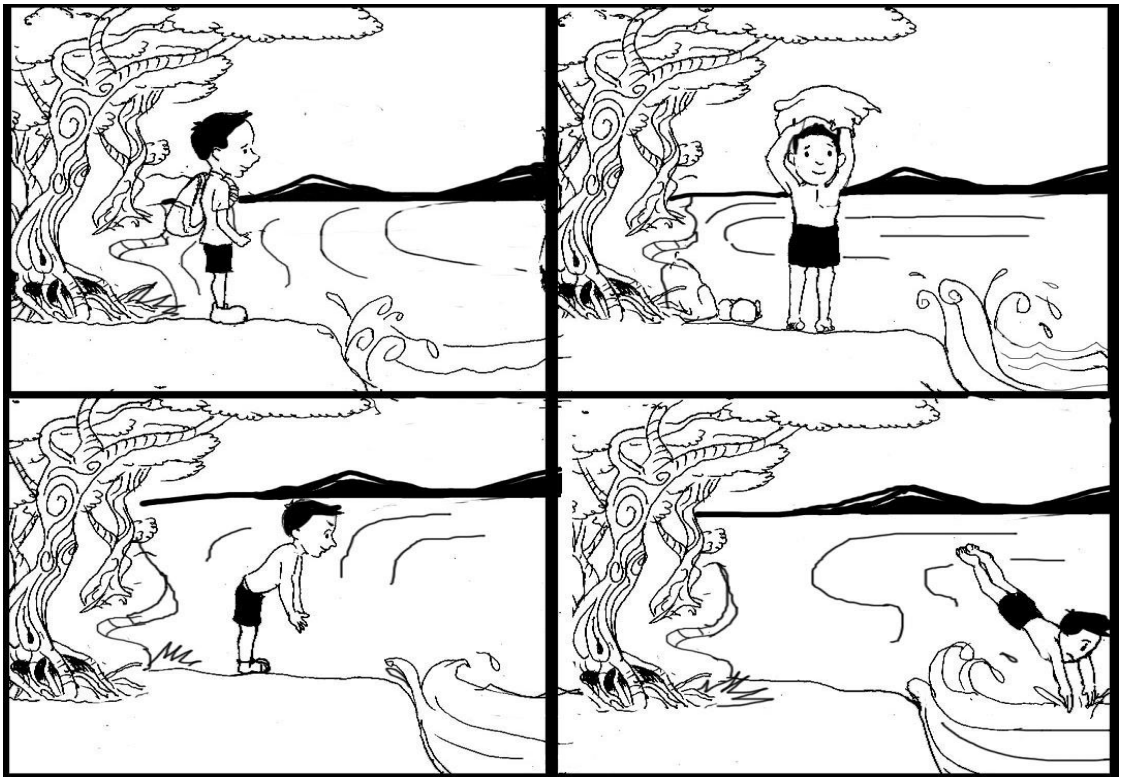
c) Picture test (Adapted from Shintani & Ellis, 2013)

Name: _____ **Address:** _____

Level: _____ **Age:** _____

Birthday: _____

Write a narrative description of the series of events provided in the picture.



_____
_____
_____
_____
_____
_____
_____
_____
_____
_____
_____
_____
_____
_____
_____
_____

Appendix G: Sample Metalinguistic Handout

METALINGUISTIC HANDOUT

Example:

- ✓ Plastic **floats** in water.
- ✓ Malaysia **is** in Southeast Asia.

Code: **PC**

PRESENT CONTINUOUS (PC)

The action is happening now. Not in the past and not in the future.

The form is *BE + VERB -ING* ***(VB)***

Example:

- ✓ You are listening to me.
- ✓ He is writing.

The action is not yet finished but still happening.

Example:

I am studying English this semester, but he isn't studying now. He is playing football outside.

Code: **SP**

PAST SIMPLE (SP)

The action ended in the past. No real connection with the current time.

The form is VERB + ED

Example:

- ✓ Joshua **tried** his new shoes *last night*.
- ✓ I withdrew my money last *Monday*.

It is also important to note of irregular verbs who have different forms in their past tense.

Example:

- ✓ The bird **flew** really fast. (fly)
- ✓ I **ate** the bread this morning. (eat)

Code: **CP**

PAST CONTINUOUS (CP)

Action happened over time in the past when something happened.

Example:

- ✓ He **was eating** when his mother **arrived**.
- ✓ The students **were talking** when their teacher opened the door.

It also indicated an action in the past that is no longer in duration.

Example:

- ✓ He **was napping** at 3:00 PM.
- ✓ They **were listening** to the music around 9 in the morning.

Using DO, DOES, and DID

Code: **D1**

DO (plural-present)

To use it in a sentence, the formation is do + verb (base).

Example:

- ✓ They **do believe** in him.

For questions, 'do' starts the sentence followed by a noun then by the **base form** of the verb.

Example:

- ✓ **Do** you **believe** in him?
- ✓ **Do** you **take** breakfast every day?

Code: **D2**

DOES (singular- present)

To use it in a sentence, the formation is does + verb (base).

Appendix H: Statistical Method Consultation Reference

1. Dr. Mahmoud Danaee is currently a visiting research fellow at University Malaya, Academic Development Center (ADeC). He is an academic member at Islamic Azad University (Iran) since 1998. He graduated from Tarbiat Modares University in Biometrical genetics in 1997 and finished his PhD in Biotechnology (Bioprocess Modeling) from UPM. As a Bio-statistician, during last 20 years. He has taught courses in statistics, experimental design, advance statistical methods, Research Methodology at undergraduate and postgraduate level in different disciplines. He conducted many workshops and courses such as Basic and advance statistical methods using SPSS, Structural Equation Modeling (AMOS& Smart-PLS), Design of Experiments (RSM) since 2008 in Malaysian universities such as UPM, UM,UKM and UTM. He has provided consultancy services and analyzing data for many Master and PhD students during 2008 until 2015.

   Taken from https://umexpert.um.edu.my/mdanaee.html

2. Ildefonso Halipa is a former Mathematics, Research, and Statistics lecturer at St. La Salle University in the Philippines. He graduated with Masters in Teaching Mathematics and attended credit hours for PhD studies in Mathematics. He has been teaching the subject for 10 years now.