WRITTEN CORRECTIVE FEEDBACK, WORKING MEMORY, AND THE DEVELOPMENT OF EXPLICIT AND IMPLICIT KNOWLEDGE OF ENGLISH PLURALS

Ornuma Chingchit

This thesis is submitted in fulfilment of the requirements

for the degree of Doctor of Philosophy in Linguistics,

Department of Linguistics and English Language,

Lancaster University

October, 2021

DECLARATION

I declare that the work in this thesis has not been submitted for any degrees at other universities, and that the work, except for the publications indicated, is entirely my own.

<div align="right">

Ornuma Chingchit

October, 2021

</div>

WRITTEN CORRECTIVE FEEDBACK, WORKING MEMORY, AND THE DEVELOPMENT OF EXPLICIT AND IMPLICIT KNOWLEDGE OF ENGLISH PLURALS

Ornuma Chingchit, M.A.

Thesis supervisors: Alison Mackey, Ph.D. & Jenefer Philp, Ph.D.

ABSTRACT

Unlike oral corrective feedback that is unanimously well accepted in the second language acquisition (SLA) field as an effective tool that helps promote learners' acquisition (e.g., Adams, Nuevo and Egi, 2011; Gass and Mackey, 2015), the role of written corrective feedback (WCF) and the extent to which it could help second language (L2) learners develop their second languages are still questionable (e.g., Polio, 2012; Shintani and R. Ellis, 2013). Few attention has also been given to test its long-term effect towards L2 development. Contributing to this gap presented in the literature, the current study examined and compared the effectiveness of direct *focused* and direct *unfocused* WCF on the development of Thai EFL learners' explicit and implicit knowledge of English plurals. Learner differences in working memory capacity (WMC) were also measured in order to explore the extent to which this cognitive factor mediated the effectiveness of WCF.

This study employed a pre-post-delayed posttest design and was carried out over the course of a 9-month period. Seventy-two low intermediate learners were randomly assigned to two experimental groups, one of which received direct *focused* WCF and one of which received direct *unfocused* WCF, and a control group which received content feedback. All groups completed batteries of pre, post and delayed posttests, involving an untimed grammatical judgement test (UGJT), a metalinguistic knowledge test (MKT), a timed essay-writing test, and a timed oral elicited imitation test (OEIT). Learner differences in WMC were measured using backward digit span and operation span

(OSpan) tasks, all of which were conducted in learners' first language (L1). In addition, two learners from each group were randomly selected to take part in qualitative interviews to explore potential variables that might mediate the effectiveness of WCF. After the 6-week treatment period, all learners completed exit questionnaire surveys. The delayed posttest was administered three months after the posttest.

The results revealed that direct *focused* and *unfocused* WCF provided for the experimental groups and content feedback provided for the control group were equally effective in facilitating the development of learners' explicit and implicit knowledge of English plurals. Learners' educational and instructional contexts as well as their proficiency level are posited to have influenced the yielded results. In addition, learner differences in WMC did not moderate the extent to which learners benefited from WCF. Feedback type, instructional context as well as learners' proficiency level are key factors attributing to the absence of WMC's moderating effect. A number of theoretical and pedagogical implications are discussed based on these findings.

ACKNOWLEDGEMENTS

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter I

INTRODUCTION

1.1. Background and motivation for the study

Written Corrective Feedback (WCF) is usually regarded by composition teachers as an effective tool to help second language (L2) learners improve their written accuracy. It is also believed that by giving L2 learners systematic WCF over time, L2 learners will eventually become more independent writers who are capable of producing more linguistically accurate L2 texts (Evans, Hartshorn, McCollum, & Wolfersberger, 2010). Even though some teachers are skeptical about the effectiveness of WCF due to the fact that some learners seemingly fail to gain their knowledge from feedback, most composition teachers do, to some extent, believe that WCF is efficacious in enhancing learners' grammatical accuracy contributing to learners' L2 development.

However, since Truscott's (1996) controversial argument claiming that WCF should be abandoned from practice because there was neither empirical nor theoretical justification attesting its effectiveness as a pedagogical tool that could help learners develop their L2 knowledge, the effectiveness and necessity of WCF have been vehemently debated and reevaluated. In his 1996 polemic argument (also in his subsequent 1999, 2004, 2007, 2010 articles), Truscott posited that WCF might at best help L2 learners obtain explicit knowledge, which is merely useful for grammar monitoring or text editing, it would not, however, contribute to learners' implicit knowledge (i.e., acquisition). Truscott's contention was based on the premises that (1) WCF or as he called it, "grammar correction", can only contribute to learners' explicit knowledge, and (2) explicit knowledge can never become implicit knowledge.

It is important to note that within the second language acquisition (SLA) field, implicit knowledge seems to play a more significant role than explicit knowledge. Implicit knowledge is usually viewed as the ultimate goal for L2 learners to achieve, since it can be rapidly accessed and therefore enabling learners to produce the target language with high automaticity, i.e., without conscious effort, similar to the knowledge possessed by a native speaker of a language (R. Ellis, Loewen, Elder, Erlam, Philp, &

Reinders, 2009). It is the tacit or unconscious type of knowledge that is acquired without learners' awareness or intentionality (R. Ellis et al., 2009) via implicit learning[1] (Rebuschat, 2013, 2015). Since implicit knowledge enables L2 learners to produce the target language with high automaticity similar to the way a native speaker produces his/her own native language, factors contributing to this type of knowledge are of interest to most SLA theorists. On the contrary, explicit or conscious knowledge which is usually acquired through explicit learning (Rebuschat, 2013, 2015) and can only be accessed through conscious attention control (R. Ellis et al., 2009), cannot help L2 learners produce the target language with high automaticity. Thus, explicit knowledge is perceived to play a less important role in SLA. Since Truscott conjectured that WCF can only contribute to explicit knowledge, not implicit knowledge, which is the ultimate goal of L2 learning, this explains why he contended that WCF is a futile practice.

Further, given that some SLA theorists (e.g., Krashen, 1982, 1985; Schwartz, 1993) have proposed that explicit knowledge (i.e., "*learnt knowledge*") can never become implicit knowledge (i.e., "*acquired knowledge*"), and that there is no interface (interaction) between the two types of knowledge, the role of explicit knowledge in L2 learning is even more disregarded. Believing in this non-interface position, Truscott presumed that WCF is ineffective for acquisition and thereby called for the renounce of the practice (i.e., providing WCF to learners). In his supposition, since WCF cannot contribute to implicit knowledge, or as some researchers call it the "genuine knowledge" of a language, and the knowledge obtained from WCF cannot aid the development of implicit knowledge, composition teachers should discontinue this tedious and time-consuming convention and adopt other effective L2 strategies or instructions instead.

Nonetheless, it is worth noting that for the past decades, the non-interface position has been countered by a growing number of researchers such as Anderson (1983, 1985, 2007), DeKeyser (1997, 1998, 2007, 2015), McLaughlin (1987, 1990), and N. Ellis

---

[1] According to Rebuschat (2013, 2015), implicit learning refers to the learning that "occurs without [learners'] intention to learn and without [learners'] awareness of what has been learned" (p. 597), and as a result, learners are usually unaware of the knowledge they have acquired. On the other hand, explicit learning refers to learning with intention or learners are aware of what they have been learned and therefore learners are more often than not aware of the knowledge they have acquired.

(2005, 2012, 2015b), all of whom strongly argue that there is interaction between these two distinct types of knowledge. Explicit knowledge plays an active role in implicit knowledge development (DeKeyser, 2015; N. Ellis, 2015b; Wulff & N. Ellis, 2018). Indeed, up until now, the interface issue is still unresolved. Further investigations are actually needed to determine if (1) the interface between the two discrete types of knowledge exists, (2) explicit knowledge helps promote the development of implicit knowledge, and (3) explicit knowledge proves futile for L2 acquisition as claimed by some groups of scholars.

Ignited by Truscott's (1996) strong proposition on the ineffectiveness of WCF based on the non-interface premise, researchers in the L2 writing field have raised a number of theoretical discussions about whether WCF can only contribute to explicit knowledge or it can also contribute to the development of implicit knowledge. A substantial body of empirical research were also carried out to investigate the effectiveness of WCF on L2 development. However, these early WCF studies (e.g., Ashwell, 2000; Kepner, 1991; Sheppard, 1992), focused more on pedagogical issues such as which type of WCF might be more effective in improving the quality of learners' written work or short-term accuracy (Polio, 2012). They did not attempt to answer the specific theoretical question raised by Truscott regarding the potential of WCF in promoting the development of learners' explicit and implicit knowledge (i.e., long term gain). That is, early WCF work did not attempt to uncover whether the provision of WCF that leads to an increase in grammatical accuracy also benefits the development of learners' explicit and implicit knowledge (Shintani & R. Ellis, 2013). As pointed out by Santos, López-Serrano Rosa, and Manchón (2010), early WCF studies focused more on feedback for accuracy rather than feedback for acquisition (also see Manchón, 2011; Manchón & Williams, 2016 for a review). Even though many of these early studies found that WCF led to learners' increases in grammatical accuracy (e.g., Ashwell, 2000; Chandler, 2003; Fathman & Whalley, 1990; Ferris, 2006; Ferris & Robert, 2001), it is as Polio (2012) suggests that an increase in grammatical accuracy, especially on revised texts (which were mostly found in these early WCF work) cannot be regarded as evidence of learning.

Additionally, no consensus has been reached concerning which type of WCF is the most effective in helping improving learners' short-term accuracy as well as their writing abilities (Kang & Han, 2015; Liu & Brown, 2015; Lee, 2019; Mao & Lee, 2020; Polio, 2012). The divergent findings are largely due to the early studies' methodological flaws and measurement inconsistencies, such as there were no strict control groups, mixed types of WCF treatments were provided for a single group, revision was used as measure of increased accuracy instead of using new pieces of writing as evidence of learning, and so on (Bruton, 2009, 2010; Bitchener & Storch, 2016; Guénette, 2007, Liu & Brown, 2015; Mao & Lee, 2020; Norris & Ortega, 2000; Polio, 2012; Van Beuningen, 2010). These inconsistencies make it difficult for the early findings to be generalized to other contexts and to be considered as robust evidence attesting the effectiveness of WCF on L2 development.

Even though many recent WCF studies successfully avoided the aforementioned methodological and measurement issues, most of them still focused on investigating the effectiveness of WCF as a pedagogical tool helping learners improve their L2 accuracy (short-term gain), and did not attempt to provide robust evidence in support of the effect of WCF on the development of explicit and implicit knowledge, despite having a theoretical reason to believe that the increase in accuracy resulting from WCF might have positive effects on learners' L2 development (DeKeyser, 1997, 1998, 2007, 2015; N. Ellis, 2005, 2011, 2015b). To date, to the best of my knowledge, only four studies (i.e., Nemati, Alavi, & Mohebbi, 2019; Rezazadeh, Tavakoli & Eslami Rasekh, 2015; Shintani & R. Ellis, 2013; Shooshtari, Vahdat, & Negahi, 2019) were actually conducted on SLA grounds examining the effect of WCF on the development of learners' explicit and implicit knowledge. It is posited that some design and methodological complications may account for limited investigations under this agenda.

One major obstacle stems from a lack of valid and reliable measures used to assess learners' implicit knowledge[2] (Akakura, 2012; DeKeyser, 2009; R. Ellis, 2005;

---

[2] Although some researchers have endorsed some implicit knowledge tests (e.g., oral elicited imitation task, timed grammatical judgment task, etc.) as reliable and valid, the reliability and validity of these tests remain questionable for others (see Bowles, 2011; Culbertson, Andersen & Christensen, 2020; R. Ellis et al., 2009; Kim & Nam, 2017; Sarandi, 2015; Spada, Shiu, Tomita, 2015; Suzuki & DeKeyser, 2015; Vafaee, Suzuki, &

Vafaee, Suzuki, & Kachisnke, 2017, also see Rebuschat, 2013, 2015 for a review). Other obstacles concern the design and methodology adopted for the study. For instance, although single shot treatments are the dominant approach in the WCF field (Kang & Han, 2015; Liu & Brown, 2015; Storch, 2018), a longitudinal design is a more appropriate approach for this research agenda, since it takes time for the effect of WCF to emerge and for the explicit and implicit knowledge to develop (DeKeyser, 2015; Paradis, 2009; Suzuki & DeKeyser, 2015, 2017). That is, a longitudinal design may allow researchers to effectively scrutinize the effect of WCF on learners' explicit and implicit knowledge development.

However, in practice, a longitudinal design may present several challenges. First of all, a vast majority of WCF studies are usually conducted in a real classroom setting to obtain high ecological validity, making it difficult to withhold or apply particular types of WCF on learners for a long period of time without affecting their actual learning, especially if the studies are administered in a writing class where learners might expect to receive certain kinds of feedback. In addition, it will be onerous to control other classroom related variables that might arise and mediate the findings if the study is prolonged. Due to these challenges, two WCF meta-analyses (i.e., Kang & Han, 2015; Liu & Brown, 2015) have discovered that most WCF research conducted in recent years usually limited the amount and frequency of feedback provided and single feedback treatment studies have become the most dominant design of all.

As mentioned, by adopting a single feedback treatment design or otherwise limiting the amount and frequency of WCF, it is more difficult to detect the effect of WCF on either learner's explicit or implicit knowledge. In part, this may explain why Shintani and R. Ellis' (2013) study, which investigated the effectiveness of WCF on L2 development and adopted the single shot treatment design, did not find any significant effect of WCF on learners' implicit knowledge development.

Indeed, Shintani and R. Ellis (2013) acknowledged earlier in their study that they might not find any effect of WCF particularly on learners' implicit knowledge because of

---

Kachisnke, 2017; Wu & Ortega, 2013; Yan, Maeda, Lv & Ginther, 2016). In other words, the results concerning which implicit knowledge test is reliable and valid remain inconclusive (hence, a lack thereof) and this may prevent researchers to pursue this research agenda until more rigorous evidence is substantiated.

the one shot WCF treatment they employed in their study. They commented that the "one-shot written error feedback of the type investigated in many studies (including the one we report below) may contribute to explicit knowledge but will have no effect on their implicit knowledge" (p. 288). Given the nature of language learning, whether an L1 or L2, Shintani and R. Ellis (2013) are well aware that time and opportunity to practice producing the output are two key factors that learners need in order to successfully acquire certain features of a language.

Apart from the fact that most WCF studies seem to disregard the potential role of WCF plays in L2 learning, another aspect that is usually overlooked within this field of study concerns the moderating effects of learner difference factors (Bitchener, 2012, 2017; Bitchener & Storch, 2016; Mao & Lee, 2020; Storch, 2018). Apparently, certain learner difference factors could either facilitate or impede learning, given learners' different motivation levels or learning styles for example. Neglecting the potential moderating effects of learner differences when conducting the study could be one of the reasons why there are conflicting results in the field.

Among a wide range of learner difference factors investigated in the SLA field, learner differences in working memory capacity (WMC) has recently been posited to play a significant moderating role in L2 learning. A plethora of oral feedback studies had investigated the relationships between WMC and the efficacy of oral feedback on learners' L2 learning outcomes and found that WMC had moderating effects on the effectiveness of oral feedback and learners' L2 learning outcomes (e.g., Goo, 2012; Kim, Payant, & Pearson, 2015; Li, 2013; Mackey, Philp, Egi, Fujii, & Tatsumi, 2002; Mackey, Adams, Stafford, & Winke, 2010; Mackey & Sachs, 2012; Révész, 2012; Sagarra & Abbuhl, 2013; Sanz, Lin, Lado, Stafford, & Bowden, 2016; Yilmaz & Granena, 2019). Notwithstanding, to the best of my knowledge, until now, only one WCF study (i.e., Li & Roshan, 2019) did empirically examine the moderating effect of WMC on the efficacy of WCF. And even though Li and Roshan (2019) also found the moderating effect of WMC on the efficacy of certain types of WCF, similar to those oral feedback studies, a firm conclusion as to whether this cognitive ability has significant impact on the effectiveness of WCF could not be drawn based solely on their study. More empirical studies under this agenda are undoubtedly needed.

Indeed, Kellogg (1996) and Kellogg, Whiteford, Turner, Cahill and Mertens (2013) have long posited that learners' WMC does have an influential effect on learners' writing processes. Olive (2004) has also suggested that learners with different WMC may benefit from WCF distinctively. Still, this line of inquiry is left unattended. To enrich our understanding about the role of this widely researched cognitive ability in WCF processing and to arrive at a more affirmative conclusion concerning this association, further study under this agenda is essential.

1.2. Focus of the study

As mentioned, there is a wide gap in the literature regarding the potential role of WCF in promoting L2 development. In response to the theoretical and pedagogical issues left unanswered in the field, the current study aimed to investigate both the theoretical question regarding the efficacy of WCF on learners' explicit and implicit knowledge development, and the more pedagogically oriented question regarding which type of WCF is more effective in facilitating learners' L2 development.

Additionally, since there is a possibility that learner differences in WMC may mediate the efficacy of WCF, this study also intended to investigate the mediating effects of learner differences in WMC on the relationships between different types of WCF and learners' L2 learning outcomes.

1.3. Structure of the study

The current study is divided into six main chapters. Chapter 1 presents the background, motivation and the main objectives of the current research.

Chapter 2 introduces a comprehensive synthesis of key theories addressing the roles of negative evidence and WCF in SLA, followed by the explicit-implicit interface proposition, and prior research investigating the effectiveness of WCF on L2 development, respectively. In the subsequent section, the mediating effects of learner differences in WMC in relation to the efficacy of WCF and learners' L2 learning

outcomes are discussed. The chapter concludes with current research gaps presenting in the literature and the study's research questions which guide the analysis.

Chapter 3 starts with an overview of the design of the study, detailed descriptions of participating populations, selected target linguistic structure, treatment tasks and instrumentation employed for the study. The final sections of this chapter provide detailed description of the research procedures and significant operationalized terms used in the study.

Chapter 4 and 5 of the study share identical structures. Both begin with detailed data analysis methods followed by quantitative findings of the study (for Chapter 4: Quantitative Results) or qualitative results deriving from qualitative interview and exit questionnaire data (for Chapter 5: Qualitative Results). The two chapters then conclude with an in-depth discussion of the emerging results and the limitations of the quantitative or qualitative part of the study.

Lastly, Chapter 6 presents the combined analysis of the quantitative and qualitative findings of the current study. The combined discussion is then followed by implications for second language learning theories and pedagogies. The chapter ends with the study's limitations and viable venues for future research.

Chapter II.

LITERATURE REVIEW

This chapter begins with a review of the roles of negative evidence and written corrective feedback (WCF) perceived in the field of second language acquisition (SLA), followed by a discussion of theoretical relevance, and an examination of previous WCF studies presenting in the literature. In the final section of this chapter, the mediating effects of learner differences in working memory capacity (WMC) on L2 learning are reviewed. Relevant WMC work is also explored.

2.1. The roles of negative evidence and corrective feedback in SLA

It is well accepted in SLA that language learners usually have access to two types of input: positive and negative input (Gass, 2003, 2015). "Positive input," or "positive evidence," refers to well-formed linguistic information to which learners are exposed through spoken or written language, which they can use to form linguistic hypotheses of the target language (Gass, 2003, 2015). That is, learners can "make use of the input (among other information) to determine the structure of the second language" (Gass, 2015, p. 187). In contrast, "negative input," or "negative evidence," refers to language input that are purposely used to inform learners about their incorrect use of the target language forms or structures. Negative input can be provided explicitly or implicitly (Gass, 2003, 2015). Notably, negative input is mostly provided through feedback following learners' erroneous linguistic production (Gass, 2015).

The distinctiveness of these two types of evidence, however, raises a controversial question among SLA researchers as to whether L2 learners need access to both types of input to acquire a second language, or access to positive input alone is sufficient, similar to how a child acquires his or her first language mostly through natural exposure to positive input.

Nonetheless, to date, the field consensus has shown that L1 and L2 acquisition processes are not exactly the same phenomenon (Van Beuningen, 2010), and the cognitive processes of first language (L1) and second language (L2) learners do not

overlap as much as previously conjectured (Doughty, 2003, Van Beuningen, 2010). As a result, it is unwarranted to presume that L1 and L2 learners need exactly the same type of language input to successfully acquire the target language. In fact, substantial studies conducted to evaluate the performance of French immersion learners in Canada (see Swain & Lapkin, 1982, 1986, 1998; Swain, 1984, 1985, 1995, 1998, 2005), have shown that while these immersion learners had developed native-like fluency after years of intensive exposure to positive input, they failed to acquire native-like grammatical competence. The studies' findings suggest that positive input alone is not sufficient for acquisition[3] (also see Gass & Mackey, 2015; Long, 1996, 2014; Mackey, Abbuhl & Gass, 2013; Swain, 1995, 2005 for a full review) and that negative evidence in the form of corrective feedback is useful in helping L2 learners improve their overall accuracy (Swain, 1985; Swain & Lapkin, 1998). That is, while positive evidence is undoubtedly necessary, it is not "the only true cause of second language acquisition" (Swain, 2005, p.472) as claimed by most nativist theorists (e.g., Chomsky, 1981; Krashen, 1982, 1985; Schwartz, 1993).

In addition, recent years have seen a growing body of empirical evidence, meta-analyses, and narrative reviews all attesting the effectiveness of negative evidence in the forms of oral and written corrective feedback in promoting L2 development, further emphasizing the potential role of negative evidence in L2 learning processes (e.g., Abbuhl, Ziegler, Mackey, & Amoroso, 2018; Adams, 2003; Adams, Nuevo & Egi, 2011; Biber, Nekrasova, Horn, 2011; Benson & DeKeyser, 2019; Bitchener, 2012; Bitchener & Ferris, 2012; Bitchener & Storch, 2016; Bonilla López, Van Steendam, & Buyse, 2017; Bonilla López, Van Steendam, Speelman & Buyse, 2018; Brown, 2016; R. Ellis, 2006, 2008a, 2012a; Farrokhi & Sattapour, 2012; Ferris, Liu, Sinha & Senna, 2013; Frear & Chiu, 2015, Gass & Mackey, 2015; Goo, 2012; Goo & Mackey, 2013; Guo, 2015; Guo & Yang, 2018; Kang & Han, 2015; Kurzer, 2018; Lee, 2019; Li, 2010; Li & Vuono, 2019; Loewen, 2012; Long, 1996, 2007, 2014; Lyster & Ranta, 2013; Lyster & Saito, 2010; Lyster, Saito & Sato, 2013; Mackey, 2012; Mackey et al., 2002; Mackey & Goo, 2007;

---

[3] "If one starts from the idea that a native-like proficiency on all possible levels, including accuracy, is the ultimate goal of L2 instruction, the conclusion should be that a fully meaning-based approach (i.e., focusing only on comprehensible of positive input) to L2 instruction does not suffice" (Van Beuningen, 2010, p. 4).

Mackey et al., 2013; Mao & Lee, 2020; Nassaji, 2016; Nassaji & Fotos, 2011; Nassaji & Kartchava, 2017; Nemati et al., 2019; Polio, 2012; Rezazadeh et al., 2015; Russell & Spada, 2006; Sagarra & Abbhul, 2013; Sato & Loewen, 2018; Sheen, 2010; Shintani & R. Ellis, 2013; Shintani, R. Ellis & Suzuki, 2014; Shooshtari et al., 2019; Stefanou & Révész, 2015; Van Beuningen, 2010; Van Beuningen, De Jong, & Kuiken, 2012; Wagner & Wulf, 2016). However, despite these findings, the roles of negative evidence and corrective feedback in L2 acquisition remain controversial.

Different groups of theorists have shown different positions towards the roles of negative evidence and feedback play in SLA. These theorists can be largely divided into two main camps: "nature" versus "nurture" (Gass, 2003). The former proposes that SLA is all about learners' innate knowledge about the language and that exposure to positive evidence alone is adequate for acquisition. Specifically, most "nature" advocates postulate that learners' L2 acquisition is either governed by learners' innate devices called Universal Grammar [UG] (Chomsky, 1981) or constrained by particular developmental trajectories (Pienemann, 1998, 2007). According to their viewpoint, positive evidence alone is sufficient for acquisition processes. All other factors, including negative evidence in the form of corrective feedback, are not pertinent to acquisition, and are, therefore, dismissed completely.

Contrary to the "nature" supposition, those arguing the "nurture" standpoint posit that SLA is "inspired and conditioned by the environment" (Gass, 2003, p. 225). In other words, positive evidence alone is not adequate for acquisition and that negative evidence is helpful (R. Ellis, 2006, 2008a, 2012a, 2012b; Gass & Mackey, 2015; Long, 1996, 2007, 2014; Mackey et al., 2013). "Nurture" proponents argue that actual learning or acquisition takes place based on the interaction of multi factors and negative evidence certainly has its role in acquisition (R. Ellis, 2007; Gass, 2003, 2015; Gass & Mackey, 2015; Long, 1996, 2007, 2014, Mackey et al., 2013).

Since the current study aimed to investigate the effectiveness of WCF (one form of negative evidence) on learners' L2 development (i.e., explicit and implicit knowledge development), it is, therefore, crucial to explore different theoretical views towards the roles of negative evidence and corrective feedback within the field of SLA, so that we can

better understand where WCF stands in the current literature. In the subsequent sections, two aforementioned theoretical camps: the nature and nurture suppositions are further discussed in more detail.

2.1.1. Theoretical views in support of the roles of negative evidence and corrective feedback in second language learning

While the nature proposition posits that mere access to positive evidence is sufficient for acquisition and negative evidence is irrelevant, the nurture proposition, i.e., the proponents of the four key theoretical underpinnings: the interaction, skill acquisition, sociocultural and usage-based approaches, perceive that negative evidence in the form of corrective feedback is essential especially for L2 acquisition. In the subsequent sections, theoretical foundations in support of the roles of negative evidence and corrective feedback in L2 acquisition are discussed in more detail.

2.1.1.1. Interaction approach

The interaction approach to L2 learning emphasizes the roles of input, output, and corrective feedback during L2 interaction, i.e., negotiation for meaning (Gass & Mackey, 2015; Long, 1996, 2014). Attention to linguistic forms is also one of the key learning mechanisms within this approach. The interactionist believes that learners have to pay attention to forms during the negotiation for meaning so that they can internalize the forms (Gass & Mackey, 2015; Long, 1996, 2014; Schmidt, 1990, 1995, 2001).

As well, learners have to attend to forms when they produce the output so that they can notice if there are gaps between their output and the target language forms (Swain, 1985, 1991, 1995, 1998 2005). Negative evidence in the form of corrective feedback is, thus, essential as it serves as a pedagogical tool inducing learners' attention to problematic forms, both in the input and the output (Gass & Mackey, 2015). Long (1996, also 2014) elucidates the tenets of interaction approach as:

> Environmental contributions to acquisition are mediated by selective attention and
> the learner's developing L2 processing capacity, and ... these resources are

12

brought together most usefully, although not exclusively, during negotiation for meaning. Negative feedback obtained during negotiation work or elsewhere may be facilitative of L2 development, at least for vocabulary, morphology, and language-specific syntax, and essential for learning certain specifiable L1-L2 contrasts. (p. 414)

The significant role attention plays in acquisition processes highlighted within this interactionist framework is also buttressed by Schmidt's Noticing Hypothesis (1990, 1995, 2001) which proposes that conscious attention to linguistic forms is a precondition to learning, since "people learn about things they attend to and do not learn much about the things they do not attend to (Schmidt, 2001, p. 30) and only through conscious attention that input can be internalized to be an intake.

Schmidt further argues that conscious attention (i.e., noticing at the attention level) is significant because it makes learners aware of the input, i.e., target structures, as well as notice a mismatch between what the learners can produce and what the native speakers of the language actually produce. Feedback which works as a consciousness raising or form-focused device can help learners consciously aware of or notice the mismatch between their interlanguage output and the target-like input, prompting the destabilization and reconstruction of learners' interlanguage grammar (Bitchener & Storch, 2016; Gass, 1997, 2003; Long, 1996, 2007, 2014; Polio, 2012; Van Beuningen, 2010, Williams, 2012). Since conscious attention to forms is facilitative to acquisition as such, feedback which induces noticing and noticing the gap (i.e., inducing attention to forms) can be viewed as a facilitative tool to acquisition (Bitchener & Storch, 2016; Brown, 2016; DeKeyser, 2007; R. Ellis, 2006; Gass & Mackey, 2015; Goo & Mackey, 2013; Long, 1996, 2007, 2014; Lyster & Mori, 2006; Lyster, Saito & Sato, 2013; Mackey, 2012; Mackey et al., 2013; Sheen, 2010). A large body of research, meta-analyses and narrative reviews investigating the effectiveness of oral feedback within this framework has affirmed the beneficial effect of oral feedback in promoting L2 development (e.g., Abbuhl et al., 2018; Adams et al., 2011; Brown, 2016; Goo & Mackey, 2013; Li, 2010; Li & Vuono, 2019; Loewen, 2012; Loewen & Philp, 2006; Lyster & Ranta, 2013; Lyster & Saito, 2010; Lyster, Saito & Sato, 2013; Russell & Spada, 2006; Mackey, 2012; Mackey & Philip, 1998; Mackey & Goo, 2007;

McDonough, 2007; McDonough & Mackey, 2006; Nassaji, 2016; Nassaji & Kartchava, 2017; Saito & Lyster, 2012; Sato & Loewen, 2018; Trifimovich, Amnar & Gatbonton, 2007).

Even though there are some fundamental differences between the nature of oral and written feedback, given that WCF is not usually provided at the exact moment of the negotiation for meaning unlike oral feedback, for instance, many researchers (e.g., Bitchener, 2012; Bitchener & Storch, 2016; Kuiken & Vedder, 2011; Polio, 2012; Williams, 2012) contend that there are theoretical reasons to believe that WCF can be beneficial for L2 development similar to oral feedback. They propose, for example, that learners who receive WCF certainly have more time and enough attention resources to notice and compare the mismatch between their own interlanguage output and the target language input provided by WCF, than oral feedback (which is fleeting by nature); thereby, increasing learners' chances of noticing the gaps in their interlanguage leading to greater opportunity for hypothesis testing and modified output (Bitchener, 2012; Bitchener & Storch, 2016; Kuiken & Vedder, 2011; Williams, 2012). It is as Williams (2012) points out that learners' "cognitive window is open somewhat wider and that learners have a richer opportunity to test their hypotheses when they write [or receive WCF] than when they speak" (p. 328).

Recently, more empirical work has been conducted within the interactionist framework applying the concept to writing practice (e.g., Adams, 2003; Coyle & Roca de Larios, 2014, 2020; Diab, 2015; Hanaoka & Izumi, 2012; Lapkin, Smith, & Swain, 2002; Qi & Lapkin, 2001; Sachs & Polio, 2007; Santos, López, & Manchón, 2010; Shintani, 2016; Swain & Lapkin, 2002; Yang & Zhang, 2010). These studies have found that WCF effectively promoted noticing leading to improvement in accuracy and writing performance similar to oral feedback. Even though all of these studies did not address WCF's long-term effect (i.e., effect on acquisition), only the short-term ones, for most interactionists, short-term improvement can be regarded as impactful (Norris & Ortega, 2003; Mackey & Polio, 2009).

In sum, within the interactionist framework, it is uncontroversial that negative evidence or corrective feedback certainly plays a significant role in acquisition processes (Gass & Mackey, 2015; Long, 1996, 2007, 2014; Mackey et al., 2013), even though it

14

remains unresolved whether WCF has the same facilitating effect on learners' acquisition (i.e., implicit knowledge) the same way oral feedback does.

2.1.1.2. Skill acquisition theory

Skill acquisition theory is a cognitive based theory accounting for how people proceed to acquire a variety of skills, from the very initial controlled stage to the end state where people can eventually execute their skills automatically (DeKeyser, 2007, 2015). In terms of second language learning, skill acquisition theorists posit that learning a language is indistinguishable from learning other skills in the way that language learners will progress through similar stages of development: from first acquiring declarative knowledge (i.e., explicit knowledge[4]), a knowledge about the skill, to finally attaining procedural knowledge (i.e., automatized procedural knowledge) in which the knowledge assists learners to execute the acquired skill faster, effortlessly, with minimal attention and fewer errors (Anderson, 1981, 1982, 1985, 1993, 2007). Generally, there are three stages of skill development within this theory framework: declarative, procedural and automatic stages (Anderson, 1981, 1982, 1985, 1993, 2007; DeKeyser, 2007, 2015).

Declarative stage is the stage where learners encode new knowledge required for skill performance to their memory (Anderson, 1981, 1982, 1985, 1993, 2007) or "acquire quite a bit of knowledge about the skill (DeKeyser, 2015, p. 95). At this stage, even though new knowledge is already encoded in learners' memory, their knowledge execution is still very slow as learners still need time to retrieve pieces of new information from their memory and assemble them together to perform the skill. Through practice, learners proceed to *procedural stage* where learners start using the knowledge they obtain with faster execution given that the new knowledge has already been established in the declarative stage and "it no longer requires the individual to retrieve bits and pieces of information from memory to assemble them into a program for a specific behavior; instead, that program is now available as a ready-made chunk (as a result of production compilation, see also Anderson, 2007)" (DeKeyser, 2015, p. 95).

---

[4] Declarative and procedural knowledge is comparable to explicit knowledge, while automatized procedural knowledge is comparable to implicit knowledge.

However, procedural knowledge acquired at the procedural stage is still not yet robust or fine-tuned and therefore learners need further intensive practice so that they can progress to the last stage which is the automatic stage where learners can use the knowledge they gain automatically, effortlessly with fewer errors (DeKeyser, 2007, 2015). Faster execution with incremental accuracy (i.e., fewer errors) is evidence of learners' progress towards automaticity (i.e., implicit or automatized procedural knowledge) (DeKeyser, 2007, 2015). However, it should be emphasized here that "Automaticity is not an all-or-nothing affair; even highly automatized behaviors are not 100% automatic, as becomes clear when we stumble walking down the stairs…or when we stumble over our words while uttering a simple sentence in our native language" (Dekeyser, 2015, p. 96).

In addition to the three stages of knowledge development, practice is a key mechanism within this approach as it assists learners to move from declarative to automatic stages. It should be noted that within this framework, declarative knowledge has a significant role in learners' acquisition as it is "a major avenue for the acquisition of procedural knowledge" (Anderson & Fincham, 1994, p. 1323). In other words, through practice, declarative knowledge can be converted into implicit or automatized procedural knowledge. Negative evidence in the form of corrective feedback is also viewed as an indispensable tool helping learners notice the gaps between the target language and their current interlanguage during declarative stage, so that they do not proceduralize inaccurate language input into their repertoire.

Within this approach, it is vital that erroneous language constructions are not proceduralized because once they are proceduralized, it is difficult to modify those pieces of knowledge (Anderson, 1981, 1982, 1985, 1993, 2007) since they are already beyond learners' conscious control and this may result in learners' language fossilization. A number of WCF studies (i.e., Hartshorn & Evans, 2012, 2015; Hartshorn, Evans, Merrill, Sudweeks, Strong-Krause, & Anderson, 2010; Evans, Hartshorn, Strong-Krause, 2011; Kurzer, 2018; Wagner & Wulf, 2016) have applied the tenets of skill acquisition theory to writing and found that the provision of WCF led to significant improvement in written accuracy affirming the effectiveness of WCF in facilitating learners' L2 development.

All in all, it could be concluded that within skill acquisition theory, negative evidence in the form of corrective feedback plays a crucial role in L2 development, as it

helps prevent wrong information to be proceduralized into learners' memory during declarative stage, and this facilitates the successful development of learners' automatized procedural knowledge.

2.1.1.3. Sociocultural theory

The sociocultural theory is originally developed based on the work of Vygotsky (1978, 1981) proposing that human cognitive development occurs through interaction with social or material environments (Engeström, 1987), or through "participation in cultural, linguistics and historically formed settings such as family life, peer group interaction, and institutional contexts like schooling" (Lantolf, Thorne, & Poehner, 2015, p. 207). Within this approach, the Zone of Proximal Development (ZPD) is the key construct of the theory that is most pertinent to L2 learning and the roles of negative evidence and feedback in particular (Storch, 2018). That is, the sociocultural approach to L2 learning proposes that learning can take place through social interaction whereby an expert member of the society (e.g., teachers, more knowledgeable peers) provides negotiated and dynamic assistance (i.e., scaffolding) to a novice member (i.e., learners, inexperienced peers) of the society (Storch, 2018). Yet, learning can actually take place only when negotiated and dynamic assistance or scaffolding is provided within learners' ZPD, the learners' current and potential level of competence. Through scaffolded assistance provided by an expert within learners' ZPD, learners can progress their learning until they are self-regulated or able to perform their learning activities independently.

Within this approach, negative evidence or corrective feedback can be used as dynamic scaffolding responding to individual learners' needs; thereby, helping learners become more independent (i.e., self-regulated in sociocultural term) in their learning (Storch, 2018). This is as Aljaafreh and Lantolf (1994) comment that "all types of feedback are potentially relevant for learning, but their relevance depends on where in the learner's ZPD" (p. 480). In term of learning development, less reliance on feedback over time (i.e., learners require less assistance on error correction/ they can self-correct their

errors) is taken as evidence of progression within this framework (see Lantolf & Poehner, 2014).

To date, there are a number of studies (e.g., Aljaafreh & Lantolf, 1994; Brooks & Swain, 2009; Erlam, R. Ellis & Batstone, 2013; Kim, Choi, Kang, Kim & Yun, 2020; Nassaji & Swain, 2000; Storch & Wigglesworth, 2010) investigated the effectiveness of oral and written corrective feedback employing this theory's tenets and reported that scaffolded feedback was more effective in helping learners improve their L2 knowledge than non-scaffolded feedback (i.e., "random feedback that was not graduated and contingently responsive", Storch, 2018, p. 266). Nonetheless, even though feedback can be used as a dynamic scaffolded tool assisting learners' L2 development as demonstrated, Polio (2012) has pointed out that the question remains whether non-scaffolded feedback can actually be effective given that all the feedback provided in the aforementioned studies were individualized and responsive to individual learners' ZPD.

2.1.1.4. Usage-based theory

According to Wulff and N. Ellis (2018), usage-based theory encompasses various L2 learning approaches that share the same two main assumptions: 1) cognitive mechanisms underlying language learning are domain general, that is they are not exclusive to language learning but all kinds of learning, and 2) linguistic input is the primary source for learning. Language learning, under this framework, associates with construction learning (N. Ellis, 2015a, 2015b). A construction is a basic unit of language representation. It is a form-function pairing, ranging from morphemes, to words, phrases, and syntax (Wulff & N. Ellis, 2018). Construction learning is, thus, the learning of the form-function connection (N. Ellis, 2015a, 2015b).

"Frequency of usage is the driving force of construction learning" (Wulff & N. Ellis, 2018, p. 42). The more frequently a construction is experienced, the more likely that that particular construction is earlier acquired comparing to the less frequently one (N. Ellis, 2012, 2015a, 2015b; Wulff & N. Ellis, 2018). That is, learners have higher tendency to acquire a high frequently linguistic feature (i.e., a construction) they encounter in the input than the less frequently one they rarely find in the input. This is because the more

often learners are exposed to the same linguistic features, the greater chance they have to analyze and extract the associations "within and between" constructions from the exposed input (i.e., associative learning), and over time, after collecting enough exemplars, learners are able to create a network of language knowledge (Wulff & N. Ellis, 2018). In other words, a theory of language learning requires an understanding of the associative learning of representations that reflects the probability of the occurrences of form–function mappings. Learners have to figure the language out: their task is, in essence, to learn the probability of an interpretation by looking at formal cues given in a particular context (N. Ellis, 2008, p. 15).

Consequently, within this approach, apart from input, which is conceived as primary source for learning, frequency of usage which facilitates associative learning is one of the key determinants of L2 acquisition. Despite that, the learnability of each construction can be affected by three main factors, that is 1) the saliency of the target form (i.e., how a particular form stands out of its context), 2) the contingency of form-function association (i.e., the reliability of the form-function mapping), and 3) learned attention (N. Ellis, 2012, 2015a, 2015b, 2016a, 2016b; Wulff & N. Ellis, 2018).

With regard to saliency, it is proposed that salient forms are more likely to be attended to and registered into learners' learning systems, while the less salient cues tend to be disregarded and therefore making it more difficult to acquire. Second, the high and low contingency of the form-function association determine the extent to which the target construction can be successfully processed and acquired (Gries & N. Ellis, 2015; Wulff & N. Ellis, 2018). That is, a learnable construction (i.e., a high contingency form-function mapping) is the one in which the form (i.e., cue) can be reliably used to predict an interpretation (i.e., outcome), whereas "Cues with multiple interpretations [i.e., low contingency form-function mappings] are ambiguous and so hard to resolve" (Wulff & N. Ellis, 2018, p. 46) posing a challenge for acquisition. The last influential factor affecting the learnability of the target forms associates with learners' learned attention.

Within this approach, it is believed that L2 learners' attention can be blocked or interfered by their prior L1 knowledge (N. Ellis, 2012, 2015a, 2015b, 2016a, 2016b; Wulff & N. Ellis, 2018). That is, learners' L1 can bias or block learners' learned attention towards associative learning in a new language. Given that learners' learning mechanisms

are already tuned to L1 cues, learners then naturally incline to rely on their L1 cues even when learning an L2; thus, disregarding some L2 cues especially those that are low in salience or contingency (N. Ellis, 2015a, 2015b; Wulff & N. Ellis, 2018). It is as Wulff and N. Ellis (2018) point out that "Since everything is filtered through the lens of the L1, not all of the relevant input is in fact taken advantage of" (p. 50). Due to selective attention among L2 learners or blocking phenomenon resulting from L1 interference, N. Ellis (2005, 2008, 2015a, 2015b, 2016a) conjectures that noticing (i.e., attention) which leads to conscious processing (i.e., explicit learning) and form-focused instruction (FFI) including form-focused feedback are necessitated for L2 acquisition.

> The limitations of second-language learning show us that this tallying is by no means guaranteed. Effects of salience and learned attention entail that, broadly, it is not until a representation has been noticed and consolidated, that the strength of that representation can thereafter be tuned implicitly during subsequent processing…There is thus a strong role for consciousness in language learning.
> (N. Ellis, 2016a, p. 247)

N. Ellis postulates that FFI and feedback can effectively induce learners' noticing and raise learners' metalinguistic awareness to the less salient and redundant linguistic forms blocked by learners' L1 cues, allowing successful consolidation of form-function mapping of new constructions. "Once a construction has been represented in this way, its use in subsequent implicit processing can update the statistical tallying of its frequency of usage and probabilities of form-function mapping" (Wulff & N. Ellis, 2018, p. 51).

In Cintrón-Valentín and N. Ellis' (2015) and Cintrón-Valentín and N. Ellis' (2016) studies which were conducted within this framework to investigate whether different types of FFI can subdue L2 learners' selective attention and blocking presented in L2 learners' online processing of L2 input, the results demonstrated that learners who received FFI treatment became more sensitive to morphological cues in both comprehension and production tasks. Further, explicit focus on form (FonF) treatment condition seemed to efficiently modulate learners' long-term blocking of verb morphology.

To date, to the best of my knowledge, there are no studies conducted within this framework directly investigating the effectiveness of oral or written feedback on learners'

selective attention similar to the two studies exemplified above. Nonetheless, it can be implicated that feedback, a form-focused device, can be equally effective as other FFI interventions, providing that the findings from a plethora of oral and written feedback studies have demonstrated that learners who received feedback generally displayed more noticing and produced more uptake episodes than those who did not (e.g., Bao, Egi, & Han, 2011; Ca Novas Guirao, Roca de Larios, & Coyle, 2015; Coyle & Roca de Larios, 2014, 2020; Diab, 2015; Hanaoka & Izumi, 2012; Kartchava & Ammar, 2014; Kim & Bowles, 2019; Loewen & Philp, 2006; Mackey, 2006; Sachs & Polio, 2007; Santos et al., 2010; Sheen, 2010; Shintani, 2016; Shintani & R. Ellis, 2013; Simard, Guénette, & Bergeron, 2015). These findings suggest that feedback can efficiently draw learners' attention to problematic forms, inducing more "noticing the gap" (i.e., noticing the mismatch between learners' L1 and L2) which over time may potentially subdue L1 blocking phenomenon, allowing successful consolidation of the form-function mappings.

In sum, the low saliency and contingency (i.e., reliability) of the target form, and the learned attention bias due to L1 influence make it more difficult for L2 learners to acquire the target form despite its frequency in the input. FFI and corrective feedback are useful in this light as they promote noticing and help raise L2 learners' conscious awareness towards the less perceptual salient or low contingency construction, making the construction become more salient while subduing learners' L1 influence. Consequently, within the usage-based approach, feedback is beneficial as it facilitates the learning of the low salient and low contingency forms and the forms blocked by learners' L1, attributing to learners' overall L2 acquisition.

2.1.2. Theoretical views in opposition to the roles of negative evidence and corrective feedback in second language learning

Generative and processability theories are the two schools of thought advocating the nature supposition and therefore barely acknowledge the roles of negative evidence and corrective feedback play in SLA. In the following sections, details regarding each of these theories' views towards negative input and corrective feedback in L2 development are summarized and discussed.

2.1.2.1. Generative theory

Generativists or proponents of generative approach are most interested in how children acquire their first languages. Its aim is "to provide a characterization of the linguistic competence of native speakers of a language and to explain how it is possible for child first language (L1) acquirers to achieve that competence" (White, 2015, p.34). The generativist view on L2 acquisition is similar, to understand the acquisition of L2 learners' interlanguage competence. Initially, generativist theory was developed to offer partial explanation about "a mismatch between the input that a child is exposed to (the primary linguistic data) and the complex knowledge the child acquires from it (the grammar) (White, 2018, p.1), as Chomsky (1987, as cited in Cook, 1988) puts it " How do we come to have such rich and specific knowledge, or such intricate systems of belief and understanding, when the evidence available to us is so meagre?" (p. 82).

Introduced by Chomsky (1981), the fundamental principle of the generative approach is that all children acquire their first languages through their innate devices called "universal grammar" (UG), which functions as universal grammar constraints. The UG constraints limit learners' "operation of linguistic rules" (White, 2015, p.38), with the parameters situated within the UG directly constrain grammar variations across languages. It is postulated that once these parameters are set through the exposure of positive evidence of a particular language, a child will be able to generate grammatically correct sentences in that language. Thus, within this approach, positive evidence serves as UG parameter trigger helping learners acquire different linguistic features of the target

language. Negative input, on the other hand, does not have any particular roles within this theory since UG theorists claim that "language is acquired on the basis of primary linguistic data [positive evidence] interacting with [or "triggering"] the principles and parameters of UG" and that "L1 acquirers do not have access to negative evidence" (White, 2018, p. 5).

As aforementioned, since the generative theory is originally developed from L1 acquisition theory, most generativists assume that what is necessary for L1 acquisition is also essential for L2 acquisition, and what is not necessary for L1, therefore, is not necessitated for L2. Although there is a subset of UG theory proposing that the provision of negative evidence is necessary, especially when the positive evidence is not adequately informative (White, 2003, 2018, also see Flynn, 1996; Schwarts & Sprouse,1996; Trahey & White, 1993), the roles of negative evidence and corrective feedback within this approach are still largely disregarded (Doughty, 2003, Polio, 2012). In essence, generativists believe that only positive evidence is necessary and adequate for acquisition (i.e., the development of implicit knowledge). Other factors like negative evidence and corrective feedback are not relevant and may only help with learning (i.e., the development of explicit knowledge) but will not contribute to acquisition.

In Truscott's (1996) most controversial article in which he fervently claimed that WCF is useless and should be abandoned as it only serves "pseudolearning" (i.e., explicit knowledge), not acquisition (i.e., implicit knowledge), Truscott grounded his argument on the generativist theory. However, Truscott's allegation towards the ineffectiveness of WCF based on the generative theory appears to be unwarranted, since even until now it is still debatable as to whether UG is fully accessible to L2 learners or completely constrains learners' L2 grammar (Tarone, 2014, Gass, 2003, White, 2003). Further, many SLA researchers (e.g., Abrahamsson & Hyltenstam, 2009; Bley-Vroman, 1989, 1990; Clahsen, Felser, Neubauer, Sato & Silva, 2010; DeKeyser, 2000, 2003; Doughty, 2003; Doughty & Williams, 1998; N. Ellis, 2015b, 2016a, 2016b; Long, 2007; Meisel, 2011; Morgan-Short & Ullman, 2012; Paradis, 2009; Schachter, 1991; Selinger, 1972; Selinger & Lakshmanan, 1992; Ullman, 2001, 2004, 2015, 2016; Wulff & N. Ellis, 2018) do not concur that this theory can entirely apply to learners' L2 acquisition system, since it is

more likely that L1 and L2 learning systems/mechanisms are rather divergent than identical.

Doughty and Williams (1998), for instance, have argued that "second language learning is not identical to first language learning, and so we do not consider leaving learners to their own devices to be the best plan" (p. 197). Accordingly, it may be precipitate to follow the generativist's claim and set L1 learning conditions as a default model for L2 learning, ruling out other potential factors such as negative evidence in the form of corrective feedback that might help L2 learners acquire the target language faster. Of importance here is the fact that it is still debatable in the field whether explicit knowledge plays any crucial roles in the development of implicit knowledge. Thus, Truscott's claim stating that WCF could only contribute to the development of explicit but not implicit knowledge seems to be fallacious for the time being, at least until there is robust evidence ratifying that the two types of knowledge do not interact.

Also, as mentioned in previous section, prior studies investigating the effectiveness of French immersion programs in Canada have shown that a mere provision of positive input in communicative-oriented immersion settings (i.e., meaningful contexts) was not sufficient to help L2 learners develop the near native like proficiency since these immersion learners could only acquire target-like fluency yet failed to reach target-like accuracy (Swain, 1985, 1995, 1998, 2005). Swain (1998), thus, suggests that in this light negative evidence in the form of corrective feedback is beneficial because it can effectively help learners notice the mismatch between the target-like form and their actual output (i.e., noticing the gap). The noticing induced by feedback then increases learners' chance to modify, reprocess and reproduce their output and that helps advancing learners' accuracy.

In addition, some other researchers (e.g., Skehan & Foster, 2001; Van Beuningen, 2010) postulate that disregarding L2 learners' accuracy could eventually lead learners to premature fossilization as they continuously proceduralize non-target like features into their language repertoire. However, providing learners with corrective feedback helps decrease such premature fossilization since corrective feedback could draw learners' attention to accurate forms preventing them to proceduralize the non-target like ones. In sum, until there is adequate evidence to prove that UG fully constrains learners' L2

grammar the same way it does with L1 learners, Truscott's claim about the ineffectiveness of WCF based on generative theory is unwarranted and should be reevaluated judiciously for the time being.

2.1.2.2. Processability theory

Similar to generative theory's view, Pienemann's (1998, 2007; Pienemann & Lenzing, 2015) processability theory is another theory that barely acknowledges the roles of negative evidence and corrective feedback in SLA (Bitchener & Ferris, 2012; Polio, 2012). Processability theory is built upon the premise that language learners' acquisition processes are constrained by "developmental trajectories", developmental sequences, or stages of development, which all language learners, both L1 and L2, have to follow through (Pienemann, 1998, 2007; Pienemann & Lenzing, 2015). Within this framework, linguistic structures will be acquired in a predictable hierarchical order and only when the features of the previous stage have been attained, learners could then acquire later stage features (Pienemann, 1998, 2007; Pienemann & Lenzing, 2015).

According to Pienemann (2007), "every procedure is a necessary prerequisite for the next procedure…Therefore, the learner has no choice but to develop along this hierarchy" (p. 141). Consequently, within this approach, external factors do not seem to have much influence on learners' acquisition since the effects of external factors will be constrained by learners' developmental sequences. That is, nothing can help learners acquire target linguistic features unless they are at their developmental readiness (i.e., at the stage where they are ready to acquire those particular features). Negative input or corrective feedback, thus, appears to have no role within this theory, even though some researchers (e.g., Ammar, 2008; Bitchener & Ferris, 2012; Goo & Mackey, 2013; Mackey & Philp, 1998; McDonough & Mackey, 2006; Polio, 2012) posit that some types of corrective feedback could speed up learners' rate of acquisition in each stage of development.

Mackey and Philp's (1998), and McDonough and Mackey's (2006) studies on the effectiveness of implicit oral feedback in the form of "recasts" on L2 development, for instance, have found that the provision of intensive recasts during negotiated interactions

increased learners' uses of developmentally more advanced structures. The result, thus, suggests that corrective feedback provided during conversational interactions has facilitative effect on learners' developmental speed, as it could help L2 learners move towards their next developmental stages faster under the condition that the feedback must be consistent with the learners' developmental levels.

The findings seem to align with Pienemann (1998, 2007; Pienemann & Lenzing, 2015)'s proposition arguing that grammar instruction can be effective for learners only when it is provided at the right developmental stage where learners themselves are ready to acquire those more advanced features naturally. Even though the two studies (i.e., Mackey & Philp, 1998; McDonough & Mackey, 2006) did not investigate the effectiveness of grammar instruction on learners' processing constraints, still it can be inferred that if the right L2 learning method or tool is provided to learners at their right stages of development, there is a chance that that particular method or pedagogical tool can accelerate learners' rate of acquisition enabling them to acquire more advanced structures faster. Extrapolating Pienemann's statement, Bitchener and Ferris (2012) assert that this condition is also applicable to WCF: WCF can be effective when it is provided to learners at their right developmental stages or when they are developmentally ready to acquire certain linguistic features appearing in their corrections.

In Truscott's (1996) argument against the effectiveness of WCF, besides grounding his claim on generativist framework as aforementioned, Truscott also applied processability theory to fortify his stance, arguing that because all language learners have to go through these rigid developmental trajectories, and nothing including feedback can alter this acquisition route unless it can be provided exactly at learners' right developmental stages, the use of WCF is, therefore, futile. However, Truscott's claim must be treated with caution for several reasons. First of all, it should be taken into account that because most processability-based findings are drawn from "learners who had massive exposure to the language and/or were young learners, which means that they were largely implicit learners" (DeKeyser, 2015, p. 102), not from instructed learners with limited exposure to the target language, the theory may not be fully applicable to instructed L2 learners elsewhere. That is, it is highly probable that the developmental trajectories of the implicit learners may be discrepant from those of instructed L2 learners

(e.g., the two groups may have different order of developmental sequences, if there is at all for the instructed learners) and therefore the two groups of learners may require different learning mechanisms for successful acquisition. In line with Dekeyser (2015), R. Ellis (2012a) comments that:

> The natural route of acquisition may only become evident in data that reflect learners' implicit knowledge of the L2. However, it seems perfectly possible for learners to develop explicit knowledge of grammatical features in any order and, if one accepts DeKeyser's arguments, such knowledge [explicit knowledge] can be proceduralized [become implicit knowledge]..Thus, FFI [including the provision of corrective feedback] may be powerless to change the course of acquisition of implicit knowledge but it can still be effective in developing functional control of specific features irrespective of learners' developmental level. (p. 280)

Further, it should be noted that until recently the advocate of the processability approach:

> never found an ordering for all or even most structures in the language [not including the fact that developmental stages of different languages can be distinctively discrepant from one another]; only for a few morphemes in some studies or for a few closely related syntactic patterns in others. (DeKeyser, 2015, p. 102);

therefore, the knowledge concerning developmental sequences seems to be quite limited for theoretical generalization and pedagogical application (DeKeyser, 2015; Van Beuningen, 2010). Thus, it is precipitate to presume at this stage that the acquisition route exists for all linguistic structures. If the developmental sequences do not exist for all linguistic features, it is then counterproductive to disregard the possibility that WCF could potentially contribute to learners' acquisition.

In addition, considering that to date there is still no syllabus developed based on this processability theory to be used in practice because the complete developmental sequences are not fully developed, it does not seem to do justice to hastily conclude that WCF (among other key factors) is ineffective and should be abandoned from practice. In fact, even if WCF could not alter the acquisition courses per se, it might speed up

27

learners' rate of acquisition (Polio, 2012), as effectively as demonstrated in Mackey and Philp's (1998), and McDonough and Mackey's (2006) oral feedback studies.

Another point worth emphasizing is the fact that Pienemann' s processability theory is developed to describe the constraining effect of developmental trajectories on learners' language acquisition, not to propose that these constraints are sufficient conditions for acquisition to take place. Pienemann (1998) elucidates this misunderstanding about his theory stating that he proposes "a constraining effect of linguistic information processing of the genesis of linguistic knowledge, [but he does not] imply that processing constraints describe *sufficient* conditions for the genesis of linguistic knowledge" (p. 2). Based on his explanation, it is conceivable that developmental stages cannot be skipped, and nothing can help learners move from one stage to another when they are not developmental ready, yet, the processing constraint, in itself, is not sufficient condition for acquisition to take place. Other factors are still necessitated for successful acquisition under the condition that they are provided at the right developmental stage. In sum, since there is evidence from oral feedback studies showing that oral feedback can help accelerate learners' rate of acquisition in each developmental stage, the same could possibly hold true for WCF and therefore the roles of negative evidence and particularly WCF in L2 learning should not be disregarded based solely on this processability ground.

2.1.3. Negative evidence, corrective feedback and the interface

One of the most controversial issues in SLA concerns the distinct relationship between explicit and implicit knowledge. Although most scholars agree that implicit knowledge is the goal for L2 learning, the field has not yet arrived at a consensus on how implicit knowledge is developed or whether explicit knowledge helps with the development of implicit knowledge. Different theories have proposed different views on how implicit knowledge can be fostered and of significant, how these two types of knowledge interact, as briefly elucidated in previous section.

Within Krashen's (1981) Monitor theory, learners' implicit knowledge is developed naturally outside their awareness through L2 interactions where their focuses are on meaningful input. Conversely, explicit knowledge is developed consciously through L2 instruction where learners' attention is mostly drawn to focus on forms. Since explicit and implicit knowledge are the byproducts of disparate processes, it is postulated within this framework that the two function distinctively and cannot interact; that is explicit knowledge can never turn into implicit knowledge or even aid the development of implicit knowledge.

In line with Krashen's non-interface position, most generativists agree that explicit knowledge can never become implicit knowledge. They also agree on the same premise that since the two types of knowledge are acquired differently (through different routes), these two types of knowledge can never interact. Both theories strongly support the non-interface position, despite having dissimilar presumptions on how implicit knowledge is acquired, because for generativists, implicit knowledge is developed unconsciously only by means of an implicit mechanism (i.e., Universal Grammar [UG]), not chiefly from interactions with meaningful input, whereas explicit knowledge is acquired through conscious memorization of rules or patterns in the input. Since the non-interface position does not believe that explicit knowledge can actually become implicit knowledge or aid the development of implicit knowledge, they deny the role of negative evidence in which they posit that can only promote the development of explicit knowledge.

In sharp contrast to the aforementioned non-interface position, other theories have stipulated that indeed there is the interface between explicit and implicit knowledge and explicit knowledge does play an active role in the development of implicit knowledge (henceforth the interface position), given the usage-based and skill acquisition theories for instance.

For usage-based theory, even though its proponents perceive that "Most knowledge is tacit knowledge; most learning is implicit; the vast majority of our cognitive processing is unconscious" (N. Ellis & Wulff, 2015, p.89), that is, implicit knowledge is developed implicitly and unconsciously for the most part, they do not concur that it is the only learning condition leading to successful development of learners' L2 knowledge. N. Ellis and Wulff (2018) have pointed out that this is because some aspects of L2 are not learnable or might take a long period of time to acquire through implicit learning alone, given some linguistic forms that lack perceptual salience and consequently usually go unnoticed and cannot be encoded into learners' implicit knowledge repertoire for example. In such case, N. Ellis and Wulff (2018) believe that explicit learning can help inducing learners' attention to the unnoticed grammatical features and this unquestionably helps with the development of learners' overall L2 knowledge. N. Ellis (2005) has elucidated the interface process of these two types of knowledge as

> …happens transiently during conscious processing, but the influence upon implicit cognitions endures thereafter. Explicit memories can guide the conscious building of novel linguistic utterances through process of analogy. Patterned practice and declarative pedagogical grammar rules both contribute to the conscious creation of utterances whose subsequent usage promotes implicit learning and proceduralization. (p.305)

In sum, the advocates of usage-base theory do believe that explicit knowledge helps with the development of implicit knowledge by introducing either non-salient forms or new grammatical knowledge into implicit learning process, which subsequently updates the new knowledge into statistical tallying of its frequency usage. Through frequent exposure, the new knowledge is bestowed into learners' implicit knowledge repertoire. The same can be said about skill acquisition theory providing that the theory also strongly supports the interface position. Within skill acquisition theory, "existing

declarative knowledge [i.e., explicit knowledge], via practice, plays a causal role in the development of procedural knowledge [i.e., implicit or automatized procedural knowledge]" (Dekeyser, 2015, p. 103). Even though the theory may focus more on conscious processing and deliberate learning, while a usage-based approach centers its attention more on statistical and implicit learning, both support the same claim that explicit knowledge does aid the development of implicit knowledge in some way. Providing that the interface position speculates that explicit knowledge can help with the development of implicit knowledge, it is within this ground that negative evidence and corrective feedback are favored and valued.

Albeit strong claims from both camps regarding the existing nature of the interface, empirical findings from each end are until now incongruent. Thus, whether negative evidence or corrective feedback could only contribute to the development of explicit knowledge or both remains an open empirical question. Further investigation is inevitably needed to shed more light on this issue.

2.1.4. Summary

As discussed in the preceding sections, while the four main approaches to L2 learning, namely, the interaction, skill acquisition, sociocultural and usage-based theories perceive that negative evidence and corrective feedback are important for SLA, the generative and processability theories believe that negative evidence and corrective feedback are irrelevant especially when acquisition is the main concern.

Precisely, while the naturalistic approach (i.e., generative and processability theories) believes that L2 acquisition can be succeeded without the provision of negative evidence or corrective feedback, similar to the way L1 learners acquire their native languages, the nurture proposition considers negative evidence in the form of corrective feedback essential for L2 development. For the nurture advocates, positive evidence alone may be sufficient to develop target-like fluency, yet it is not adequate for the development of target-like accuracy. Their hypothesis is substantiated by French immersion studies (see Swain & Lapkin, 1982, 1986, 1998; Swain, 1984, 1985, 1995, 1998, 2005) which demonstrated that despite several years of exposure to comprehensible input in immersion

31

settings, immersion learners still failed to develop native-like grammar when compared to their native peers. The finding also lends support to the claim that L1 and L2 acquisition may not share the same learning system or mechanisms. Therefore, most nurture researchers (e.g., Doughty, 2003; R. Ellis, 2008b, 2012; Lightbown & Spada, 1990; Long, 1998, 2000, 2007, 2014; Long & Robinson, 1998; Nassaji & Fotos, 2007; Norris & Ortega, 2000, 2001; Spada, 1997, 2011) propose that some focuses on forms or form-focused instruction (FFI) is necessary for L2 acquisition, especially for the development of target-like accuracy. Also, negative evidence in the form of corrective feedback, one type of FFI interventions, which helps draw learners' attention to the target form, is facilitative to L2 development (R. Ellis, 2008b, 2012a; R. Ellis, Basturkmen, & Loewen, 2001; R. Ellis, Loewen & Erlam, 2006; DeKeyser, 2007, 2015; Gass & Mackey, 2015; Long, 1998, 2007, 2014; Norris & Ortega, 2000, 2001; Schmidt, 2001; Spada, 1997, 2011; Spada, Jessop, Tomita, Suzuki & Valeo, 2014; Swain, 2005).

Findings from a plethora of oral and written feedback studies demonstrated that learners who received feedback generally displayed more noticing and produced more uptake episodes than those who did not (e.g., Bao et al., 2011; Ca Novas Guirao, et al., 2015; Coyle & Roca de Larios, 2014, 2020; Diab, 2015; Hanaoka & Izumi, 2012; Kartchava & Ammar, 2014; Kim & Bowles, 2019; Loewen & Philp, 2006; Mackey, 2006; Sachs & Polio, 2007; Santos et al., 2010; Sheen, 2010; Shintani, 2016; Shintani & R. Ellis, 2013; Simard et al., 2015). These findings suggest that feedback can efficiently draw learners' attention to problematic forms, inducing more "noticing the gap" (the mismatch between learners' L1 and L2) which subsequently facilitates learners' L2 learning processes.

In particular, most nurture advocates believe that some focus on forms is necessary especially for the learning of non-salient and communicatively redundant forms (Schmidt, 2001). The usage-based approach, for example, proposes that feedback that draws learners' attention to the non-salient or low contingency form helps facilitate the consolidation of form-function mappings; thus, contributing to overall acquisition processes. Besides serving as a noticing facilitator directing learners' attention to the problematic form, feedback also functions as a consciousness raising device prompting learners to notice the mismatch between their own interlanguage and the target-like

production. "Noticing the gap" then encourages learners to test or retest their hypothesis leading to the destabilization and reconstruction of their interlanguage knowledge (Bitchener & Storch, 2016; R. Ellis, 2006; Gass, 1997, 2003; Long, 1996, 2007, 2014; Lyster, 2004; Polio, 2012; Swain, 2005; Van Beuningen, 2010, Williams, 2012). Specially for those interactionists, the notions of noticing and attention are of significant as they believe that noticing is a pre-requisite condition for learning (i.e., acquisition) and learning cannot take place without noticing. That is, without noticing, input cannot become intake (Gass & Mackey, 2015; Long, 1996, 2007, 2014; Mackey et al., 2013; Robinson, 2002; Schmidt, 2001, Swain, 2005).

For skill acquisition theory, feedback is also useful for acquisition since it helps prevent learners to proceduralize inaccurate linguistic forms into their learning systems. Sociocultural approach also views feedback as dynamic scaffolding helping learners to become more independent in their learning. In short, for the nurture advocates, negative evidence and corrective feedback are unequivocally beneficial for L2 development and relevant to learning.

However, for the nature supposition, that is, the generative and processability approaches, two theoretical objections are raised against the roles of negative evidence and feedback in SLA. For generativists, given that negative evidence and feedback only facilitate explicit knowledge development (i.e., pseudolearning, conscious grammar knowledge) as opposed to implicit knowledge development (i.e., acquisition, native-like knowledge), and explicit knowledge does not contribute to acquisition (the non-interface position), the effectiveness of negative evidence or feedback is quite limited and even futile when acquisition is concerned. Nevertheless, as already mentioned, some generativists postulate that the provision of negative evidence maybe necessary for L2 as opposed to L1 acquisition (e.g., White, 2018, Schwarts & Sprouse,1996) because it is probable that L1 and L2 acquisition systems might not fully overlap and in such case the generative theory cannot be fully accounted for learners' L2 acquisition processes (Bley-Vroman, 1989, 1990; Clahsen et al., 2010; N. Ellis, 2015b, 2016a; Meisel, 2011; Morgan-Short & Ullman, 2012; Paradis, 2009; Selinger, 1972; Selinger & Lakshmanan, 1992; Ullman, 2015, 2016; Wulff & N. Ellis, 2018).

Indeed, until the generative assumption that L1 and L2 learning share the same learning mechanisms is proved valid, it does not seem legitimate to conclude that negative evidence is irrelevant and not facilitative to acquisition, as what is not required for L1 learning[5] may be essential for that of the L2. In addition, even if negative evidence can only facilitate explicit knowledge development as claimed, it is still open for discussion whether explicit knowledge plays any roles in the development of implicit knowledge (i.e., the interface issue). If explicit knowledge does in fact contributes to the development of implicit knowledge, negative evidence and feedback can also contribute to SLA.

With regard to the processability theory, negative evidence and feedback are also irrelevant to acquisition. Based on the theory ground, learners will acquire linguistic features based on their own developmental readiness following the already pre-setting developmental trajectories, and nothing including negative evidence or feedback can alter their course of developmental hierarchy. Despite that, evidence from oral feedback studies (e.g., Ammar, 2008; Mackey & Philp, 1998; McDonough & Mackey, 2006) have shown that while feedback may not be able to alter the course of development per se, the provision of feedback can accelerate learners' rate of acquisition in each stage of development and therefore feedback should be regarded as a useful tool contributing to L2 acquisition in some way.

Recent years have also witnessed a growing body of studies, meta-analyses and narrative reviews in support of the effectiveness of corrective feedback, both in the oral and written forms, on learners' L2 development (e.g., Abbuhl et al., 2018; Adams, 2003; Adams et al., 2011; Benson & DeKeyser, 2019; Bitchener & Storch, 2016; Bonilla López et al., 2018; Brown, 2016; R. Ellis, 2006, 2008b, 2012a; Gass & Mackey, 2015; Goo, 2012; Goo & Mackey, 2013; Kang & Han, 2015, Li, 2010; Li & Vuono, 2019; Loewen, 2012; Long, 1996, 2007, 2014; Lyster & Ranta, 2013; Lyster & Saito, 2010; Lyster, Saito & Sato, 2013; Mackey, 2012; Mao & Lee, 2020; Nassaji, 2016; Nassaji & Kartchava, 2017; Nemati et al., 2019; Polio, 2012; Sagarra & Abbhul, 2013; Sato & Loewen, 2018; Shintani & R. Ellis, 2013; Shintani, R. Ellis & Suzuki, 2014; Stefanou & Révész, 2015).

---

[5] It is still debatable in the field as to whether negative evidence also plays a part in L1 learning.

Nonetheless, while most of the above studies investigated the effectiveness of oral feedback in relation to L2 development, the efficacy of WCF on L2 acquisition is under-researched, leaving the huge gap in the literature. To fill the gap in the literature and shed more light on this underexplored topic, the current study aimed to further explore the efficacy of WCF on learners' L2 development, that is, investigating "feedback for acquisition" as opposed to "feedback for accuracy" aspect (Manchón, 2011). In the subsequent sections, relevant studies investigating the effectiveness of WCF on learners' L2 development, the main topic of the current investigation, are explored in more detail.

2.2. The efficacy of written corrective feedback

In previous sections, the French immersion studies (see Swain & Lapkin, 1982, 1986, 1998; Swain, 1984, 1985, 1995, 1998, 2005) have demonstrated that despite several years of intensive exposure to positive evidence in immersion settings, similar to how L1 learners acquire their native languages, most immersion L2 learners still failed to develop target-like accuracy. The findings, thus, suggest that positive evidence alone is not sufficient for SLA (Gass & Mackey, 2015; Long, 1996, 2014; Mackey et al., 2013; Swain, 1995, 2005) and that form-focused instruction (FFI) is necessary for the development of target-like accuracy (Doughty, 2003; R. Ellis, 2008b, 2012a; Lightbown & Spada, 1990; Long, 1998, 2000, 2007, 2014; Long & Robinson, 1998; Nassaji & Fotos, 2007, 2011; Norris & Ortega, 2000, 2001).

According to Spada (1997), FFI refers to "any pedagogical effort which is used to draw the learner's attention to language form implicitly or explicitly" (p.73). Negative evidence in the form of corrective feedback is, thus, categorized as one type of FFI (see R. Ellis, 2008b, 2012a) as it also helps drawing learners' attention to the target form facilitating L2 learning processes (R. Ellis, Basturkmen, & Loewen, 2001; R. Ellis, Loewen & Erlam, 2006; R. Ellis, 2006, 2008b, 2012b; Gass & Mackey, 2015; Long, 1998, 2007, 2014; Norris & Ortega, 2000, 2001; Schmidt, 2001, Spada, 2011; Swain, 2005). And because FFI is necessary for L2 learning as such, the effectiveness of corrective feedback, one type of FFI interventions, has become the focus of a growing number of empirical studies in SLA, most of which have attested its efficacy in

promoting the development of L2 learning (see Abbuhl et al., 2018; Biber et al., 2011; Bitchener & Storch, 2016; Brown, 2016; Goo & Mackey, 2013; Kang & Han, 2015; Lee, 2019; Li, 2010; Li & Vuono, 2019; Mao & Lee, 2020; Russell & Spada, 2006; Van Beuningen, 2010 for a full review).

Nonetheless, most corrective feedback studies focus more on exploring the efficacy of oral feedback as opposed to written corrective feedback (WCF), since most theorists believe that spontaneous oral language (i.e., speaking skill) is a better approximation of learners' target-like competence and oral feedback appears to be able to contribute directly to that said competence (Bitchener & Storch, 2016). Written language or writing, on the other hand, is usually viewed as a product of acquisition, rather than the skill that aids acquisition; hence, WCF which is a written response to learners' written errors is also disregarded as many do not believe that it could contribute to acquisition. Yet, for the past two decades, there has been a growing interest as to 1) whether WCF can also contribute to SLA the same way oral feedback does, 2) whether the development of L2 knowledge that occurs in oral context can also occur in written context, 3) whether certain type of WCF (direct vs. indirect and focused vs. unfocused) is proved more effective for L2 development than others, and 4) whether certain individual factor facilitates or impedes the efficacy of WCF (Bitchener & Storch, 2016). In effect, some WCF studies have already been conducted under these agendas but the conclusion is still far from conclusive. Since WCF is the main topic of the current investigation, the sections that follow provide a thorough review of previous WCF studies.

2.2.1. Definitions of written corrective feedback

SLA researchers have been using various terms to refer to written corrective feedback (WCF). According to Bitchener and Storch (2016), WCF refers to "a written response to a linguistic error that has been made in the writing of a text by an L2 learner" (p. 1). It is also included a response on lexical or non-grammatical errors like punctuation or spelling (Bitchener & Storch, 2016). In addition, according to Erlam, R. Ellis, and Batstone (2013), WCF can also be provided in an oral form responding to learners' written work during individual conferences or in class sessions.

On the explicitness/implicitness continuum, Polio (2012) points out that WCF is always explicit in nature and refers to WCF as "any explicit attempts to draw a learner's attention to a morphosyntactic or lexical error" (p.376). Polio's definition of "explicit" is based on DeKeyser's (1995) term stating that a technique/method/learning is explicit if some rules are presented or learners are drawn to pay attention to specific forms or structures. Based on his definition, it can be conjectured that all forms of WCF are explicit, regardless of when learners are given substituted corrections by teachers (i.e., direct corrective feedback), or when they are not given corrections and are encouraged to correct their own errors (i.e., indirect corrective feedback), since learners are given overt linguistic signals indicating their linguistic errors, they acknowledge that what is marked on their writing constitutes corrections and they are told to use their observation to correct their errors respectively (Polio, 2012). In this respect, it is worth noting that WCF is, therefore, dissimilar to oral feedback which "can be implicit or explicit depending on whether learners are made aware of the problematic nature of their speech performance…the implicit-explicit distinction does not apply to written CF" (Li & Vuono, 2019, p. 94).

Given that most WCF studies under "feedback for acquisition" agenda focus more on written responses to learners' morphosyntactic errors rather than content, lexical, or other mechanical types of errors, and that the current study only aimed to investigate the efficacy of WCF on learners' grammatical development, the term "WCF" operationalized in this present study will accordingly only refer to "a written response to learners' linguistic errors".

2.2.2. Objection to the provision of written corrective feedback
(Case against the provision of WCF)

The provision of WCF is an entrenched practice in most L2 composition classrooms as most teachers view WCF as an effective pedagogic strategy that could help improve learners' written accuracy as well as writing ability in general. Even though some teachers may be skeptical about its effectiveness in helping learners develop their L2 knowledge, most still believe that WCF can contribute to learners' L2 development in some way (Bitchener & Storch, 2016; Ferris, 2010; Mao & Lee, 2020). In the early years, the predominant focus of WCF research is on how WCF could help L2 learners develop their overall writing ability or "feedback for accuracy" aspect, situated within the "learning-to-write" dimension (Manchón, 2011; Santos et al., 2010, also see Hyland, 2010 for an overview). However, after Truscott's (1996) strong opposition towards the effectiveness of WCF contending that "grammar correction has no place in writing courses and should be abandoned (p.328), the focus of WCF research has been shifted more towards the potential role of WCF in aiding learners' interlanguage development or "feedback for acquisition" aspect under the "writing-to-learn" agenda, which views writing as the skill that could assist L2 development (Manchón, 2011). It is as Li and Vuono (2019) remark that the momentum of WCF studies have commenced since Truscott's (1996) article.

Since Truscott's (1996, 1999, 2004, 2007, 2010, also see Truscott & Hsu, 2008) provoking argument on the ineffectiveness of WCF as a pedagogical tool due to a lack of empirical and theoretical justifications attesting its efficacy and utility in composition classrooms and his call for the abolishment of the practice, there has been a surge of WCF research conducted to verify his claim. Truscott's claim is based on two main theoretical grounds, those of generative and processability theories (see sections 2.1.2.1. and 2.1.2.2.). In his first theoretical thesis, Truscott argues that since explicit knowledge cannot become implicit knowledge (Krashen, 1982, 1985) and WCF (i.e., negative evidence) can at best promote the development of explicit knowledge, WCF is then futile for acquisition (i.e., implicit knowledge) and therefore should be renounced. Nonetheless, as previously mentioned (see section 2.1.2.1.), Truscott's call for the abandonment of

WCF based on this ground appears precipitous, providing that it is still debatable in the field whether 1) L1 and L2 acquisition share the same learning mechanisms and 2) explicit knowledge plays a role in the development of implicit knowledge (i.e., acquisition).

Another theoretical argument raised by Truscott is that because all learners are constrained by the pre-determined developmental trajectories, and nothing including WCF can alter this fixed developmental course, the provision of WCF is fruitless, unless feedback can be provided exactly at learners' developmental readiness, which in practice, is too challenging to implement (Polio, 2012). In other words, WCF that does not target features that are at learners' developmental readiness is ineffective and such feedback should be discarded from practice since it consumes too much time and energy from both teachers and learners. Nevertheless, as already noted (see section 2.1.2.2.), most processability-based findings were drawn from learners who had intensive exposure to L2 input or "implicit learners" (DeKeyser, 2015), not from instructed L2 learners with limited exposure to the target language; thus, the theory may not be fully applicable to instructed L2 learners. It is also as R. Ellis (2012a) points out that "FFI [including corrective feedback] may be powerless to change the course of acquisition of implicit knowledge but it can still be effective in developing functional control of specific features irrespective of learners' developmental level" (p. 280).

A growing body of FFI studies and reviews have also attested the facilitative effect of corrective feedback as a form-focused intervention on L2 development (R. Ellis, 2008b, 2012a; R. Ellis, Basturkmen, & Loewen, 2001; R. Ellis, Loewen & Erlam, 2006; DeKeyser, 2007; Gass & Mackey, 2015; Long, 1998, 2007, 2014; Lyster, 2007; Nassaji & Fotos, 2007, 2011; Norris & Ortega, 2000, 2001; Schmidt, 2001; Spada, 1997, 2011; Spada et al., 2014; Swain, 2005). In particular, findings from oral feedback studies have demonstrated that even though feedback may not be able to alter learners' developmental stages, the provision of feedback can accelerate learners' rate of acquisition in each stage of development. Thus, the above evidence has shown that feedback is not as useless as claimed by Truscott. Instead, it is evident that feedback contributes to learners' L2 development in some way. More evidence on the effectiveness of WCF in promoting L2 development is provided in the next section.

It is also worth emphasizing that besides claiming that WCF practice is futile because the practice neglects the theoretical basis of SLA as previously discussed, Truscott further argues that there is no empirical evidence indicating that WCF could de facto contribute to learners' L2 development. Instead, there is compelling evidence showing that the provision of WCF does not help with the improvement of learners' writing accuracy. The research evidence he refers to is, however, limited to only four WCF studies (i.e., Kepner, 1991; Semke, 1984; Sheppard, 1992; Robb, Ross & Shortreed, 1986). Truscott used these studies to strengthen his ground since they all showed that there were no significant differences in learners' writing accuracy gains after receiving WCF and this implicated that WCF was ineffective in helping learners improve their L2 accuracy. However, it should be noted that the reliability and validity of these studies' findings are criticized by many researchers in the field for their design and methodology shortcomings (e.g., Bitchener & Ferris, 2012; Bitchener & Storch, 2016; Guenette, 2007; Mao & Lee, 2020; Polio, 2012; Van Beuningen et al., 2012).

With regard to Kepner's (1991) study, learners' pre-treatment ability was not measured due to the lack of the pretest task and this raises the questions as to how learners' improvement after treatment was determined in this study and whether the control and treatment groups were equivalent to begin with.

> Kepner analyzed the sixth set of journal entries (out of eight total) for comparison purposes and she did not look at the first set of journal entries to see where the students started out on this type of writing. (Ferris, 2003, p. 60)

Additionally, the author did not report whether there was any control for the writing task performing outside of class time and the length of the given written task was not controlled making it another "variable that could affect both errors and preposition counts" (Ferris, 2003, p. 60).

In Semke's (1984) study, there are also some design and execution shortcomings. The two main ones involve 1) the employment of different measurements for the control and other WCF treatment groups, and 2) the learners' incentive issue. Guenette (2007) has pointed out that since the control group's written work was graded based only on the number of written words, the control group may not feel worried or pressured to write or attempt to write less as their points would not be deducted even if they made some errors.

On the other hand, because the treatment groups' work was graded based on the ratio of errors to the number of written words, this may affect their mindset and the learners in the WCF groups may attempt to "write less for fear of making many mistakes" (Guenette, 2007, p.50). These two issues are confounding factors that can apparently affect the validity and the results of the study; thus, the conclusion drawn from this study must be interpreted with caution.

In Sheppard's (1992) study, the question has also been raised concerning the one-on-one conference each learner from the control group (which received only content feedback) had with the teacher, as it is unclear whether there was a discussion of content issue in relation to learner's linguistic errors or not (Bitchener & Storch, 2016). Ferris (2003) also mentioned a lack of inter-rater reliability check on the feedback coding data and this can seriously affect the overall findings. With regard to Robb, Ross and Shortreed 's (1986) study, the only main issue is that there was no real control group since the control group in this study also received marginal feedback (counting total number of errors per line) which in actual can also be regarded as one type of corrective feedback as it still helped raise learners' awareness on the target forms, albeit indirectly. A lack of a real control group, thus, inevitably has some impact on the results.

Because of such shortcomings, many L2 writing researchers (e.g., Bitchener & Ferris, 2012; Bitchener & Storch, 2016; Guenette, 2007; Mao & Lee, 2020; Van Beuningen et al., 2012) contend that these studies' findings should not be considered as evidence against the effectiveness of WCF and Truscott's claim against the effectiveness of WCF based merely on these four studies should not be viewed as valid.

In fact, a number of other WCF studies (e.g., Ashwell, 2000; Chandler, 2003; Fathman & Whalley, 1990; Ferris, 1995, 1997, 2006; Ferris & Helt, 2000; Ferris & Robert, 2001; Lalande, 1982) have shown positive results attesting the effectiveness of WCF in helping improve learners' writing accuracy. Yet, these studies are also suffered from some design and execution shortcomings, namely a lack of a real control group or the use of learners' revision task rather than the new piece of writing as WCF post-treatment measurement; thus, the studies' findings cannot be used as robust evidence in support of the effectiveness of WCF either (Bitchener & Storch, 2016).

However, mounting evidence of recent WCF research which already overcomes

the early research's design and methodological shortcomings has provided evidence in support of the effectiveness of WCF in promoting the development of L2 accuracy (e.g., Bitchener & Knoch, 2008; 2010a, 2010b; Benson & DeKeyser, 2019; Bonilla et al., 2018; R. Ellis, Sheen, Murakami & Takashima, 2008; Farrokhi & Sattapour, 2012; Ferris et al., 2013; Frear & Chiu, 2015; Guo, 2015; Hartshorn et al., 2010; Mohammadreza, 2020; Nemati et al., 2019, Rezazadeh et al., 2015; Rummel, 2014; Saeb, 2014; Sheen, 2007; Sheen, Wright, & Moldawa, 2009; Shintani & R. Ellis, 2013; Shintani et al., 2014; Shooshtari et al., 2019; Stefanou & Révész, 2015; Wagner & Wulf, 2016). Recent meta-analyses and synthesis reviews on WCF (i.e., Biber et al., 2011; Bitchener & Storch, 2016; Kang & Han, 2015; Li & Vuono, 2019; Mao & Lee, 2020; Russell & Spada, 2006) have also attested the facilitative effect of WCF on learners' L2 development. These WCF studies and meta-analysis reviews are thoroughly discussed in the subsequent sections.

2.2.3. The effectiveness of written corrective feedback

Learning from the shortcomings of early research, most recent WCF studies have successfully avoided those aforementioned design and methodology flaws, providing more valid and reliable evidence on the effectiveness of WCF. Some WCF meta-analyses and synthesis reviews are also conducted to justify the role of WCF in L2 learning (e.g., Biber et al., 2011; Bitchener & Storch, 2016; Kang & Han, 2015; Li & Vuono, 2019; Mao & Lee, 2020; Truscott, 2007). Most of these studies, though not all (except only for Truscott, 2007), have affirmed the effectiveness of WCF as a pedagogical tool helping learners develop their L2 knowledge.

To date, three meta-analyses[6] (i.e., Biber et al., 2011; Kang & Han, 2015; Truscott, 2007) have been conducted to investigate the effectiveness of WCF on L2

---

[6] Liu and Brown's (2015) meta-analysis did not investigate the overall effectiveness of WCF on L2 development but explored methodological aspects of WCF research. In the same vein, Mao and Lee (2020) only focused on "feedback scope" review (the efficacy of focused and unfocused WCF).

learning (excluding the ones that included oral feedback studies in their analyses[7]). The pioneer one and the only study that found a negative effect of WCF on learners' written accuracy is Truscott's (2007) meta-analysis. In his study, Truscott had accumulated the findings of L2 WCF studies in which he later analyzed and reported a small negative effect of WCF and made a conclusion that:

> The primary conclusion, based on the controlled experiments, is (a) the best estimate is that correction has a small harmful effect on students' ability to write accurately, and (b) we can be 95% confident that if it actually has any benefits, they are very small. (p. 270)

However, contrary to Truscott's (2007) findings, subsequent meta-analysis and synthesis review studies (e.g., Biber et al., 2011; Kang & Han, 2015; Li & Vuono, 2019; Mao & Lee, 2020) have shown that WCF is effective in helping learners develop their L2 knowledge. In addition, Truscott's (2007) study has been considered "notably narrow in scope" (p. 3, Kang & Han, 2015) since only published empirical studies were examined. Besides, the studies investigated in his analysis were mainly drawn from narrative reviews published by Ferris' (1999, 2003, 2004) and himself (1996, 1999), instead of using more extensive research available in the literature (Li & Vuono, 2019). Consequently, the result of his study with regard to the negative effect of WCF has to be viewed with caution.

In Biber, Nekrasova, and Horn's (2011) work which examined 25 published studies, the study found that the provision of WCF led to learners' gains, with a moderate to large effect size, indicating the facilitative effect of WCF in promoting L2 development. The study further revealed that lower proficient learners benefited more from WCF than those from higher proficiency levels. This finding, nonetheless, should also be viewed with caution due to an unequal sample size between the low and high proficiency sample groups and the fact that each proficiency level was not operationalized within the study.

In addition to learners' proficiency levels that appeared to mediate the efficacy of

---

[7] Russell and Spada's (2006) meta-analysis included both oral and WCF studies in their analysis. They combined and reported the overall effect size of both feedback types.

WCF, Biber et al. (2011) have found that "comment feedback"[8] was more efficacious than "error location"[9] and feedback targeting both forms and content appeared to be more effective than feedback exclusively targeting linguistic forms. Notwithstanding, the results of this study must be exercised with caution since the data used in this analysis was drawn from both L1 and L2 WCF studies and therefore the results may not be used as robust evidence in support of the role of WCF in L2 development.

In line with Biber et al.'s (2011) analysis, a recent meta-analysis conducted by Kang and Han (2015) which examined 21 published and unpublished WCF studies (to avoid the publication bias), also found that WCF did have a moderate to large positive effect on learners' L2 accuracy development. This, to some extent, affirms the potential of WCF in promoting learners' L2 development. Kang and Han's (2015) meta-analysis also revealed that the effect of WCF was mediated by a number of variables. Among the 9 variables investigated in this analysis, learners' proficiency was found to be the most influential factor moderating the effect of WCF, with high proficient learners benefited more from feedback than the lower ones. This finding differs from that of Biber et al. (2011) which reported that lower proficient learners seemed to benefit more from feedback than the higher proficient ones. However, as mentioned earlier, caution must be taken with regard to Biber et al.'s result, due to an unequal sample size representing each proficiency group. In addition, Li and Vuono (2019) comment that it is unclear how each proficiency level was operationalized among those sample studies used in the analysis.

Apart from learners' proficiency, the genre of the writing task was found to be another key moderator as WCF seemed to yield a larger effect size (i.e., more effective) when provided for composition writing and yielded a significantly lower effect size when provided for journal writing. Kang and Han (2015) believe that because journal writing is usually viewed as a form of free writing; thus, grammatical error does not seem to be the primary concern of this writing genre and error correction is then viewed as less important or even unnecessary. The analysis also showed that longer term treatment (more than three WCF treatment sections) did not lead to a larger effect size than short-

---

[8] It is unclear what Biber et al. (2011) termed as "comment feedback" actually refers to. It is speculated that it might refer to "metalinguistic comment".
[9] "Error location" in Biber et al.'s (2011) term might refer to "indirect WCF" which is the feedback that locates learners' errors but correct forms are not provided.

term treatment (one to three WCF treatment sections) and the finding aligns with that of Li's (2010) meta-analysis on oral feedback. Kang and Han posited that the short-term treatment study yielded a larger effect size because the short-term treatment was prone to focus on the narrower ranges of errors and, thus, WCF became more salient than that of the long-term one which had a tendency to focus on the wider ranges of errors making the feedback less salient to learners. In addition, it is possible that the saliency of WCF maybe impinged by other confounding factors the longer the treatment phase is prolonged.

Another interesting finding of this synthesis is that English as a second language (i.e., ESL) learners appeared to benefit more from WCF than English as a foreign language (i.e., EFL) learners. The finding, however, contradicts Li's (2010) meta-analysis finding on oral feedback studies. The discrepancy in the result is attributed to the possibility that EFL learners may be less concerned about their L2 writing than ESL learners (Ferris, 2010) and therefore they are not motivated to revise or edit their work following the given WCF (Hedgcock & Lefkowitz, 1994, 1996). Consequently, these EFL learners may not fully benefit from WCF. On the other hand, these learners may be more concerned about their speaking skill[10] and as a result, they tend to pay more attention to oral feedback than WCF. In sum, the results of Kang and Han's (2015) meta-analysis have suggested that WCF is facilitative to L2 development and that learners' L2 proficiency, writing genres, length of treatment and educational contexts are significant moderators of the efficacy of WCF.

Recent synthesis reviews (i.e., Bitchener & Storch, 2016; Li & Vuono, 2019; Mao & Lee, 2020) also provide mounting empirical evidence in support of the effectiveness of different types of WCF on learners' L2 development (e.g., Benson & DeKeyser, 2019; Bitchener, 2008; Bitchener & Knoch, 2009a, 2009b; 2010a, 2010b; Bonilla López et al., 2018; Diab, 2015; R. Ellis et al., 2008; Evans et al., 2011; Farrokhi & Sattapour, 2012; Ferris et al., 2013; Frear & Chiu, 2015; Guo, 2015; Hartshorn et al., 2010; Kim et al.,

---

[10] EFL learners tend to value speaking than writing skill as good speaking skill is instrumental for their future careers as most companies require their employees to have decent speaking skill but only some (very limited ranges of careers) ask for good writing skill. Consequently, most EFL learners may be more motivated to develop their speaking as opposed to writing skill.

2020; Kurzer, 2018; Mohammadreza, 2020; Nemati et al., 2019, Rezazadeh et al., 2015; Rummel, 2014; Saeb, 2014; Santos et al., 2010; Sheen, 2007; Sheen et al., 2009; Shintani, 2016; Shintani & R. Ellis, 2013; Shintani et al., 2014; Shooshtari et al., 2019; Stefanou & Révész, 2015; Van Beuningen et al., 2012; Wagner & Wulf, 2016). In the subsequent sections, empirical evidence in support of the effectiveness of different types of WCF is explored in more detail.

2.2.3.1. The effectiveness of different types of written corrective feedback
2.2.3.1.1. The effectiveness of direct and indirect written corrective feedback

Besides theoretical motivated to discover whether WCF is contributing to L2 acquisition, most L2 writing researchers are also keen to uncover whether some types of WCF are more facilitative to L2 learning. In the early years, the most common types of WCF found in L2 writing literature are "*direct*" and "*indirect*" WCF. At that time, L2 writing researchers tended to show more interest in indirect WCF (i.e., learners' errors are located but correct linguistics forms are not provided) since they believe that this type of feedback engages learners in "guided learning and problem solving" processes (Lalande, 1982, p.143); thereby, helps foster learners' long-term acquisition (Ferris et al., 2013).

However, advocates of direct WCF (i.e., correct linguistic forms are substituted to learners' errors) believe otherwise. They contend that direct WCF is more efficacious because it reduces learners' confusion towards error correction (Bitchener & Ferris, 2012; Ferris & Hedgcock, 2014, Ferris et al., 2013), unlike the indirect one that sometimes makes learners, especially the ones with limited L2 knowledge, even more confused because they may not be able to self-correct their own errors and they may start to self-doubt their hypothesized corrections. Chandler (2003) asserts that while learners can potentially internalize correct forms provided by direct WCF instantly and therefore, fully benefit from the given feedback, the delay in access to the target forms as a result of indirect WCF may, on the other hand, leave out the opportunity for learners to uptake the input provided by feedback. Moreover, it is proposed that only learners with adequate metalinguistic background can actually self-correct their own errors using indirect WCF (Hyland & Hyland, 2006; Sheen, 2007; Van et al., 2012). Ferris et al. (2013) comment

that indirect WCF may be more useful for learners' writing development (within the learning-to write aspect) since this type of feedback may promote learners' self-monitoring ability. Notwithstanding, when acquisition is the main concern (within the writing-to-learn aspect), direct WCF may be more efficient because it provides unambiguous and comprehensible information about the target structure to learners and learners can immediately benefit from this kind of information (Bitchener & Knoch, 2010a, 2010b; Ferris et al., 2013; Manchón, 2011; Van Beuningan et al., 2012).

However, R. Ellis (1994, 2009b) has posited that in fact the effectiveness of either direct or indirect WCF depends largely on the current state of learners' grammatical knowledge (i.e., proficiency levels), not on the feedback itself. R. Ellis' (1994, 2009b) speculation has later been confirmed by the finding of Kang and Han's (2015) meta-analysis suggesting that proficiency level is one of the most influential variables moderating the effectiveness of WCF. The result of Kang and Han's (2015) study also showed that there was no significant difference between direct and indirect WCF, even though direct WCF yielded a larger effect size than indirect WCF (g =.598 vs. g =. 361).

Thus far, it remains inconclusive as to whether direct or indirect WCF is more effective for L2 learning as research conducted under this agenda yields mixed findings. Some previous studies reported the superiority of direct over indirect WCF (e.g., Bitchener & Knoch, 2010a, 2010b; Chandler, 2003; Van Bueningen et al., 2008), while others found the advantage of indirect over direct WCF (e.g., Lalande, 1982; Storch & Wigglesworth, 2010; Tan & Manochphinyo, 2017). In addition, some neither found significant differences in the effects of direct and indirect WCF in revised texts (e.g., Ferris, 2006) nor in new written texts (e.g., Robb et al., 1986; Semke, 1984; Van Bueningen et al., 2012; Vyatkina, 2010). In part, a firm conclusion cannot be drawn because the validity and reliability of some of these findings were critically compromised due to the design and execution shortcomings, for example, an absence of a control group in Chandler's (2003), Lalande's (1982), Robb et al.'s (1986), and Storch and Wigglesworth's (2010) studies or the use of revised texts as accuracy measurement in Ferris's (2006) study. Besides, the discrepancy in research designs, methodology, operationalization of WCF, accuracy measurements, varied population is speculated to attribute to the contradictory results presenting in the literature (Hyland & Hyland, 2006;

Kang & Han, 2015; Li & Vuono, 2019; Liu & Brown, 2015; Mao & Lee, 2020; Shintani et al., 2014; Van Bueningen et al., 2012).

For example, in Storch and Wigglesworth's (2010) study which explored the effectiveness of direct and indirect WCF and the impact each feedback type had on learners' uptake and retention within collaborative writing context, the investigation found an advantage of indirect WCF over direct WCF. Forty-eight advanced ESL learners took part in this study. In the first session of the study, learners were asked to work in pairs. Each pair wrote a data commentary text following the graphic prompt. In the second session, each pair received either direct WCF in the form of reformulation, or indirect WCF in the form of editing symbol (providing codes for different error types), both targeting grammatical, lexical and mechanics (spelling and punctuation) errors. Each pair then had 15 minutes to look at the feedback and discussed it with his/her peer. After 15 minutes, feedback was removed, and each pair was asked to revise their original text (text produced in session 1). In the last session, individual learner was asked to write a new data commentary text using the same prompt. Learners' revised texts and individual texts produced in the last session were analyzed to find evidence of uptake and feedback retention.

The results of the analysis revealed that learners tended to be more engaged with indirect WCF than direct WCF leading to greater uptake and retention episodes. In particular, the study found that learners who extensively engaged with feedback showed high levels of uptake whereas those who had limited or no engagement with feedback did not produce uptake episodes or repeatedly produced persistent errors. In the same vein, retention of feedback also associated with the level of feedback engagement, with the higher level of engagement led to the higher level of uptake and retention episodes. It was also speculated that learners' beliefs and attitude towards the given feedback as well as their learning goals may have some influence on learners' ability on feedback retention.

Drawing on these findings, Storch and Wigglesworth concluded that indirect WCF was more beneficial for L2 development than direct WCF. It is, however, worth noting here that Storch and Wigglesworth's (2010) study is also subject to some methodological and design flaws. First of all, since the study used learners' revised texts composed in pairs after feedback provision (during the 2nd session) as evidence of uptake

while using new texts composed individually (during the last session) as evidence of retention, the inconsistency inevitably impacted the results to some extent. In addition, since the study did not have a control group, the credibility of the study's finding may be critically compromised.

In contrast to Storch and Wigglesworth's (2010) analysis, Van Beuningan, De Jong, and Kuiken's (2012) subsequent study found an advantage of direct WCF over indirect WCF. Van Beuningen et al.'s (2012) study investigated the effects of direct and indirect unfocused (i.e., comprehensive) WCF targeting both grammatical errors (e.g., article, inflectional, word order errors) and non-grammatical errors (e.g., lexical, orthographical, mechanics errors) on learners' accuracy. Two hundred and sixty-eight ESL learners participated the study. The proficiency level of these learners was, however, not clearly indicated. The study comprised 4 sessions: pretest, treatment, posttest, and delayed sessions, over a 6-week period. In this study, each learner completed three new written tasks during the pretest, posttest, and delayed posttest and completed one revised text during treatment.

The results revealed that both types of WCF were effective in helping learners improve their accuracy in their writing during the posttest and delayed posttest. The WCF gains were also shown to be superior than those of the self-editing and the sheer practicing control groups. Nonetheless, only direct WCF was shown to be efficient in promoting durable improvement in grammatical accuracy (i.e., maintained significant gains during the delayed posttest). Indirect WCF, on the other hand, only led to durable non-grammatical improvement. Based on the findings, Van Beuningen et al. suggest that different types of WCF may respond differently to different types of errors: direct WCF is more effective for linguistic errors whereas indirect WCF is more efficacious for non-grammatical types of errors.

Based on the two studies exemplified above, evidently, the conflicting results can be attributed to the discrepancy in designs and methodology adopted in each study. For instance, while Storch and Wigglesworth's (2010) study used revised texts as evidence of uptake, Van Beuningen et al.'s (2012) study used new pieces of writing as evidence of L2 development. In addition, whereas Storch and Wigglesworth's (2010) study combined learners' grammatical and nongrammatical errors into one single set of error category and

analyzed the two types of errors together in one single analysis, Van Beuningen et al.'s (2012) study separated grammatical from non-grammatical errors and performed individual analysis.

In addition, the conflicting results of the two studies can also be attributed to differences in learners' factors such as age, proficiency level, etc. For example, while learners in Storch and Wigglesworth's (2010) study were predominantly advanced graduate learners, learners in Van Beuningen et al.'s (2012) were in their second year of secondary education and their proficiency level was unclear (the authors used the term "limited language proficiency"). This, in some way, makes the two studies incomparable. This, to some extent, also explains why early WCF research which compared the effectiveness of direct with indirect WCF yielded conflicting results. Providing that only a limited number of studies under this agenda (e.g., Bitchener & Knoch, 2010a, 2010b; Van Bueningen et al., 2008, 2012) have successfully avoided early research's shortcomings, more robust investigation under this agenda is needed for the findings to be generalizable.

2.2.3.1.2. *The effectiveness of the more explicit types of written corrective feedback*

While it is still inconclusive as to whether direct or indirect WCF is more effective in promoting L2 development, feedback explicitness has recently become a focal research focus among L2 writing researchers as they speculate that the degree of explicitness of WCF may have influential effects on the efficacy of WCF (Mao & Lee, 2020). That is, the more explicit type of feedback may be more effective in promoting learners' L2 development than the less explicit one (Bitchener & Storch, 2016; Suzuki et al., 2019), because it might potentially induce learners' noticing at the level of understanding; thus, more facilitative to learning (than noticing at the level of attention) (Shintani et al., 2014; Stefanou & Révész, 2015). As a result, the efficacy of "*metalinguistic explanation*" (ME[11]) feedback, the feedback that provides learners with grammatical rules and

---

[11] R. Ellis (2009a) categorizes ME as another type of WCF apart from direct and indirect WCF, while some other L2 writing researchers include ME either in the direct or indirect WCF category depending on research contexts.

examples of correct usage (R. Ellis, 2009a), which was previously under-investigated, has been extensively explored either individually or in combination with other types of WCF (to make the feedback even more explicit[12]).

As elucidated, many researchers became more interested in ME due to its explicitness and its potential to foster deeper metalinguistic understanding about the target structure (Nassaji, 2015; R. Ellis et al., 2008), given that it provides learners with grammatical rules and examples of correct usage. A number of studies have been conducted to investigate the effectiveness of "metalinguistic" feedback or "metalinguistic" feedback in combination with other types of WCF (e.g., Benson & DeKeyser, 2019; Bitchener, 2008; Bitchener & Knoch, 2010a, 2010b; Shintani & R.Ellis, 2013; Shintani et al., 2014; Stefanou & Révész, 2015). Most of these studies are rigorously designed and overcome the shortcomings of previous studies since all had a real control group, used a new piece of writing as evidence of improved accuracy and adopted the pre-post-delayed posttest structure, etc. Still, it is inconclusive as to whether "metalinguistic" feedback is more effective and whether it has any beneficial values when combined with other types of WCF. Some studies (e.g., Bitchener & Knoch, 2010a; Guo, 2015; Sheen, 2007; Shintani & R. Ellis, 2013) found that the more explicit WCF is more beneficial, while some (e.g., Bitchener & Knoch, 2008, 2010b; Stefanou & Révész, 2015; Suzuki et al., 2019) did not find any beneficial effect adding ME with other types of WCF, implicating that the degree of feedback explicitness does not significantly mediate the overall effectiveness of WCF as conjectured. Several other studies even found the advantage of direct WCF over ME (e.g., Benson & DeKeyser, 2019; Rummel, 2014; Shintani et al., 2014).

However, albeit the mixed findings, there seems to be a consensus among these studies that direct WCF is effective in aiding learners' accuracy development and its effect is durable (e.g., Benson & DeKeyser, 2019; Bitchener & Knoch, 2010b; Shintani et al., 2014; Stefanou & Révész, 2015; Suzuki et al., 2019). Kang and Han's (2015) meta-analysis also found that direct WCF yielded a substantially larger effect size than indirect

---

[12] For instance, direct WCF combined with metalinguistic feedback is perceived as even more explicit than direct WCF only, or indirect WCF combined with metalinguistic comments is more explicit than indirect WCF only (Suzuki et al., 2019).

WCF, implicating that direct WCF may be more effective in helping learners develop their L2 knowledge.

In Bitchener and Knoch (2010b), the differential effects of different types of WCF on the two functional uses of English articles ("a" and "the") were investigated over a 10-month period. Fifty-two low-intermediate ESL learners participated in this study. These learners were assigned into one of the four groups; control, direct (focused[13]) WCF, direct (focused) WCF with written ME, or direct (focused) WCF with written and oral (30-minute lesson on the two article uses) ME. During the experiment, each learner was required to produce 5 pieces of written work: during the pretest, immediate posttest, and three delayed posttests. Learners received feedback one week after the pretest, had 5 minutes to review their feedback and then completed the immediate posttest. The first delayed posttest was performed on the eighth week, the second and third delayed posttests were administered during the sixth and tenth month, respectively. The results revealed that all WCF treatment groups outperformed the control group in the immediate posttest and all three delayed posttests, and there were no significant differences among the WCF treatment groups implicating that all WCF types were equally effective. Overall, the study suggested that WCF was effective in promoting learners' grammatical accuracy and the effect was durable over the 10-month period; therefore, receiving WCF was more beneficial than a lack thereof. However, since all types of WCF were equally effective, this indicated that the degree of feedback explicitness may not have any influential effect on the efficacy of WCF. The finding is in line with previous findings (e.g., Bitchener, 2008; Bitchener & Knoch, 2010a) which also found no advantages of additional provision of ME.

The findings of Bitchener and Knoch (2010b) are later corroborated by the findings of Stefanou and Révész's (2015) investigation which compared the effects of direct WCF only and direct WCF plus ME on the development of English article uses on specific and general plural references of the 89 intermediate EFL Greek learners, and found that the two WCF were equally effective in helping learners improve their written accuracy given that no significant difference between the two feedback types was detected. In this pre-post-delayed posttest design, learners were assigned into the control,

---

[13] The feedback only provided corrections to errors of the two article uses.

direct WCF only or direct plus ME groups. They were then required to complete the pretest, attend two WCF treatment sessions and complete the immediate and delayed posttests, respectively. Overall, in this study, each learner completed two written tasks and received two feedback treatment episodes during treatment.

The results of the study demonstrated that both WCF groups significantly outperformed the control group on article uses with specific plural references on both the text summary task and the truth value judgement test (TVJT) during the immediate and delayed posttests. However, there were no significant differences among all groups (both treatment and control) with regard to their article uses on general plural references. Stefanou and Révész (2015) proposed that the already high pretest scores of all the participating groups, i.e., the ceiling effect, may account for WCF groups' small gains leading to the non-significant difference between the treatment and control groups. The result of the study also showed that when the effects of direct WCF and direct WCF plus ME were compared, there were no significant differences between the two types of WCF detected.

Stefanou and Révész's (2015) finding concurs with Bitchener's (2008), Bitchener and Storch's (2008), Bitchener and Knoch's (2010a, 2010b) previous studies which did not find any beneficial effect combining ME with direct WCF. The finding, however, contradicts the result of Sheen's (2007) work which found beneficial effect of complementing direct WCF with ME. In sum, the findings of Stefanou and Révész (2015) have affirmed the effectiveness of direct WCF with or without ME in promoting learners' grammatical development, while implicating that the degree of explicitness of WCF may not significantly mediate the overall effectiveness of WCF as previously conjectured.

In line with Stefanou and Révész's (2015) findings, the study of Suzuki, Nassaji and Sato (2019) also found that WCF with or without ME were equally effective in helping learners develop their L2 knowledge. In their study, Suzuki et al. (2019) examined the effectiveness of different types of direct and indirect WCF (with ME and without ME) on 88 Japanese EFL learners'[14]uses of English indefinite articles and past perfect tense. The learners in this study were assigned into one of the four groups: direct

[14] The authors did not define the learners' proficiency level. They only reported learners' TOEIC score ranges.

WCF only, direct plus ME, indirect WCF, and indirect plus ME. Each learner was required to write two new narratives and received one feedback episode during treatment.

The results of the study showed that both types of WCF in combination with ME or without ME were equally effective in helping learners improve their past perfect tense during the delayed posttest (two weeks after treatment). Yet, all types of WCF did not lead to gains on indefinite articles. Suzuki et al. (2019) believe that the type of linguistic structure is a significant factor mediating the effectiveness of WCF. With regard to the degree of feedback explicitness, they found that those who received direct WCF only and direct plus ME outperformed those who received indirect WCF only and indirect plus ME on the use of past perfect tense. Suzuki et al. (2019) attributed the superiority of the direct WCF groups (with ME and without ME) to their higher pretest scores (51%-56% accuracy) when compared to the indirect WCF groups (approximately 41%), although the difference was not significant. Further, since direct WCF only and direct WCF plus ME were equally effective, Suzuki et al. (2019) concluded that the degree of feedback explicitness may not be a determining factor influencing the efficacy of WCF. All in all, the findings of Suzuki et al.'s (2019) study have attested the impartial effects of direct and indirect WCF with ME or without ME on learners' L2 learning and suggested that target linguistic structure may play a more significant role in mediating the overall efficacy of WCF. However, it should be noted that the study did not have a control group; therefore, the results must be interpreted with caution.

In contrast to the above results, the findings of Shintani, R. Ellis, and Suzuki's (2014) study, which explored the effectiveness of direct WCF and ME on the use of indefinite articles and hypothetical conditional structures of 171 low intermediate Japanese EFL learners, showed that direct WCF was more effective than ME providing that its effect was found to be more durable.

In this study, learners were assigned into one of the five groups: direct WCF only, direct WCF plus revision, ME only, ME plus revision and control. In the first session, all groups wrote their first written task (used as the pretest task) which was then returned in the second session with feedback. In the second session, after examining the feedback, the direct WCF and ME only groups wrote a new written piece (used as the immediate posttest task), while the direct WCF and ME plus revision groups needed to rewrite their

original pieces first before writing the new pieces. The control group, on the other hand, did not receive any feedback and only had to write a new written piece in this session. In the third session, all groups completed the third written task (used as the delayed posttest task). Overall, each learner in this study completed three written tasks and received one feedback treatment.

The results revealed that all WCF treatment groups significantly outperformed the comparison group on hypothetical conditional usage during the immediate posttest but only the direct WCF plus revision group outperformed the comparison group during the delayed posttest. In other words, even though the two types of WCF (with or without revision) were effective for hypothetical conditional considering learners' significant increases in accuracy during the posttest, only direct WCF had shown to have a long-term effect on this structure. This suggests that direct WCF may be more effective than ME in tackling a more complex syntactic feature especially if the long-term gain is to be expected. Further, except for the direct WCF plus revision group, Shintani et al. (2014) concluded that "there was no long-term advantage for the feedback" (p. 123). This finding, however, runs contrary to the findings reported in Bitchener and Knoch's (2010a) and Van et al.'s (2012) studies which detected the durable effect of direct WCF on learners' L2 learning outcomes. Shintani et al. (2014) argued that the discrepancy in findings is attributable to the fact that the hypothetical conditional targeted in this study is more complex than the linguistic features targeted in those studies.

In addition, while all types of WCF provided in this study were effective in helping learners improve their hypothetical conditional usage, none of the feedback had any positive effect on the development of learners' indefinite article uses. The authors explained that when WCF is aimed at two different linguistic structures, learners are likely to pay more attention to the more salient structure that contributes more to the global meaning of their written texts (e.g., the hypothetical conditional) than the non-salient redundant feature (e.g., the indefinite articles).

The result of Shintani et al.'s (2014) study, nonetheless, differs from that of Shintani and R. Ellis (2013) which also investigated the effectiveness of direct WCF and ME but found that only ME had a positive effect on learners' indefinite article usage. Yet, the effect of ME was not durable as it wore off after the posttest. Shintani et al. (2014)

posited that ME was effective in Shintani and R. Ellis' (2013) study because in their study, learners received only WCF targeting one structure (i.e., indefinite articles), they can then pay their full attention to their infinite article usage and benefited from ME, unlike this study that targeted two structures at the same time. Shintani et al. (2014) also proposed that ME had shown to be more effective in improving learners' outcomes than direct WCF in Shintani and R. Ellis' (2013) study because ME provided learners with explicit rules of article usage, which helped deepening learners' understanding towards the target structure aiding their overall progress. In contrast, direct WCF may not provide enough information for learners to extrapolate the rules for article usage; as a result, learners failed to make subsequent correction and improvement on the target structure (Shintani et al., 2014).

The finding of Shintani et al.'s (2014) study which indicated the superiority of direct WCF over ME is subsequently substantiated by Benson and DeKeyser's (2019) recent research which also found an advantage of direct WCF over ME. In Benson and DeKeyser's (2019) work which explored the effects of direct WCF and ME on the simple past tense and present perfect tense, 151 low intermediate to advanced ESL learners took part in this pre-post-delayed posttest designed study. During the treatment phase, each learner from the WCF treatment groups received feedback on his/her two essays while the control group received only content feedback on their writing.

The results of the study revealed that both WCF groups outperformed the comparison group (which showed no gain) on both structures during the immediate posttest. However, only the direct WCF group was able to maintain their significant gains on simple past tense, outperforming the control group, during the delayed posttest (4 weeks after the immediate posttest). The ME group could not maintain their posttest gains to the delayed posttest. Nonetheless, both WCF treatment groups could not retain their gains on present perfect tense. With regard to the more durable effect of direct WCF on simple past tense, Benson and DeKeyser (2019) explained that since the learners in this study may already have some metalinguistic knowledge of the simple past tense, direct WCF was then more effective than ME as it provided learners with positive evidence that learners can use to confirm or reject their hypothesis testing instantly (leading to the consolidation or destabilization of their knowledge). However, for the present perfect

tense that the effects of the two WCF could not be sustained after the posttest, Benson and DeKeyser (2019) attributed this to a lack of prior declarative knowledge about the structure and the fact that many learners may not have been ready to develop the knowledge of this particular structure.

In sum, the findings of Benson and DeKeyser's (2019) study corroborate the results of previous research which affirm the beneficial effect of WCF in promoting learners' grammatical accuracy (e.g., Bitchener & Knoch, 2010a, 2010b; Diab, 2015; R. Ellis et al., 2008; Evans et al., 2011; Farrokhi & Sattapour, 2012; Frear & Chiu, 2015; Hartshorn et al., 2010; Rummel, 2014; Saeb, 2014; Santos et al., 2010; Sheen, 2007; Shintani & R. Ellis, 2013; Shintani et al., 2014; Stefanou & Révész, 2015; Van Beuningen et al., 2012) and attest the superior effect of direct WCF over ME (e.g., Rummel, 2014; Shintani et al., 2014). Nonetheless, more studies under this agenda are needed to warrant the claim that direct WCF is de facto more effective than ME or other less explicit types of WCF (e.g., indirect WCF).

2.2.3.1.3. *The effectiveness of focused and unfocused written corrective feedback*

Another dichotomy that attracts researchers' interests is that between "*focused*" (only specific error types are corrected while the rest is disregarded) and "*unfocused*" (all or most grammatical errors are corrected) WCF. Different conjectures have been proposed regarding the effectiveness of both types of feedback. R. Ellis (2008a) argues that focused WCF is more effective than unfocused WCF because learners are likely to notice and understand corrections better when they only have to pay attention to a few types of grammatical errors. His assumption is based on SLA theoretical premise that noticing and understanding are the prerequisite for language acquisition and that noticed input is more likely to become intake (Schmidt, 1994). Bitchener (2008) and Sheen et al. (2009) corroborate R. Ellis' (2008a) premise, contending that unfocused WCF is more likely to overload learners' attentional and cognitive capacities, as learners need to pay attention to a wide range of corrections at the same time; therefore, limiting their feedback processing. Lee (2019) also criticizes that it is doubtful if unfocused WCF can be effective for learning, since all errors are corrected without considering learners'

proficiency and developmental readiness (correction is ineffective for learning if learners are not at the stage where they can comprehend the corrected features). Focused WCF, on the other hand, seems to be a more promising feedback technique considering that only a selective number of errors are targeted. As a result, learners can notice the feedback more easily and this is more facilitating to L2 learning processes. Truscott (2001) himself also contends that for the feedback to be more effective (rather than harmful), "correction must be used selectively" (p. 93). For these reasons, researchers in favor of focused WCF assume that focused WCF is more beneficial for L2 learning than the unfocused one.

Notwithstanding, the proponents of unfocused WCF argue that the attentional capacity problem might be more critical in online (i.e., speaking) rather than offline processing such as in writing. Ferris (2010) asserts that correcting only specific types of errors while ignoring the rest may confuse learners and does not help with their overall writing ability. Hartshorn et al. (2010) postulate that focused WCF may disappoint learners who expect to know all their errors or have all their errors corrected. Van Beuningen et al. (2012) further comment that even though focused WCF may better enable learners to restructure their interlanguage since they are exposed to corrections of the same errors repeatedly, unfocused WCF corresponds to actual practice and, consequently, seems to have higher ecological validity. In Kang and Han's (2015) meta-analysis, even though focused WCF had shown to yield a larger effect size on learners' L2 learning outcomes than unfocused WCF (g = .690 vs. g = .329), the difference between the two was not statistically significant.

To date, it remains unsettled as to whether focused or unfocused WCF is more effective in promoting L2 accuracy. Most studies that only investigated the effectiveness of focused WCF (e.g., Benson & DeKeyser, 2019; Bitchener 2008, 2012; Bitchener & Knoch, 2008, 2009b; Nemati et al., 2019; Rezazadeh et al., 2015; Sheen, 2007; Shintani & R. Ellis, 2013; Shintani et al., 2014; Shooshtari et al., 2019; Stefanou & Révész, 2015) found that focused WCF is beneficial for L2 learning. Likewise, studies that only explored the effectiveness of unfocused WCF (e.g., Bonilla López et al., 2017, 2018; Santos et al., 2010; Van Beuningen et al., 2012; Wagner & Wulf, 2016) also found unfocused WCF effective for L2 development. Only a small number of studies have actually been conducted to compare the effectiveness of these two types of WCF in one

single study (e.g., R. Ellis et al., 2008; Farrokhi & Sattapour, 2012; Frear & Chiu, 2015; Saeb, 2014; Sheen et al., 2009) and the results are incongruent, leaving the dispute as to whether focused or unfocused WCF is more effective remains open for discussion and further investigation.

For example, in R. Ellis, Sheen, Murakami, and Takashima's (2008) study which compared the effects of direct focused and direct unfocused WCF on the use of English definite and indefinite articles of 49 intermediate Japanese EFL learners, the study reported the comparable effectiveness of both focused and unfocused WCF in improving learners' accuracy. In this pre-post-delayed posttest designed study, learners completed the pretest, attended 3 treatment sessions, and completed the immediate and delayed posttests 4 weeks later. In each treatment session, each learner was required to write one narrative in which the teacher provided the feedback on and returned it back to the learner the subsequent week. While the two WCF groups received either direct focused WCF (correcting only English definite and indefinite article errors) or direct unfocused WCF (correcting article, other grammatical and vocabulary errors), the control group only received general comments on their content (e.g., what happened next? good etc.). Overall, each learner completed 3 narratives and received 3 feedback treatment episodes during treatment.

The results of the study showed that both WCF groups significantly outperformed the control group at the delayed posttest, and no significant difference was found between the focused and unfocused WCF groups, suggesting the indistinguishable effectiveness of both types of WCF in improving learners' grammatical accuracy. This finding is subsequently corroborated by the finding of Saeb (2014) which also found no differential effect between focused and unfocused WCF. In Saeb's (2014) study, the effects of direct focused and direct unfocused WCF on the use of English third person singular 's' morpheme of 79 EFL beginners were examined. Learners were assigned into one of the three groups: control (received general comments on content), direct focused (received only correction on English third person singular 's') and direct unfocused WCF (received correction on all types of errors). Learners then completed the pretest and attended two treatment sessions. During the first treatment session, each learner was asked to watch a film and write a summary about what happened. The feedback on the summary was

returned a week after and each learner was asked to examine his/her own feedback before beginning the second treatment session: watching another film and writing another summary. After the treatment phase, learners completed the posttest.

The results of this study showed that both WCF groups performed significantly better on the use of English third person singular 's' morpheme from pretest to posttest. The two WCF groups also outperformed the control group, suggesting that the provision of grammar feedback was more beneficial than a lack thereof. However, no significant difference between the two feedback types was detected, indicating the comparable efficacy of both feedback types in aiding learners' accuracy development. Notwithstanding, it should be noted that in this study only the error correction test was employed as accuracy measurement during the pretest and posttest (no writing task); therefore, this, to some extent, may appear to compromise the validity of the findings. In addition, given that the delayed posttest was not administered, the long-term effect of WCF on this particular structure was not confirmed.

In Frear and Chiu's (2015) study, the effects of indirect focused (targeting weak verbs) and indirect unfocused (targeting all errors) WCF were explored. Forty-two Taiwanese EFL learners[15] participated in this study. Each learner was asked to produce three writing tasks: the first one was written during the pretest which was returned with feedback a week later, the second one was written in class after receiving the feedback and this one was taken as the immediate posttest task, and the third one was written two weeks later during the delayed posttest. The control group did not receive any feedback during the experiment. The results of the study showed that both WCF groups significantly outperformed the control group during the posttest and delayed posttest. Further, there were no observable differences between indirect focused and indirect unfocused WCF detected in this study.

Based on the findings of the three studies above (i.e., R. Ellis et al., 2008; Frear & Chiu, 2015; Saeb, 2014), even though these studies are similar in their research findings demonstrating the indistinguishable effects of focused and unfocused WCF on L2 learning, the discrepancy in learners' proficiency, target linguistic structures, feedback operationalization and techniques (while R. Ellis et al. and Saeb used "*direct*" technique,

---

[15] The authors did not indicate the proficiency level of the learners.

Frear and Chiu used "*indirect*" technique), research designs and measurements, to name just a few, make them incomparable. This, thus, leads to a lack of adequate comparable evidence to draw a conclusive answer from, causing a gap in WCF literature.

Variations in target population, target linguistic structures, feedback techniques, outcome measures as well as research designs as described above are also found in the studies that found the advantage of direct focused over direct unfocused WCF making it more difficult to draw any firm conclusions from the existing findings.

In Sheen, Wright, and Moldawa's (2009) study which compared the effects of direct focused (directing at article errors only) and direct unfocused WCF (directing at article and other grammatical errors) on accurate uses of articles, copula 'be', regular and irregular past tense forms and preposition of 80 intermediate ESL learners, the result revealed that direct focused WCF was more effective than direct unfocused WCF.

In this study, 80 learners were assigned into one of the four groups: control, writing practice only, direct focused and direct unfocused WCF. Learners attended two treatment sessions where each of them was asked to write two short narrative essays based on two Aesop's fables. After treatment, all groups completed the posttest and delayed posttest. The results of the study demonstrated the advantage of direct focused over direct unfocused WCF. More importantly, the direct focused WCF group demonstrated to obtain the highest gain not only on their target structure, i.e., articles, but also on other targeted linguistic features, i.e., copula 'be', regular and irregular past tense forms and preposition. Further, only the direct focused WCF group significantly outperformed the control group at the posttest and delayed posttest while the unfocused WCF did not. Sheen et al. (2009) postulated that direct unfocused WCF was ineffective due to the amount of correction that may overwhelm learners' cognitive capacities leading to inefficacious feedback processing.

Farrokhi and Sattarpour's (2012) finding concurs with that of Sheen et al. (2009) which found an advantage of direct focused over direct unfocused WCF. In their study, Farrokhi and Sattarpour (2012) explored differential effects of direct focused and direct unfocused WCF on the use of English definite and indefinite articles by 120 Iranian EFL

learners from "low and high" levels of proficiency (based on TOEFL Scores)[16]. Sixty learners from each proficiency level were assigned to either the control, direct focused WCF (directing only at article errors) or direct unfocused WCF (directing at article, copula 'be', regular and irregular past tense and preposition errors) group. After that, learners completed the pretest, attended the three-week treatment sessions where each learner was asked to produce 5 narratives[17] based on provided fables and received feedback in return. One week after treatment, learners completed the posttest. The results of this study revealed that for both low and high proficient groups, direct focused WCF significantly outperformed both the control and direct unfocused WCF groups, indicating the more effectiveness of direct focused over direct unfocused WCF.

While Farrokhi and Sattarpour's (2012) result appears to align with that of Sheen et al. (2009), differences in their target population (ESL vs. EFL), proficiency levels (intermediate vs. low and high) and research designs, also make the two studies incomparable. As previously mentioned, such variations limit the amount of comparable research leading to insufficient evidence to draw the affirmative conclusion from regarding the more or less effectiveness of direct focused or direct unfocused WCF on L2 development. This line of research is, thus, still in need of further investigation so that a firm conclusion could be made.

---

[16] The authors did not clearly verify the score ranges for the low and high level of proficiency. That is, learners' proficiency levels were not clearly defined.
[17] The treatment procedure was not clearly elucidated.

2.2.3.2. The effectiveness of written corrective feedback on the development of learners' explicit and implicit knowledge

As illustrated in previous sections, to date there is much empirical evidence confirming the effectiveness of WCF as an FFI tool helping learners improve their L2 knowledge. Its positive effect on accuracy and presumably on explicit knowledge development is widely accepted (Benson & DeKyser, 2019; Bitchener & Ferris, 2012; Bitchener & Storch, 2016; Nemati et al., 2019; Polio, 2012; Rezazadeh et al., 2015; Shintani & R. Ellis, 2013; Shintani et al., 2014; Shooshtari et al., 2019; Stefanou & Révész, 2015; Wagner & Wulf, 2016; Williams, 2012). Nonetheless, most of the studies mentioned in previous sections primarily focus on "feedback for accuracy" (i.e., whether WCF assists learners to write more accurately) as opposed to "feedback for acquisition" aspect (i.e., whether feedback can facilitate L2 acquisition, Manchón, 2011). Until now, only a few studies (i.e., Nemati et al., 2019; Rezazadeh et al., 2015; Shintani & R. Ellis, 2013; Shooshtari et al., 2019) have attempted to investigate the effect of WCF on learners' explicit and implicit knowledge development, and the findings are incongruent.

In Shintani and R. Ellis' (2013) pioneer study, which compared the effects of direct focused WCF and ME on the development of learners' explicit and implicit knowledge of English indefinite articles, the study reported the beneficial effect of ME in aiding learners' development of explicit knowledge. Forty-nine low intermediate ESL learners took part in this study. They were assigned into one of the three groups: control (received no feedback), direct focused (received feedback on both definite and indefinite articles)[18], and ME (received handout explaining grammatical rules of both definite and indefinite articles). An error correction test (ECT) was used to measure learners' explicit knowledge, while a timed picture composition task (under 20-minute limit) was adopted as measure of learners' implicit knowledge. In the first session of the study, all groups completed the ECT and the first writing task (served as the pretest task). In the second

---

[18] It is unclear why the authors decided to give feedback on both definite and indefinite articles while their only target structure was indefinite articles and learners' progress on the use of definite articles was not analyzed. It will be more fruitful if the learners' uses of both definite and indefinite articles were analyzed, given that they already received feedback on both article usage.

session (two days after the first session), WCF groups received feedback on their first written task, revised the task, and completed the second writing task (served as the immediate posttest task). The control group, on the other hand, did not receive any feedback but was asked to revise their work and completed the second writing task. In the last session (two weeks after the second session), all groups completed the ECT and the third writing task (served as the delayed posttest task). Overall, each learner in this study completed three writing tasks and received one feedback treatment.

The results of the study revealed that ME had a positive effect on the development of learners' explicit knowledge as learners in the ME group significantly improved their ECT scores and writing performance during the posttest. The ME group also outperformed both the control and direct focused groups, both of which made no significant improvement at all. Shintani and R. Ellis believed that ME was effective for the development of explicit knowledge because it induced noticing at the level of understanding and the understanding of the target structure rules enabled learners to successfully develop their explicit knowledge.

In contrast, direct focused WCF was ineffective because it only induced noticing at the attention level which was adequate to make learners aware of their errors yet was inadequate to help them establish an understanding of the target structure rules necessary for the development of explicit knowledge. Nonetheless, since the effect of ME wore off after the posttest, Shintani and R. Ellis concluded that ME did not have any effect on implicit knowledge development, arguing that if ME had had any effect on learners' implicit knowledge, the effect must be evident in both immediate and delayed posttests. Shintani and R. Ellis attributed the ineffectiveness of both types of WCF in promoting implicit knowledge development to the one-shot treatment provided in the study and suggested that multiple corrections (more than one WCF treatment session) may be more facilitative given that implicit knowledge generally needs more time and intensive exposure to input to be developed.

Subsequent work by Rezazadeh, Tavakoli and Eslami Rasekh (2015) which investigated the effects of direct focused WCF and ME on the development of learners' explicit and implicit knowledge of English definite and indefinite articles, has, however, found that both direct focused WCF and ME were equally effective in promoting the

development of learners' explicit and implicit knowledge. Still, ME had shown to have a more durable effect than direct focused WCF. Ninety-four intermediate Iranian EFL learners participated in this study. They were randomly assigned into one of the three groups: control (received no feedback), direct focused (received corrections on definite and indefinite article errors) and ME (received metalinguistic handout adopted from Shintani and R. Ellis, 2013). An error correction test (ECT) and an untimed grammaticality judgment test (UGJT) were used to measure learners' explicit knowledge, whereas a timed narrative writing task (also similar to that of Shintani and R. Ellis, 2013) and a speeded dictation were employed to measure learners' implicit knowledge. In the first week, learners from all groups were asked to complete the ECT, UGJT, timed narrative writing and speeded dictation tasks. One week later, learners from the two WCF treatment groups received feedback on their previous narrative writing tasks, reviewed their feedback for 5 minutes, revised without access to the feedback, and wrote a new narrative task (served as the immediate posttest task). The control group learners, on the other hand, did not receive any feedback so they were asked to write a new narrative straight away. After completing the second narrative writing task, all learners completed the ECT, UGJT and speeded dictation task. Three weeks later during the delayed posttest, learners completed the same set of tasks. The definite and indefinite article combined scores were then analyzed. Overall, during the study, each learner from the WCF group wrote three new narratives, revised one text and received one feedback treatment. The control group learner, on the other hand, only wrote three new narratives but were not asked to revise their texts and did not receive any feedback.

The results of the study revealed that both WCF were effective in aiding the development of learners' explicit and implicit knowledge. Both WCF groups also outperformed the control group during the immediate posttest. Nonetheless, only the ME group could sustain their gain scores three weeks after the immediate posttest, outperforming both the control and the direct focused groups. Rezazadeh et al. (2015), thus, concluded that ME was more effective than direct focused WCF as its positive effect on learners' explicit and implicit knowledge was sustained over time, while the effect of direct focused WCF wore off after three weeks. Similar to Shintani and R. Ellis (2013), Rezazadeh et al. (2015) speculated that ME was more effective because it

efficiently promoted noticing at the understanding level helping learners form the metalinguistic knowledge of the target structure necessary for the advancement of learners' explicit and implicit knowledge.

Nonetheless, the findings of Rezazadeh et al. (2015) run counter to those of Shintani and R. Ellis (2013) which did not find any effect of direct focused WCF on learners' explicit and implicit knowledge development but the positive effect of ME in promoting learners' explicit knowledge. Rezazadeh et al. (2015) speculated that the fact that their study analyzed the combined scores of the definite and indefinite articles while Shintani and R. Ellis (2013) only analyzed the indefinite scores may, to some extent, attribute to the discrepancy in their findings. Differences in target population (EFL vs. ESL) were also postulated to partially account for the divergent findings. It is also worth noting that in Rezazadeh et al.'s study, the WCF groups' improvements on timed writing and speeded dictation tasks during the immediate posttest were considered evidence of learners' implicit knowledge development. In contrast, in Shintani and R. Ellis' study, the WCF groups' improvements on the timed writing task during the immediate posttest were not considered evidence of learners' implicit knowledge development. This, thus, also attribute to their contradictory results.

In effect, the conflicting results can be expected given that the two studies targeted different linguistic structures (indefinite vs. definite and indefinite articles), proficiency levels (low intermediate vs. intermediate), and educational settings (ESL vs. EFL). As mentioned earlier, the finding of Kang and Han's (2015) meta-analysis has also shown that learners' proficiency level is the strongest moderator of the overall WCF efficacy. Besides, differences in measurements and analysis methods used in each study can considerably attribute to the incongruent results limiting the generalizability of the findings. Most importantly, some limitations presented in these two studies seem to markedly affect the reliability and validity of their findings, making it difficult to draw any firm conclusions from their data.

The first limitation concerns the one shot WCF treatment design the two studies adopted. Since the two studies aimed to investigate the effect of WCF on the development of learners' explicit and implicit knowledge (long-term gains), the one-shot treatment the two studies provided may not be sufficient for the learners to fully benefit from the

feedback since it takes time for the effect of WCF to emerge and for the explicit and implicit knowledge to develop (DeKeyser, 2015; Paradis, 2009; Suzuki & DeKeyser, 2015, 2017). A longitudinal design in which multiple feedback treatment sessions are provided seems to be a more valid approach under this agenda.

In addition to the one-shot treatment issue, the use of the timed narrative writing task as a single tool measuring learners' implicit knowledge in Shintani and R. Ellis' study may, to some extent, compromise the reliability of the yielded findings, since the construct validity of the timed writing task as measure of implicit knowledge is still up for debate. Most scholars do not view timed writing task as a valid tool. Polio (2012) points out that even under a time constraint, writing task still allows learners to tap into their explicit knowledge. That is, even though the task may bias the use of implicit knowledge (the task focus is on meaning and there is a time limit), learners still, to some extent, rely on their explicit knowledge when writing. Shintani and R. Ellis also acknowledged this shortcoming in their study stating that:

> writing tasks – even when pressured and learners are primarily focused on meaning – cannot afford a convincing measure of implicit knowledge given that they allow for controlled processing and monitoring for accuracy. In future studies, therefore, it may be necessary to make use of the kinds of instruments that SLA researchers have employed to measure implicit knowledge (e.g., oral elicited imitation or free oral production). (Shintani & R. Ellis, 2013, p. 301)

Further, the fact that Shintani and R. Ellis only used the timed writing task as the sole measurement measuring learners' implicit knowledge, instead of using multiple measurements, seems to affect the credibility and reliability of their findings to some extent. This is because the data obtained from a single measurement may be more susceptible to errors. The researchers can only assume that the obtained data represent the truth without a chance to cross-check their data for any analysis errors with the data obtained from other sources. Thus, using multiple measurements so that data obtained from multiple sources can be triangulated and compared seems to be a more valid design, increasing the overall credibility and reliability of the yielded findings, particularly in the case where the validity of the chosen test is still up for discussion.

In Rezazadeh et al.'s (2015) study, the validity and reliability of their overall findings are also subjected to debate. First of all, even though this study employed the speeded dictation test as another measure of implicit knowledge in addition to the timed writing task, the construct validity of the speeded dictation task, similar to the timed writing task, is still questionable. To the best of my knowledge, the test's validity has never been attested. Hence, the results yielded from these two tests may not appear to be very robust. Further, since Rezazadeh et al. analyzed the combined scores of the definite and indefinite articles, instead of analyzing the definite apart from indefinite scores, this made their findings unclear as to whether WCF was truly effective in aiding the development of both articles usage as claimed. It could be that WCF was only effective in helping learners develop their definite (or indefinite) article knowledge but because the combined scores were used for the analysis, the results were interpreted as WCF was effective for both.

In an attempt to overcome the limitations of previous studies, recent work by Nemati, Alavi, and Mohebbi (2019) and Shooshtari, Vahdat, and Negahi (2019) adopted a more longitudinal design providing leaners with multiple WCF treatment episodes. Nonetheless, in Nemati et al.'s (2019) study, the issue concerning the implicit knowledge measurement persists.

In Nemati et al.'s (2019) study, the effects of direct focused and indirect focused WCF on the development of learners' explicit and implicit knowledge of English simple past tense (regular past tense form -ed) were explored. Eighty-seven Iranian EFL beginners took part in this study and were assigned into the control (received no feedback), direct focused or indirect focused WCF group. Learners were then asked to complete the pretest, posttest, delayed posttest, which consisted of a narrative writing task, text summary task, timed grammaticality judgment test (TGJT), untimed grammaticality judgment test (UGJT), and metalinguistic knowledge test (MKT). The narrative writing and text summary tasks were used to measure learners' improved accuracy. The UGJT and MKT were used to measure learners' explicit knowledge and the TGJT was used to measure learners' implicit knowledge. Learners in this study attended five WCF treatment sessions. In each session, each of them was required to write one text summary task. It was only during the last session that each learner revised his/her

own text instead of writing the new one. Overall, the learners in this study wrote four new texts and received four feedback treatment episodes during treatment sessions.

The study's findings revealed that both WCF groups outperformed the control group on both the narrative writing and text summary tasks, indicating the effectiveness of both types of WCF in enhancing beginner learner's grammatical accuracy. The direct focused WCF group also outperformed the indirect focused WCF group on these two tasks. Regarding learners' explicit and implicit knowledge development, the results showed that only the direct focused WCF group outperformed the control and indirect focused groups on the UGJT, MKT, and TGJT and there were no significant differences between the control and indirect focused groups' test performance. Based on the results, Nemati et al. concluded that only direct focused WCF effectively helped develop beginner learners' explicit and implicit knowledge. The finding aligns with that of Rezazadeh et al.'s (2015) study which also found direct focused WCF effective in helping learners develop their explicit and implicit knowledge. Indirect focused WCF, on the other hand, was ineffective in this study because the beginner learners participated in this study may not have adequate explicit knowledge to self-correct their own errors and therefore could not make any significant improvement on their explicit and implicit knowledge tests. The study also endorses the strong interface position contending the possibility that explicit knowledge can become implicit knowledge through meaningful practice and conscious reflection and examination.

While Nemati et al.'s (2019) study provided multiple WCF treatment episodes in place of the predominant one-shot treatment (Kang & Han, 2015; Liu & Brown, 2015) responding to the previous call for longitudinal investigation under this research agenda, the validity of the TGJT employed as measure of implicit knowledge in this study is still debatable (Plonsky, Marsden, Crowther, Gass, & Spinner, 2020). Even though some empirical evidence lends support for the construct validity of the TGJT test, (see Bowles, 2011; R. Ellis, 2005; Erçetin & Alptekin, 2013; Godfroid, Loewen, Jung, Park, Gass, & R. Ellis, 2015; Gutiérrez, 2013; Zhang 2015), others do not view TGJT as valid measure of implicit knowledge (see Suzuki, 2017; Vafaee, Suzuki & Kachisnke, 2017). Thus, the employment of the TGJT test as a sole measure of learners' implicit knowledge might pose a potential threat to the validity of the study's findings as explained earlier. Nemati

et al. (2019) acknowledged this limitation of their study and asserted that other fine-grained measures of implicit knowledge should be validated for future usage. The instrument issue, however, seems to be ameliorated in Shooshtari et al.'s (2019) investigation where a TGJT was employed together with an oral imitation test (OIT), a better measure of learners' implicit knowledge endorsed by many SLA researchers (e.g., Bowles, 2011; R. Ellis, 2005; Erçetin & Alptekin, 2013; Godfroid et al., 2015; Gutiérrez, 2013; Kim & Nam, 2017; Zhang 2015).

In Shooshtari et al.'s (2019) study, the effects of direct and indirect unfocused WCF on the explicit and implicit knowledge of 17 grammatical structures were examined. Ninety EFL learners[19]participated in the study. They were randomly assigned into one of the three groups: control (receiving general comments on the organization of their writing), direct unfocused, and indirect unfocused WCF. A TGJT and OIT were used to measure learners' implicit knowledge, whereas an MKT and UGJT were used to assess learners' explicit knowledge. A writing test was employed as measure of learners' written accuracy. During the study, learners completed the pretest, attended 12 WCF treatment sessions and completed the posttest and delayed posttest, respectively. Overall, each learner in this study was required to complete 12 descriptive essays during treatment sessions.

The results of the study showed that both direct and indirect unfocused WCF were effective in promoting the development of learners' explicit and implicit knowledge as well as increasing learners' grammatical accuracy. There were also no significant differences found between these two WCF types. The study also assumed the interface position.

All in all, the findings from the four studies conducted under this agenda provide some empirical evidence in support of the role WCF plays in L2 development, even though the results concerning which type of WCF can efficiently aid the development are still conflicting. In Shintani and R. Ellis' (2013) study, the findings demonstrated that ME was effective for explicit knowledge development, while direct focused WCF had null

---

[19] The authors did not specify the proficiency level of the learners but stated that these learners took the Michigan Examination for the Certificate of Proficiency in English (ECPE) and scored one standard deviation below the mean scores of the initial 380 volunteer pool.

effect on the development. However, Rezazadeh et al.'s (2015) subsequent study found the positive effects of both ME and direct focused WCF on the development of explicit and implicit knowledge even though ME had shown to have a more durable effect than direct focused WCF. The finding of Nemati et al.'s (2019) study corroborates the finding of Rezazadeh et al. (2015) affirming the effectiveness of direct focused WCF on the development of explicit and implicit knowledge. This study, nonetheless, did not find any effect of indirect focused WCF on learners' L2 development. In Shooshtari et al.'s (2019) recent study, the authors also found positive effects of direct and indirect unfocused WCF on learners' L2 development. Differences in target linguistic structures, research designs, methodology and measurements each study adopted are posited to largely attribute to the contradictory results. Learner difference factors may also have influenced the results to a greater or lesser extent, depending on the educational and instructional contexts each study situated in. Given that the findings are still mixed, and research conducted under this agenda is limited restricting the generalizability of the findings, it is within this line of inquiry that the current study aimed to pursue to shed more light on this under-researched topic moving the field forward.

2.2.4. Summary

In previous sections, a number of studies have attested the effectiveness of WCF on learners' L2 development. These studies provide evidence in support of the facilitative role WCF plays in L2 development and affirm the advantage of providing learners with correction over the no correction condition. Nonetheless, even though it seems to be warranted that the provision of WCF is beneficial for learners' L2 development, the main pedagogical and theoretical questions concerning the efficacy of WCF remain unanswered. Pedagogically, a firm conclusion regarding which type of WCF is most beneficial for L2 learning could not be made, even though based on Kang and Han's (2015) meta-analysis, direct and focused WCF appear to be most effective given that they have shown to have greater effects on L2 learning than indirect and unfocused WCF (based on an effect size index), albeit the difference is non-significant.

71

Concerning the direct/indirect WCF dichotomy, some studies reported the superiority of direct over indirect WCF (e.g., Bitchener & Knoch, 2010a, 2010b; Chandler, 2003; Van Bueningen et al., 2008), while others found the advantage of indirect over direct WCF (e.g., Lalande, 1982; Storch & Wigglesworth, 2010; Tan & Manochphinyo, 2017). Some did not find any significant differences between the two (e.g., Ferris, 2006; Robb et al., 1986; Semke, 1984; Van Bueningen et al., 2012; Vyatkina, 2010). However, in recent years, researchers (e.g., Benson & DeKeyser, 2019; Bitchener & Knoch, 2010a, 2010b; Ferris et al., 2013; Nemati et al., 2019; Shintani et al., 2014; Stefanou & Révész, 2015; Rezazadeh et al., 2015; Rummel, 2014; Suzuki et al., 2019) seem to unanimously agree that direct WCF is more effective in aiding learners' L2 development. This is because, empirically, an increasing amount of recent research provides more evidence in support of the effectiveness of direct WCF in improving learners' grammatical accuracy. These studies also found that the effect of direct WCF has shown to be more durable. Theoretically, from an SLA perspective, direct WCF could be more efficient given that it promptly provides unambiguous comprehensible information (i.e., positive evidence) about the target structure to learners and learners can instantly tally the input it provides into their cognitive systems (Bitchener & Knoch, 2010a, 2010b; R. Ellis, 2009a; Ferris et al., 2013; Manchón, 2011; Van Beuningan et al., 2012). In contrast, learners cannot immediately internalize indirect WCF because they need to spend some time figuring out their own correction. Such a delay in an uptake of the information may leave them benefit less from the given feedback.

In addition, direct WCF better promotes L2 learning than indirect WCF because direct WCF provides learners with both positive evidence (i.e., correct linguistic forms) and negative evidence (i.e., indication of unacceptable information). Conversely, indirect WCF can only provide learners with negative evidence.

Similarly, it is still inconclusive as to which type of feedback, focused or unfocused WCF, is more efficient pedagogically. Apart from the mixed findings, a limited number of comparative studies comparing these two types of WCF in one single study (e.g., R. Ellis et al., 2008; Farrokhi & Sattapour, 2012; Frear & Chiu, 2015; Saeb, 2014; Sheen et al., 2009) and variations in their research designs, methodology and learners' factors make it more difficult for the results to be generalized. Despite that, from

most researchers' views, focused WCF appears to be more promising than unfocused WCF since it responds well to SLA theories (Lee, 2019), providing that its narrow focus on limited ranges of target structures makes the target structures become more salient and that facilitates learners' noticing and L2 acquisition, respectively (Bitchener, 2008; R. Ellis, 2008a; Lee, 2019; Sheen et al., 2009). On the contrary, unfocused WCF may require a high attentional load from learners (Frear & Chiu, 2015; Mao & Lee, 2020) and therefore has a high possibility to overload learners' limited cognitive capacities (Lee, 2019; Sheen et al., 2009) leading to unsuccessful uptake. Lee (2019) also criticizes that unfocused WCF may not be effective for learning since all errors are corrected without considering learners' proficiency levels and their developmental readiness.

As illustrated above, to date, pedagogically, a firm conclusion regarding which type of WCF is most effective could not be borne out of the current data just yet. Further, theoretically, L2 writing researchers are still unable to address the key question as to whether WCF only contributes to learners' explicit knowledge or it also contributes to learners' implicit knowledge. A limited number of empirical studies under this line of research (i.e., Nemati et al., 2019; Rezazadeh et al., 2015; Shintani & R. Ellis, 2013; Shooshtari et al., 2019) is a primary reason why the conclusive answer could not be made. As Polio (2012) suggests, it is important to establish a research agenda investigating the role of WCF plays in the development of L2 learners' explicit and implicit knowledge. A lack of empirical research under this agenda indicates the need for further investigation in order to fill in this gap presented in the literature.

Future studies, however, need to make an effort to avoid the limitations of previous research, such as the use of revised tasks as measure of learners' progress, the use of one-shot treatment, to name just a few. In particular, under the explicit-implicit agenda, a more valid and fine-grained measure of implicit knowledge should be employed. The use of multiple measures in this respect can also prove beneficial. Of equal importance here is that future WCF research should begin to explore the moderating effects of a wider ranges of learner difference factors, since they could impact the extent to which learners benefit from WCF and therefore may also account for the conflicting results presenting in the field (Bitchener & Storch, 2016; Li & Roshan, 2019; Storch, 2018). Thus, investigating the efficacy of WCF in relation to learner difference

effect is beneficial as it might lead to more robust and consistent results across studies moving the field forward (Bitchener & Storch, 2016; Li & Roshan, 2019, Storch, 2018). In what follows, the mediating effects of learner differences in working memory capacity on L2 learning and the efficacy of WCF, another topic investigated in this study, are discussed in more detail.

2.3. Learner differences in working memory capacity

In recent years, there has been a growing interest in the role of individual differences in SLA. Researchers are intrigued to empirically explore these learner difference variables since they are posited to account for differential success among L2 learners. Among these learner difference factors, working memory capacity (WMC) has recently been identified as one key factor that has influential impact on learners' L2 learning success (see Linck, Osthus, Koeth & Bunting, 2014; Wen, Mota, & McNeill, 2015 for a full review). However, despite a growing interest in the relationships between WMC and different aspects of L2 learning, and the fact that an extensive amount of research was carried out to explore the extent to which WMC mediates the effectiveness of oral feedback (e.g., Goo, 2012; Kim, Payant, & Pearson, 2015; Li, 2013; Mackey, Philp, Egi, Fujii, & Tatsumi, 2002; Mackey, Adams, Stafford, & Winke, 2010; Mackey & Sachs, 2012; Révész, 2012; Sagarra, 2007; Sagarra & Abbuhl, 2013; Sanz, Lin, Lado, Stafford, & Bowden, 2016; Trofimovich, Amnar, & Gatbonton, 2007; Yilmaz, 2013; Yilmaz & Granena, 2019), to date, to the best of my knowledge, there is only one study, i.e., Li and Roshan (2019), attempting to investigate the extent to which WMC mediates the effectiveness of WCF and its outcomes. Consequently, to provide further insight on this relatively under-investigated research avenue, the current study aimed to investigate the extent to which WMC may moderate the effect of WCF on learners' L2 learning outcomes. In the sections that follow, the definitions and description of working memory model are reviewed, followed by discussion of current research findings presented in the WCF literature.

2.3.1. Working memory and second language learning

Baddeley (2003) defines "working memory" (WM) as "the temporary storage and manipulation of information that is assumed to be necessary for a wide range of complex cognitive activities" (p. 189). Others view WM as a multicomponent cognitive system responsible for the control, regulation and active maintenance of information during ongoing processes and/or distractions (e.g., Conway, Kane, Bunting, Hambrick, Wilhelm, & Engle, 2005; Conway, Jarrold, Kane, Miyake & Towse, 2007; Juffs & Harrington, 2011; Linck et al., 2014; Williams, 2012; Unsworth, Schrock, & Engle, 2004). Based on the definitions above, it can be inferred that the main functions of WM involve storing and processing information (Juffs & Harrington, 2011). The most influential and widely accepted WM model is the model first proposed by Baddeley and Hitch (1974, also Baddeley, 1986). Based on Baddeley and Hitch's (1974) classic model, this multicomponent model can be divided into two separable subsystems: the storage-based system (i.e., slave systems) and the central executive system. The storage-based system constitutes of the *visual spatial WM* (i.e., visuo-spatial sketchpad), responsible for maintaining visual and spatial information, and *verbal WM* (i.e., phonological loop), responsible for retaining verbal representations in the domain (Baddeley, 2000, 2007, 2010; Baddeley & Hitch, 1974). On the other hand, the central executive system regulates attentional resources and information flow between the slave systems and long-term memory stores (Baddeley, 1986; Baddeley & Hitch, 1974). Its most significant functions are to control the attention needed to maintain focus and inhibit distractions which may interfere with task performance (Baddeley, 2000, 2003, 2007, 2010; Cowan, 2005; Engle & Kane, 2004; Juffs & Harrington, 2011; Linck et al., 2014). Later, Baddeley (2000, also see Baddeley, 2017) added the fourth subsystem component called the *episodic buffer* into the original model. The episodic buffer is the place where information from the phonological loop, visual-spatial sketchpad, and long-term memory are temporarily stored and processed into a single episode. The episodic buffer is also controlled by the central executive system.

Figure 2.1 Baddeley's (2000) Working Memory Model
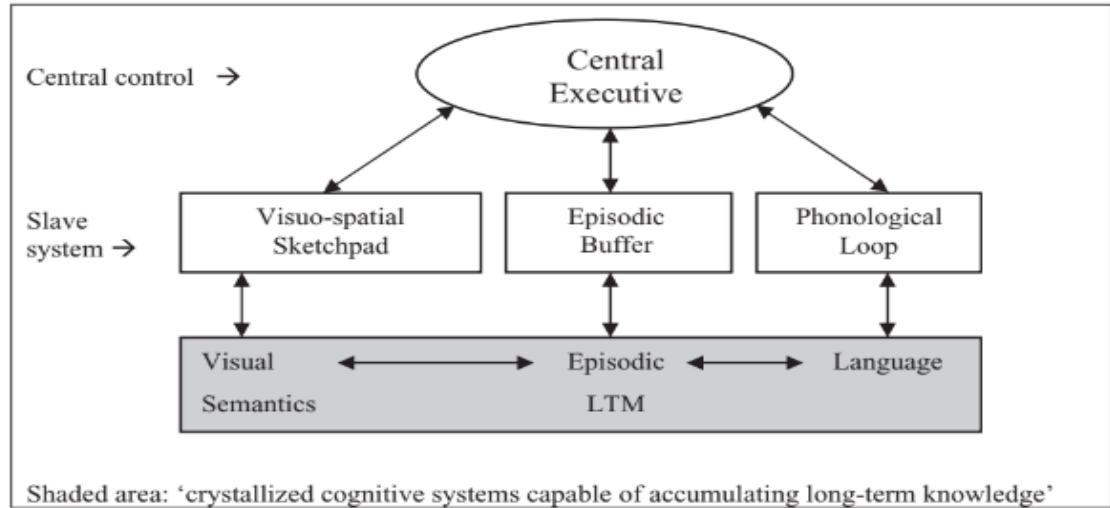
Nonetheless, in contrast to long-term memory, WM is limited in terms of the amount of information it can maintain and process each time. This limited capacity is presumed to constrain individuals' cognitive performance (Conway et al., 2007; Linck et al., 2014). This limited capacity of individual's working memory to actively maintain and process a certain amount of information available for direct access at one time is referred to as "working memory capacity" (WMC) (Oberauer, Martin Süβ, Wilhelm, & Sander's, 2007). Individual WMC can be measured either through "simple span tasks" (i.e., phonological short-term memory tasks), which assess an individual's ability to store and rehearse information (i.e., storage capacity only), or through "complex span tasks" (i.e., complex working memory tasks), which assess an individual's ability to simultaneously store and process information (i.e., processing plus storage capacity) (Juffs & Harrington, 2011; Linck et al., 2014). Nevertheless, the finding of Linck et al.'s (2014) WM meta-analysis has demonstrated that complex WM measures are stronger predictors of L2 outcomes than those phonological short-term memory measures. Their finding is also in line with the finding of Daneman and Merickle's (1996) meta-analysis on WM and L1 reading comprehension.

In cognitive psychology, WMC is considered a critical construct as a substantial body of empirical research has shown that it is a reliable predictor of a wide range of learning behaviors and complex cognitive task performance (Juffs & Harrington, 2011;

Linck et al., 2014). Precisely, the higher WMC individuals have shown to be better at attentional control (Kane, Bleckley, Conway & Engle, 2001), following direction (Engle, Carullo & Collins, 1991), reasoning, and general fluid intelligence (e.g., Ackerman, Beier, & Boyle, 2005; Conway, Cowan, Bunting, Therriault, & Minkoff, 2002; Conway, Kane, & Engle, 2003), multitasking (Hambrick, Oswald, Darowski, Rench & Brou, 2010), to name just a few. Likewise, there is also a large amount of research evidence indicating significant correlations between WMC and L2 comprehension (e.g., Adams & Mohammadtaghi, 2014; Alptekin & Erçetin, 2010; Harrington & Sawyer, 1992; Leeser, 2007; Walter, 2004), grammar learning (e.g., Martin & N. Ellis, 2012; Robinson, 1997, 2005; Verhagen & Leseman, 2016), morphosyntactic processing (e.g., Juffs, 2004, 2005; Miyake & Friedman, 1998), oral feedback processing (e.g., Goo, 2012; Kim et al., 2015; Li, 2013; Mackey et al., 2002; Mackey et al., 2010; Mackey & Sachs, 2012; Révész, 2012; Sagarra, 2007; Sagarra & Abbuhl, 2013; Sanz et al., 2016; Yilmaz & Granena, 2019), vocabulary learning (e.g., N. Ellis, & Bywater, 2004; Martin & N. Ellis, 2012; Masoura & Gathercole, 2005; Ruiz, Rebuschat & Meurers, 2019; Speciale,; Wen, 2016), writing (e.g., Baoshu & Chuanbi, 2015; Baoshu & Luo, 2012; Bergsleithner, 2010; Kormos & Sáfár, 2008; Michel, Kormos, Brunfaut, & Ratajczak, 2019; Révész, Michel, & Lee, 2017; Service, 1992; Zabihi, 2018) etc. Nevertheless, even though a large amount of empirical evidence has suggested that WMC has influential impact on various aspects of L2 learning including the extent to which learners benefit from oral feedback (see also Williams, 2012, for a review of WM and L1-L2 learning), thus far, only one WCF study, i.e., Li and Roshan's (2019) study, has been conducted to investigate the potential moderating effect of WMC on the efficacy of WCF. The findings of the study are discussed in the subsequent section.

2.3.2. Research on working memory and the effectiveness of written corrective feedback

As previously mentioned, a large amount of research has been conducted to determine if WMC has any moderating effect on the effectiveness of oral feedback, and most found that the effect of oral feedback, to certain extent, is moderated by WMC. The findings on oral feedback studies implicate the possibility that WMC may also moderate the extent to which learners benefit from WCF (Li & Roshan, 2019). Specifically, the findings of previous oral feedback studies have demonstrated that WMC has considerable influence on feedback noticing, processing, and storing processes (e.g., Goo, 2012; Kim et al., 2015; Li, 2013; Mackey et al., 2002, 2010; Mackey & Sachs, 2012; Révész, 2012; Sagarra, 2007; Sagarra & Abbuhl, 2013; Sanz et al., 2016). With respect to feedback noticing, WMC has been shown to have an influence on the extent to which learners notice the given feedback (Goo, 2012; Mackey et al., 2002, 2010; Sagarra, 2007). This is because for learners to notice the given feedback and the gaps between their erroneous production and the target-like structures provided by feedback, they first need to direct their conscious attention regulated by the central executive system of WM to the given feedback. In other words, conscious attention is necessary for the given feedback to be noticed. Since the central executive component of WM controls the allocation of attentional resources, WMC seems to inevitably link to the extent to which learners notice the given feedback (Skehan, 2002, 2016). The finding in Mackey et al.'s (2002) study, for example, revealed a significant relationship between WMC and the extent to which learners noticed the given feedback, with high WMC learners tended to notice more feedback than those of lower WMC.

Apart from feedback noticing, it is posited that learners rely on their WM (both the central executive and slave systems), when making a cognitive comparison (comparing new input with what they already archive in their systems), as well as when processing and storing new information into their systems (Li, 2013; Sagarra & Abbuhl, 2013; Sanz et al., 2016). All in all, previous research evidence on WMC and oral feedback has suggested that WMC has influential effect on the extent to which learners benefit from feedback. Given similar nature of oral and written feedback, it is conceivable that WMC may also have a moderating effect on the efficacy of WCF. Despite that fact,

to date, only Li and Roshan's (2019) study has examined the relationship between WMC and WCF, leaving a big gap in the L2 writing literature.

In Li and Roshan's (2019) study, the associations between learners' WMC, both complex WM and phonological short-term memory, and the efficacy of different types of WCF, namely direct WCF, direct WCF plus revision[20], metalinguistic explanation (ME), metalinguistic explanation (ME) plus revision, were examined. English passive voice was the target structure of this study. Seventy-nine intermediate EFL learners participated in this study. They were randomly assigned into one of the four feedback groups: direct WCF, direct WCF plus revision, ME, and ME plus revision. In the first week, learners completed the pretest writing task. In the second week, learners received feedback on their pretest writing tasks and immediately after reviewing their feedback, they completed the immediate posttest writing tasks. In the third week, learners completed complex WM (i.e., a reading span test) and phonological short-term memory (i.e., a non-word span test) tests. In the fourth week, learners completed the delayed posttest writing tasks. The results revealed that 1) complex WM was a positive predictor of the effectiveness of ME and direct WCF plus revision and 2) phonological short-term memory was a negative predictor of the effectiveness of direct WCF plus revision.

Regarding the positive relationship between complex WM and ME, Li and Roshan (2019) explained that learners needed to heavily rely on their complex WM to process ME because they needed to comprehend the ME handouts and keep the information instantly accessible in their WM so that they can use the information to detect errors in their texts and corrected their errors accordingly. On the contrary, with direct WCF, learners did not need to identify nor correct their own errors as the errors were already identified and the correct forms were provided for them and this was postulated to alleviate learners' reliance on their complex WM operations. This explains why the study did not find significant relationship between WMC and direct WCF. It was posited that WM effect will only be evident when there is the demand for deep cognitive processing.

However, Li and Roshan (2019) explained that the reason why complex WM was

---

[20] The groups that were required to revise their work following the feedback treatment session (i.e., direct WCF plus revision and ME plus revision) had to rewrite their work without access to the given feedback.

a positive predictor of direct WCF plus revision was because revision processes (i.e., rewriting the whole text), similar to writing processes (see Kellogg, 1996 for a full review), likely tax heavily on complex WM operations such as during the retrieval of stored information, during the review and editing of the text, for instance. Accordingly, because the leaners in the direct WCF plus revision group needed not only to review the given feedback but also to revise their texts based on the feedback, the learners in this group then expended more cognitive resources than the learners in the direct WCF only group. However, while complex WM positively correlated with direct WCF plus revision, the opposite was true for the ME plus revision. Li and Roshan (2019) offered two possible explanations explaining why complex WM was not predictive of the effect of ME plus revision: "(1) rule explanation alleviated learners' cognitive burden (Sanz et al., 2016) and (2) production practice neutralized the effects of cognitive ability (Erlam,2005)" (p. 11). Nonetheless, Li and Roshan (2019) did not seem to agree that rule explanation alleviated learners' cognitive load since complex WM was previously found to be a positive predictor of the performance of the ME (without revision) group. Concurring more with Erlam's (2005) explanation, Li and Roshan (2019) concluded that it was more probable that under deductive learning condition[21], production practice (i.e., essay revision) may neutralize the effects of cognitive ability and that explained why complex WM was not a significant predictor of the effect of ME plus revision.

With regard to the negative relationship between phonological short-term memory and direct WCF plus revision, Li and Roshan (2019) attributed the result to the possibility that learners with high phonological short-term memory may memorize direct WCF as discrete items (because they can memorize many); therefore, they may not attempt to comprehend the underlying grammatical rules of the target structures leading to their lower gain scores. On the contrary, given that learners with lower phonological short-term

---

[21] The authors believed that the learning condition of the ME plus revision group (receiving rule explanation followed by revision) was similar to deductive learning condition where grammatical rule is first introduced in class followed by practicing activities. And because a substantial amount of previous research did not find the moderating effects of learners' cognitive abilities within deductive learning condition, this may explain why the study did not find significant relationship between WMC and ME plus revision. That is, the authors speculated that the learning condition of the ME plus revision group may neutralize WM effect.

memory can only memorize a smaller amount of feedback each time, it is probable that these learners may be able to extrapolate the underlying grammatical rules from the limited amount of input they can remember more easily (i.e., they were not overwhelmed by the received input), and as a result, they can achieve higher gain scores. Li and Roshan (2019) further argued that their exposition aligned with Newport's (1990) 'less is more' hypothesis, proposing that because of children's lower memory capacities, children are able to meticulously analyze the received input making them better language learners than adults.

Taken together, Li and Roshan (2019) suggest that complex WM is a key predictor of the efficacy of WCF. Meanwhile, phonological short-term memory seems to have inconsequential effect on the effectiveness of WCF. The results confirm Kellogg's (1996, see also Kellogg et al., 2013) model of WM in writing which proposes a more important role complex WM plays during the monitoring stage of writing (involving the application of feedback)[22] and corroborate Linck et al.'s (2014) meta-analysis finding which suggests greater impact complex WM has on L2 learning compared to that of phonological short-term memory. Nonetheless, since this study is the only study investigating this association within WCF literature, the results yielded from the study cannot be viewed as conclusive. More comparable evidence is needed for a firm conclusion to be made.

---

[22] According to Kellogg's (1996, also Kellogg et al., 2013) model of WM in writing, feedback prompts learners to review and edit their written texts during the monitoring stage and the reviewing and editing processes make a demand on both central executive and phonological loop resources, i.e., complex working memory.

2.3.3. Summary

As aforementioned, within the field of cognitive psychology as well as L2 learning, WMC has been implicated as a key predictor of a wide range of learning behaviors, including L2 learning (Juffs & Harrington, 2011; Linck et al., 2014; Ruiz et al., 2019; Wen et al., 2015). A plethora of studies also found that WMC is predictive of the efficacy of oral feedback. Despite that fact, within the WCF field, until recently, only one study, i.e., Li and Roshan (2019), has investigated the association between WMC and WCF. In this study, Li and Roshan (2019) examined the effects of complex WM and phonological short-term memory on the effectiveness of different types of WCF. The results of the study showed that only complex WM was a positive predictor of the effectiveness of WCF, both direct WCF and ME. The finding aligns with the finding of Linck et al.'s (2014) meta-analysis which suggests the more important role complex WM plays in various aspects of L2 learning. However, since there are no other comparable studies to compare these results with, a firm conclusion could not be reached. More research under this agenda is apparently needed. To address this gap presented in the literature shedding more light on this under-investigated topic, the current study aimed to further explore this relationship.

2.4. Goals and research questions

As previously mentioned, most WCF research is conducted within a pedagogical framework rather than an SLA one, leading to a paucity of theoretical-based research, which in turn leaves a big gap in WCF literature. Even though few investigations attempted to discover whether WCF is facilitative to the development of learners' L2 explicit and implicit knowledge (i.e., Nemati et al., 2019; Rezazadeh et al., 2015; Shintani & R. Ellis, 2013; Shooshtari et al., 2019), a conclusion could not be reached. Since the primary goal of SLA is to understand factors contributing to learners' acquisition, it is, thus, crucial to find out if WCF, a form of negative evidence and form-focused instruction, can contribute to the development of learners' L2 explicit and implicit knowledge (i.e., acquisition). If WCF truly has positive effects on L2 development, theoretically, this implies that (1) positive evidence is not the only requisite input for successful L2 acquisition and (2) the provision of negative evidence or form-focused instruction is beneficial as it may speed up learners' acquisitional processes and/or prevent learners from premature language fossilization. Pedagogically, the finding could help practitioners make a sound decision regarding which type of feedback they should adopt and the extent to which it should be used in L2 composition classrooms.

Apart from a paucity of theoretical-based WCF research as illustrated, the mediating effects of learner differences on the effectiveness of WCF are also largely overlooked within the WCF context and this, to some extent, attributes to the mixed findings presented in the field (Bitchener, 2012; Bitchener & Storch, 2016; Mao & Lee, 2020; Storch, 2018). Thus, another goal of this current study was to explore the potential moderating effect of these learner difference factors in relation to the efficacy of WCF. Among a plethora of individual difference factors attributing to the disparity level of L2 learning success, the current study sought to investigate the extent to which the effectiveness of WCF could be moderated by variations in learners' WMC because learner differences in WMC have recently been posited to have considerable influential effects on various aspects of L2 learning (Linck et al., 2014; Wen et al., 2015). The results found in a substantial body of oral feedback studies and that of Li and Roshan (2019), the only study investigating the relationship between WMC and WCF to date,

also corroborate the possibility that learner differences in WMC might mediate the efficacy of WCF. Thus, it is vital to expand the existing research discovering whether or not WMC does moderate the efficacy of WCF so that the role of WMC in WCF processing could be better comprehended advancing the field forward. The findings are expected to be useful for both theory and practice.

Based on the aforementioned gaps presented in WCF literature, the aim of this study is twofold: (1) to examine if WCF has any effect on learners' explicit and implicit knowledge development, and (2) to explore if there is a relationship between WMC and learners' L2 learning outcomes resulting from WCF.

Precisely, first of all, this study attempted to investigate the efficacy of WCF, specifically direct *"focused"* WCF (correcting one to a few types of grammatical errors) and direct "*unfocused*" WCF (correcting a wide variety of grammatical errors), and how they may affect learners' explicit and implicit knowledge over time. These two types of WCF were selected for the study because both types of WCF have been shown to be effective in promoting learners' L2 development (see section 2.2.3.1.), even though when comparing these two types of WCF in one single study, the results regarding which one of them is more effective are still inconclusive. Indeed, to date, only a small number of studies have actually been conducted to compare the effectiveness of these two types of WCF in one single study (e.g., R. Ellis et al., 2008; Farrokhi & Sattapour, 2012; Frear & Chiu, 2015; Saeb, 2014; Sheen et al., 2009) and the results are incongruent. In addition, none of these studies examined whether these two types of WCF affected learners' explicit and implicit knowledge development. And even though a few subsequent studies (i.e., Nemati et al., 2019; Rezazadeh et al., 2015; Shintani & R. Ellis, 2013; Shooshtari et al., 2019) had attempted to investigate the effect of WCF on learners' explicit and implicit development, thus far, these studies also yielded mixed results as noted earlier. To shed more light on this topic and expand the existing research, this study proposed to investigate how these two types of WCF may potentially affect learners' explicit and implicit knowledge development and explore whether there is any difference in the effects these two types of WCF have on learners' learning outcomes.

Secondly, this study attempted to examine whether there is a relationship between learner differences in WMC and the effectiveness of WCF. Learner differences in WMC

are addressed in this study because (1) a plethora of oral feedback studies have suggested that WMC mediates the extent to which L2 learners benefit from oral feedback and this implicates a high probability that WMC might also moderate the efficacy of WCF, and (2) there are insufficient WCF studies investigating the relationship between WMC and learners' L2 achievement resulting from WCF, posing a big gap in WCF literature. To address this lack of attention towards the role of WMC in WCF context and shed more light on this unattended area of WCF research, this potential relationship was also investigated within the scope of this study.

Research Questions

RQ1: Do direct focused and unfocused written corrective feedback have any effect on the development of L2 learners' explicit knowledge of English plurals?

RQ2: Do direct focused and unfocused written corrective feedback have any effect on the development of L2 learners' implicit knowledge of English plurals?

RQ3: Is there any difference in the effects of direct focused and unfocused written corrective feedback on the development of L2 learners' explicit and implicit knowledge of English plurals?

RQ4: Does working memory capacity moderate the extent to which L2 learners benefit from written corrective feedback?

Regarding research question 1 and 2, it was predicted that both types of WCF might have positive effect on the development of learners' explicit and implicit knowledge, considering that a large number of WCF studies have shown that these two types of WCF are effective in helping learners improve their grammatical accuracy over time. Importantly, WCF research conducted under "feedback for acquisition" aspect has also demonstrated the effectiveness of these two types of WCF in helping learners develop their explicit and implicit knowledge, even though none of these studies compared these two types of WCF together in one single study. Based on previous empirical evidence presented in the literature, it was then hypothesized that these two types of WCF would be effective in helping learners in this study improve their explicit and implicit knowledge of English plurals.

Regarding research question 3, it was postulated that there might be differences in the effects of direct focused and direct unfocused WCF on learners' learning outcomes.

Given that a substantial amount of recent WCF research and WCF meta-analysis study have suggested the possibility that focused WCF might be more effective than unfocused WCF due to its narrower focus, it was then expected that learners in the direct focused WCF group might perform significantly better than those in the direct unfocused group.

For research question 4, it was predicted that WMC (i.e., complex WM) might have moderating effects on the extent to which these learners benefited from WCF, providing that learner differences in WMC have been found to be a key predictor of various L2 learning aspects. A large number of oral feedback studies also found the moderating effects of WMC on the efficacy of oral feedback. Similarly, positive associations between WMC and different types of WCF were also observed in Li and Roshan's (2019) recent WCF study. As a result, it was highly probable that WMC might have moderating effects on the efficacy of WCF within this study context as well.

2.5. Operationalizations
2.5.1. Second language development

In this study, second language (L2) development referred to changes in learners' L2 explicit and implicit knowledge, especially concerning an improvement in learners' English plural knowledge. Learners' explicit and implicit knowledge of English plurals were considered developed or improved when learners made fewer plural errors on the explicit and implicit knowledge measurements employed in this study. In other words, learners' progress in producing fewer plural errors on the explicit and implicit knowledge tests was accounted as evidence of learners' L2 development.

2.5.2. Explicit knowledge

In this study, explicit knowledge referred to the linguistic knowledge about a language that can be reported or described, following Anderson's (1993) and DeKeyser's (1998) definitions. That is, explicit knowledge is a type of knowledge that "is potentially reportable" (Anderson, 1993, p. 21). With explicit knowledge, learners know the rules of the language and how to apply the rules to produce accurate language, even though they

still make many grammatical errors when producing the language and cannot use the language with high automaticity (i.e., fast execution/processing speed) especially when they are under demanding conditions (e.g., time limitation).

2.5.3. Implicit knowledge

In this study, implicit knowledge referred to the linguistic knowledge of a language that enables learners to produce the target language with "complete fluency or spontaneity, rarely showing any errors" (DeKeyser, 2007, p. 95). However, as DeKeyser (2007) also points out, "even highly automatized behaviors are not 100% automatic" (p. 96). Hence, it should be emphasized here that acquiring this type of knowledge does not mean that learners will not make any linguistic errors any further, since errors could easily be made even when people converse in their native languages. Consequently, in this study, a low percentage of errors was not viewed as evidence against the acquisition of implicit knowledge. Learners who develop implicit knowledge of a particular structure are still able to make some errors on that structure, yet they are able to produce the structure with high automaticity even when they are under demanding conditions (e.g., time limitation).

Chapter III.
METHODOLOGY

This chapter begins with the description of the methodology followed by the participating population, target linguistic structure, treatment tasks, assessment tasks and scoring, procedure, and operationalizations sections, respectively.

3.1. Methodology
3.1.1. Pilot study

Before the main study was conducted, a pilot study was carried out to uncover the types of grammatical errors Thai low-intermediate learners often made when writing argumentative essays. The results were used to determine the types of grammatical structure the main study would focus on. In addition, the pilot study was conducted to assess the appropriateness of the test instruments developed to be used in the main study.

To uncover which types of grammatical errors this group of learners (i.e., low-intermediate) usually made when writing argumentative texts, a small-scale error analysis study was carried out. Thirty-six argumentative essays from 19 Thai low-intermediate learners studying at a public university in Thailand were collected for the analysis. The essays were analyzed to identify the most frequent grammatical errors this group of learners made when writing. The results of the error analysis are summarized in the target linguistic structure section (see section 3.1.1.3).

Regarding the assessment of the instruments, two versions of the three test instruments that is a metalinguistic knowledge test (MKT), oral elicited imitation test (OEIT) and untimed grammatical judgment test (UGJT), a Thai-version operation span task (OSpan), and a background and exit questionnaire were administered to 17 Thai low-intermediate learners.

After the assessment of these learners' performance, some modifications on the test instruments were made. First of all, based on item analyses (means and item difficulty-discrimination values), some items, both the main items targeting plural uses and distractors targeting other grammatical structures in the UGJT and OEIT which most

learners made most errors on (having low discrimination values) were eliminated[23]. After that, the tests were reassessed for internal consistency using Cronbach's alpha (see subsequent 3.1.2.5. assessment tasks and scoring section for more detail concerning the reliability coefficient values of these tests). Second, some wording in the background and exit questionnaires which appeared to be ambiguous to learners[24] were revised to make the survey clearer and easier to understand. Learners who did not give accurate information corresponding to the questions in their first attempt, were asked to read and complete the revised questionnaire again. The revision was completed when all learners were able to give accurate information that rightly matched the designated questions.

3.1.2. Current study
3.1.2.1. Research design

This experimental study employed a pre-post-delayed posttest design and was carried out over the course of a 9-month period. All 130 volunteering learners at the beginning of this study were asked to complete the pretest which also served as their placement test in order to see if all of them shared the same level of proficiency. The data of the learners whose pretest scores were not ranked within the 1.5 standard deviation range of the whole group scores (implying that they may not share the same level of proficiency with the majority), were not included in this study analysis[25].

Of the initial 130 volunteering learners from two intact classes, after the pretest assessment, only 109 learners were qualified to participate (they were assumed to share the same proficiency level). All qualified 109 learners were then randomly assigned to

---

[23] For assessment purposes, the tests administered in the pilot study were originally designed to have more test items than the ones expected to be used in the main study.
[24] Learners either came up to ask the researcher for clarification or gave wrong information on the survey form.
[25] These learners, however, were still allowed to take part in the study following the same procedure as other eligible learners. The reason why the researcher let them participate even though their data cannot be used was because they were from the same intact class as others whose data can be used, and the researcher did not want them to feel segregated. Moreover, it seemed unethical for these learners to do nothing for 40 minutes during class time while others had a chance to practice writing during treatment sessions.

one of the two experimental groups: direct *focused* WCF (n =37), direct *unfocused* WCF (n = 36), or a comparison group which received only content feedback (n = 36).

Learners from these three groups completed batteries of pre-post-delayed posttest tasks and took part in the 6-week treatment sessions. At each testing session, learners completed a timed argumentative essay writing task, MKT, OEIT, and UGJT. Two WM tests, the backward digit span and OSpan tasks, were also administered to the learners during the pretest and posttest. The design of the study is illustrated in Figure 3.1. below.

Figure 3.1 The design of the study

| Pretest | **Direct Focused** (n = 25) | **Direct Unfocused** (n = 23) | **Control** (n = 24) |

Timed Argumentative Essay Writing task, MKT, OEIT, UGJT, WM tests (Backward Digit span and OSpan), Background Questionnaire

Treatment (6 sessions): Direct Focused feedback / Direct Unfocused feedback / Content and Organization feedback

Posttest: Timed Argumentative Essay Writing task, MKT, OEIT, UGJT, WM tests (Backward Digit span and OSpan), Exit Questionnaire

Delayed Posttest: Timed Argumentative Essay Writing task, MKT, OEIT, UGJT tests

In addition to the quantitative method, qualitative interviews were also included in the study in order to explore potential variables that may mediate the overall results. Consequently, two learners from each group were randomly selected to take part in individual qualitative interviews during treatment and immediately after the posttest.

3.1.2.2. Participants

As mentioned earlier, the original pool of participants comprised 130 learners. However, given that those whose pretest scores were not in line with the majority of the group (above or below the 1.5 standard deviation range) were excluded[26] from the analysis, only 109 learners were qualified to participate at the onset of the study.

Nevertheless, since the study was conducted over the course of a 9-month period, the final sample size pool of the study only consisted of 75 learners (n = 75: direct focused WCF [n =26], direct unfocused WCF [n = 23], and control [n = 26]), who were able to complete the entire set of research procedure and measurements, that is completing the batteries of both language and WM tests during the pre-post-delayed posttests as well as attended the 6-week treatment sessions within the designated week. The data of the learners who were not able to complete all the tests and attended all the treatment sessions within the designated time were excluded from the analysis.

In part, the attrition is posited to be a result of the absence of learning incentive (i.e., extrinsic motivator) to motivate learners to attend all the experimental sessions. This is because whether the learners completed all the tests or attended all the treatment sessions or not did not have much influence on their course grades, even though their instructors who allowed the researcher to collect the data from their courses, agreed to give them extra credits (about 1% of their course grades) if they could complete the pretest, posttest and all treatment sessions[27]. As a result, after a few weeks, some learners decided not to attend the treatment sessions any further (see chapter 5 qualitative results for detailed analysis on this issue).

Moreover, since the pretest, posttest and delayed posttest were administered outside of learners' class time, this inevitably conflicted with some of the learners' normal learning schedules (all had different minors and selective courses to attend). This,

---

[26] These learners still attended the experiment, but their data were not included in the analysis.

[27] Since the delayed posttest was conducted in the subsequent semester (3 months after) and that was when the instructors did not teach these groups of learners anymore, the instructors, therefore, cannot offer extra credits to motivate these learners any further. This may partly explain why more attrition occurred during the delayed posttest.

thus, made it more difficult for some learners to complete the batteries of post and delayed posttests, either with the group or individually within the designated time frame. The researcher, however, still allowed these learners to complete the tests when they were available (since some of them still wanted some extra credits from the instructors or wanted to finish the whole experimental processes). Notwithstanding, their data were not included in the analysis since variations in the length of time intervals between each testing session may become another confounding variable affecting the overall results of the study. In sum, partially, the attrition occurred due to learners' unavailability to meet outside of class time to attend all the required testing sessions.

Of these 75 learners in the final pool, 17 were male and 58 were female and their ages ranged from 19 to 25 ($M = 20.37$, $SD = .11$). Their native language (L1) was Thai. They were low intermediate second (n = 32) and third year (n = 43) English major learners studying at a large public university in Thailand. These learners had been receiving formal English instruction for at least 10 years prior to the study ($M = 15.03$, $SD = .24$). Six learners had visited other English-speaking countries before. The results of one way between group ANOVAs revealed that there were no significant differences among the three participating groups in terms of age ($F (2, 72) = .273$, $p = .762$) and length of their previous formal English instruction ($F (2, 60) = .925$, $p = .402$).

These learners were from two intact English classes the English department at the university allowed the researcher to collect the data from. Since these two classes were seminar classes focusing on the discussion and oral presentation of Asian and Western cultures, English grammar was, therefore, not the emphasis of the classes, minimizing the chance the learners become too focused on English grammar, affecting the overall results of the study[28]. Apart from that, the main instructors of these two classes were advised to avoid reviewing grammatical structures particularly plural uses in class. It should also be noted that these learners already passed requisite English grammar courses since their freshman year and their second to fourth year curriculums were content-based rather than language (i.e., grammar) focused, thus, they were familiar with the meaning-based

---

[28] It was necessary that the learners did not become too focused on grammar or forms, because these learners might attempt to review their grammar during the experiment, and this may, to some extent, bias the data collected and influence the results of the study.

learning approach adopted in this study. All learners participating in this experiment were also familiar with argumentative essay writing; thus, there was no need to introduce essay writing basics to them at the onset of the study, even though most had never practiced timed writing before (they usually wrote their assigned essays outside of class time).

Low-intermediate learners are the target population of this study because the recent WCF meta-analysis (see Kang & Han, 2015) has demonstrated that intermediate and advanced learners are the two groups of learners that benefit most from WCF, whereas WCF seems to be less beneficial for beginners. Based on this finding, providing that the study aimed to investigate the effect of WCF on the development of learners' explicit and implicit knowledge, intermediate and advanced learners seemed to be a more proper population group for the study (given that they have a high probability to benefit more from feedback than beginners). However, intermediate learners were chosen to be the target population of this study instead of advanced learners. Even though the two groups of learners may equally benefit from feedback as pointed out by the meta-analysis, it is posited that most Thai intermediate learners may not fully develop their English plural knowledge (i.e., the target linguistic structure of this study) just yet, whereas it is more likely that many advanced learners may already successfully master the structure; thus, advance learners did not seem to fit the design of the study that aimed to evaluate the effect of WCF specifically on learners' English plural knowledge.

In addition, the study targeted EFL instead of ESL learners because in the EFL setting, learners have a tendency to be less affected by "the-out-of-experimental threat" (Lyster & Ranta, 2013, p. 179), unlike in the ESL setting, where learners are unavoidably exposed to potentially great amount of target linguistic input outside of experimental time and this may inevitably impact the reliability of the yielded results.

3.1.2.3. Target linguistic structure – Plural nouns

The target linguistic structure of the current study was English plurals, the ones that end with -s, -es morphemes and those of irregular forms (e.g., men, women, children). The selection of English plurals as the target feature was motivated by the fact that plural errors are ones of the most frequent grammatical errors most Thai learners made when writing, even though they have been receiving formal English instruction on plural uses for years (since primary education[29]). Consequently, it is worthwhile to examine as to whether WCF, the form-focused tool employed in this study, could efficaciously address this challenge.

The findings from a good number of error analysis (EA) studies within the Thai context examining errors most Thai learners frequently made when writing (e.g., Amnuai, 2020; Bennui, 2008; Khumpee & Yodkhamlue, 2017; Pongsiriwet, 2001; Thep-Ackrapong, 2005; Watcharapunyawong & Usaha, 2013) have demonstrated that plural errors are ones of the most frequent grammatical errors found in most Thai learners' L2 writing. According to Watcharapunyawong and Usaha's (2013) study, plural errors were ranked as the seventh most frequent grammatical errors Thai learners made when writing narrative essays, whereas they were ranked fourth for the descriptive writing genre. Plural errors, however, were the most frequent errors Thai learners made when writing argumentative essays (16.95%), followed by word choice (13.23%), articles (13.16%), subject-verb agreement (12.09%), and sentence structures (9.37%), respectively. Similarly, in Amnuai's (2020) study, plural errors were also ranked as the fourth most frequent errors made by Thai undergraduates. The most frequently made errors in this study were word choice, followed by preposition, sentence structures, plural nouns and quotation marks,[30] respectively. In consistence with Watcharapunyawong and Usaha's (2013) findings, the results of the current study's preliminary data analysis of the 36 pieces of written work from 19 Thai EFL undergraduate learners also evidently indicated that plural errors were the most frequent errors low-intermediate Thai learners made when

---

[29] Most learners received instruction on English plurals as earlier as third or fourth grade.
[30] In this study, mechanical errors (e.g., comma, capitalization) were also included in the analysis.

94

writing argumentative essays. According to the preliminary results, plural errors were the most frequently made errors (19.14%), followed by syntactic structures (11.68%), articles (10.33%), preposition (10.24%), word choice (9.43%), conjunction (7%), subject verb agreement (5.30%), adjectives (4.76%), pronouns (4.13%), tenses (3.50%), and adverbs (3.23%), respectively.

In part, it is posited that because of learners' L1 interference/ transfer, most Thai learners have trouble acquiring this linguistic structure despite several years of instruction. This is because unlike English, in Thai, only numbers are used to indicate plurality (i.e., there are no plural inflectional morphemes in Thai) as exemplified in examples 1- 6 below:

(1) In English: I need to drink 1 _glass_ of water before going to bed.
(2) In Thai: I need to drink 1 _glass_ of water before going to bed.
(3) In English: I need to drink 2 _glasses_ of water before going to bed.
(4) In Thai: I need to drink 2 _glass*_ of water before going to bed.
(5) In English: I ate 5 _oranges_ yesterday.
(6) In Thai: I ate 5 _orange*_ yesterday.

The findings from a large body of research on L1 transfer and morpheme studies have evidently shown that learners' L1 influences their L2 morpheme acquisition, especially when there is a lack of corresponding features in learners' L1 (e.g., Cintrón-Valentín & N. Ellis, 2015; N. Ellis, 2005, 2012, 2016a, 2016b; N. Ellis & Sagarra, 2010, 2011; Ionin & Montrul, 2010; Luk & Shirai, 2009; Murakami & Alexopoulou, 2016; Phoocharoensil et al., 2016; Slabakova, 2014; Watcharapunyawong & Usaha, 2013; Wulff & N. Ellis, 2018)). Luk and Shirai's (2009) review study, for example, found that Japanese, Korean and Chinese learners acquired plural forms much later than predicted by the universal natural order of acquisition (see Dulay & Burt, 1974; Ortega, 2009, Meisel, 2011, for a review of the natural order of morpheme acquisition) due to the absence of equivalent L1 features in their native languages suggesting strong L1 effects on their L2 morpheme acquisition. Similarly, Murakami and Alexopoulou's (2016) corpus study also found clear effect of L1 interference on the acquisition of some L2

morphemes. The study, thus, concluded that learners' L1 was a key predictor of learners' L2 grammatical accuracy. More specifically, the findings also indicated that the absence of equivalent morphemes in learners' L1 "nearly always" led to a low accuracy rate in an L2 (p. 394), while "the morphemes with equivalent forms in the L1 mark higher accuracy [in an L2]" (p. 396). Based on previous findings of these morpheme studies, it could be perceived that the absence of equivalent plural morphemes in Thai language may also negatively influence or delay Thai learners' acquisition of English plurals to some extent.

Many SLA scholars postulate that learners' L1 affects their L2 acquisition by interfering or biasing learners' learned attention[31] (i.e., selective attention) making learners look for cues (i.e., grammatical features) in an L2 that are similar to their L1 first, and if there is no cue in an L2 that is similar to their L1, learners may overlook or even have difficulty noticing the L2 cue leading to unsuccessful mastery of certain L2 forms (see Cintrón-Valentín & N. Ellis, 2015; N. Ellis, 2005, 2012, 2016a, 2016b; N. Ellis & Sagarra, 2010, 2011; N. Ellis et al., 2014, for the discussion of this topic). In other words, if there is a corresponding grammatical category in an L1, it is easier for learners to notice this cue in an L2. On the other hand, learners may completely overlook or may not notice certain grammatical categories in an L2 if there are no corresponding cues existing in their L1. However, many SLA scholars (R. Ellis, Loewen & Erlam, 2006; R. Ellis, 2008b, 2012a; Gass & Mackey, 2012, 2015; Long, 1998, 2007, 2014; Norris & Ortega, 2000, 2001; Schmidt, 2001, Spada, 2011; Swain, 2005) believe that form-focused instruction (FFI) can help direct learners' learned attention to problematic linguistic aspects that may be disregarded by increasing learners' chance of noticing the problematic forms. FFI is posited to help change learners' learned attention bias, making learners attend to input that they may overlook or have difficulty noticing before due to their L1 interference (Cintrón-Valentín & N. Ellis, 2015). Consequently, many researchers have endorsed FFI as useful intervention that is facilitating to L2 development. Corrective feedback, the main subject of this study, is also considered as one type of FFI tools and has also been

---

[31] Prior knowledge (in this case learners' L1) biases learners' learned attention by shifting their attention to the input that they are familiar with or direct them to correspond to the input that shares similar features to their prior knowledge first before paying attention to input that they share no background knowledge with (Cintrón-Valentín & N. Ellis, 2015; Wills, 2005).

attested to be useful in helping learners develop their L2 knowledge (see Abbuhl et al., 2018; Biber et al., 2011; Bitchener & Storch, 2016; Brown, 2016; Goo & Mackey, 2013; Kang & Han, 2015; Lee, 2019; Li, 2010; Li & Vuono, 2019; Mao & Lee, 2020; Russell & Spada, 2006; Van Beuningen, 2010 for a full review). Taken together, it is then worthwhile to examine if WCF, a form-focused tool employed in this study, could efficaciously help Thai learners acquire this problematic linguistic structure most Thai learners have difficulty acquiring due to their L1 interference, expanding the existing literature. In addition, to the best of my knowledge, given that none of the previous WCF studies has investigated the effectiveness of WCF exclusively on plural usage[32]before, this lack of empirical evidence emphasizes the need to examine the effect of WCF on this particular feature in order to address this gap in WCF literature. Based on these considerations, English plurals were selected to be the target linguistic structure of this study.

In this study, the direct focused WCF group only received corrections on plural errors, whereas the direct unfocused WCF group received corrections on 7 types of grammatical errors: plural nouns, conjunction, subject-verb agreement, adjectives, pronouns, tenses and adverbs. The control group did not receive any grammar correction, but content feedback. Details regarding each group's feedback are summarized in Table 3.1. below.

It should also be noted here that even though sentence structure, article, preposition and word choice errors are also frequently found in Thai learners' L2 writing, in this study, these errors were not given feedback upon because of the following reasons. First of all, sentence structure error is a type of syntactic errors and is beyond the scope of this study which only concentrated on morpheme types of errors. Late acquired grammatical forms such as articles were also excluded, as it is evident in the literature that it is developed much later (Huebner, 1983; Parrish, 1987; Shintani & R. Ellis, 2013).

---

[32] Indeed, there are studies that investigated the effect of WCF on multiple grammar structures including plural nouns but none of these previous studies has examined the effect of WCF exclusively on learners' plural usage. Theoretically, it is vital to explore the effect of WCF on one rather than multiple structures in a single study in order to truly distinguish the effectiveness of WCF on each linguistic structure, since it is probable that WCF may not be effective for all linguistic structures and the degree of effectiveness of WCF may greatly vary depending on the type of linguistic structure it is provided upon.

In addition, a large number of WCF studies have already investigated the effectiveness of WCF on article usage. In the same vein, preposition and word choice errors were also excluded since they are idiosyncratic errors and it might take longer time for these features to be remedied, unlike other rule-governed grammatical features that are potentially remedied within the limited time and resources of the current study.

Table. 3.1 Feedback Guideline

Experimental Groups	Target Structure (s)
1. Direct *Focused* WCF Group	**1 type of error** (only received corrections on plural errors)
2. Direct *Unfocused* WCF Group	**7 types of errors** (received corrections on plural nouns, conjunction, subject-verb agreement, adjectives, pronouns, tenses and adverbs)
3. Control Group	only received content feedback

3.1.2.4. Treatment

3.1.2.4.1. Treatment tasks

During the 6-week treatment sessions, learners from all experimental groups wrote six argumentative essays. Argumentative writing was selected for this current study for several reasons. First, this writing genre was selected because it seemed to well suite the target population of this study as when comparing to other writing genres (e.g., narrative, descriptive genres), argumentative writing is a more common writing practice for university students. In addition, practicing writing argumentative essays is posited to enhance the development of learners' academic literacy (Stavans, Seroussi, & Ehrlich, 2019). Thus, by practicing writing it during the experiment, learners had more opportunity to master their argumentative writing skill necessary for their higher education achievement. Second, since argumentative writing is part of many standardized tests, allowing learners to practice writing argumentative texts benefited learners in some way. Third, argumentative texts require learners to reason, generate and organize ideas with evidence to justify their arguments and to write this kind of complex tasks[33], learners may have to rely heavily on their working memory operations (Kim & Schatschneider, 2017; Stavans, Seroussi, & Ehrlich, 2019; Révesz, 2011; Robinson, 2005). Consequently, this genre was suitable for the current study, given that the study aimed to investigate the mediating effect of WMC on learners' L2 learning outcomes.

Regarding the argumentative topic chosen for the treatment task, the topic was selected if it was pertinent to (1) social controversial issues occurred in Thailand (the topic was widely discussed in the society) at the time of data collection, and (2) student life. It was anticipated that with familiar topics that involved their own student life or the

---

[33] It is postulated that tasks that demand reasoning are more cognitively demanding than tasks that do not require reasoning (Révesz, 2011; Robinson, 2005), because the tasks may heavily tax on learners' cognitive abilities including WMC when learners need to think of the reasons to support their arguments (Kim & Schatschneider, 2017; Stavans et al., 2019; Robinson, 2005). Therefore, it is perceivable that argumentative writing tasks which require learners to reason and support their claims in a logical manner could be a more cognitively demanding task than other types of writing tasks, given the narrative or descriptive writing tasks for example. Hence, the argumentative writing task may better reflect learners' use of WMC when writing than other types of written tasks.

society they lived in, learners would have adequate background knowledge to critically analyze the assigned topics and were able to support their opinions with concrete evidence. On the other hand, learners may not have adequate knowledge to critically discuss or thoroughly express their opinions if the topics are unfamiliar to them, as writing is an activity that can be restricted by the content knowledge of the writing topics (MacArthur, Graham, & Fitzgerald, 2006). In addition, since the topics involved their lives, learners may become more engaged with the writing task; thus, producing better quality written work.

In each treatment session, learners were required to write a 300-word argumentative essay responding to each week's assigned topic within the 30-minute time limit (see Appendix B for all assigned essay topics and instruction for the treatment). At the beginning of each treatment session, the assigned essay topic was displayed on the classroom screen (through the use of Microsoft PowerPoint computer program) so that learners could see the topic clearly. The instruction sheet was then distributed to all learners individually. To ensure that all learners comprehended the topic thoroughly, the researcher did a brief comprehension check asking some learners to explain what they were expected to write about and their opinions about the topic. After 5 minutes, learners received a piece of paper to write the essay on. Learners had 30 minutes to complete the 300-word essay. After 30 minutes, whether they finished the essay or not, their work was collected immediately (learners were asked to put down their pen/pencil immediately).

It should be noted that implicit instruction (i.e., meaning-focused instruction) was adopted during treatment to prevent learners to excessively focus on their grammar which may add another intervening variable to the study's outcomes.[34] Table 3.2. below summarizes characteristics of implicit instruction guideline adopted for this study.

---

[34] Some learners may review their grammar before the test section if they knew the real goal of the study. They may also become more aware of and monitor their grammar when performing the tests and this may especially affect the implicit knowledge test results.

Table. 3.2 Implicit instruction guideline adopted from Housen and Pierrard, (2006, p. 10) (cf. Norris & Ortega 2000; DeKeyser 1995; R. Ellis 2001, 2002).

Implicit FFI

- Attracts attention to target forms

- Is delivered spontaneously (e.g., in an otherwise communication-oriented activity)

- Is unobtrusive (minimal interruption of communication of meaning)

- Presents target forms in context

- Makes no use of metalanguage

- Encourages free use of the target form

In addition, during treatment, learners were also allowed to use electronic dictionaries when writing but they were not allowed to discuss their writing with their peers. Further, learners were informed before writing that those who wrote the three best content essays of each week would be awarded with a gift voucher (each received 100 THB = 3 GBP) from the researcher. The reason for awarding the learners was two folds: 1) to motivate learners to work harder on their writing (giving them some extrinsic motivators), and 2) to make them become more focused on meaning, rather than forms (essays were graded based on content).

3.1.2.4.2. Feedback treatment

Learners' argumentative writing tasks (i.e., treatment tasks) were given feedback upon every subsequent week. After reviewing their feedback for 5 minutes, learners started writing new written texts. Treatment procedure is illustrated in Figure 3.2. below.

With regard to feedback treatment, learners in the two experimental groups, the direct *focused* WCF and the direct *unfocused* WCF, received WCF on their targeted grammatical errors (see Table 3.1. for feedback guideline), while learners in the control group only received content feedback.

In the direct *focused* WCF group, learners' plural errors were underlined, and plural correction was supplied above plural errors. In the same manner, in the direct *unfocused* WCF group, all targeted grammatical errors were underlined, and correction was supplied above each targeted error.

For the control group, content feedback was provided when learners' arguments were ambiguous or incomprehensible, or when their writing organization was incoherent or lacking in some ways, such as when there was no introduction, no thesis statement, no conclusion, or when they used supporting details or examples as the paragraph's main idea etc.

Precisely, for content feedback, words or sentences that were incomprehensible or ambiguous, were underlined, and the question mark symbols "???" were placed above them to indicate ambiguity. Learners also received comments either on the margin or at the end of their written texts addressing the lack of organizational elements of a good essay such as a lack of topic sentence or proper conclusion, for example.

Figure 3.2 Treatment procedure

Week 1
Practicing Writing 1

Week 2
Reviewing Feedback
Practicing Writing 2

Week 3
Reviewing Feedback
Practicing Writing 3

Week 4
Reviewing Feedback
Practicing Writing 4

Week 5
Reviewing Feedback
Practicing Writing 5

Week 6
Reviewing Feedback
Practicing Writing 6

Week 7
Reviewing Feedback
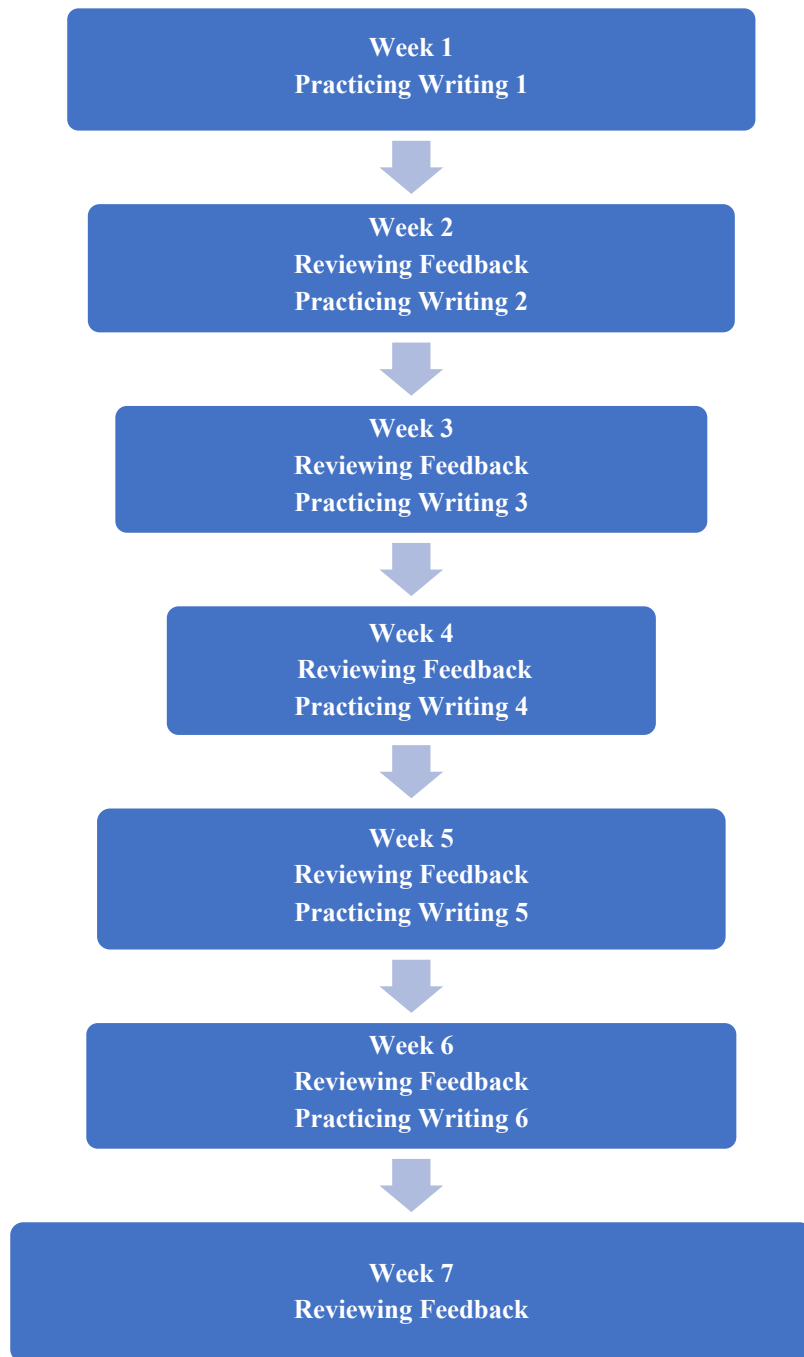
3.1.2.5. Assessment tasks and scoring
3.1.2.5.1. Questionnaires

In this study, self-report background and exit questionnaires were employed to collect learners' personal data and their opinions about the project. The information obtained through the questionnaires was used to assess whether the target population was homogeneous or heterogeneous and if there were other variables (besides the controlled variables) that may potentially influence the results of the study.

Precisely, a background questionnaire was used to collect learners' bio data such as gender, age, field of study, number of years of instruction taught in English, abroad experience, etc. (see 3.2.2.2. participants section for detailed information about learners' bio data). Meanwhile, an exit questionnaire was designed to obtain information regarding learners' perceptions about their writing progress, their writing concerns, what they think is the goal of the current study, etc. (see chapter 5 qualitative results, section 5.2.2. for the summary of exit questionnaire data). Samples of background and exit questionnaires can be found in Appendix C.

3.1.2.5.2. Timed argumentative writing tasks

The timed argumentative writing task was designed to measure learners' implicit knowledge. In line with R. Ellis et al.'s (2009) criteria for tests of implicit knowledge, the primary focus of the timed argumentative writing task was on meaning (similar to when learners practiced their writing during treatment) as learners were instructed to pay their most attention to develop strong argumentative essays on the given topic (i.e., the task did not invite learners to use their metalinguistic knowledge). The task was also time pressured.

Nonetheless, despite all the above (i.e., meaning-focused instruction and time pressure), the researcher acknowledged that during writing, whether timed or untimed, learners still had opportunity to consciously monitor their grammatical usage, relying on their explicit knowledge. However, while the timed writing may not appear to be the best measure of implicit knowledge since it may or may not only measure learners' implicit

but explicit knowledge, this test had high face validity corresponding to the focus of this study which aimed to measure learners' grammatical progress on their writing. Besides, while it may compromise the credibility of the study's results if the timed writing was used as a sole measurement of implicit knowledge, the use of the timed writing together with an oral elicited imitation test (OEIT), a widely accepted measure of implicit knowledge (see Bowles, 2011; Culbertson, Andersen & Christensen, 2020; R. Ellis, 2005; R. Ellis et al., 2009; Erlam, 2006; Kim & Nam, 2017; Sarandi, 2015; Spada, Shiu, Tomita, 2015; Wu & Ortega, 2013; Yan, Maeda, Lv & Ginther, 2016; Zhang, 2015) (has low face validity for writing research but high construct validity, see Shintani & R. Ellis, 2013) in this study, can help strengthen the validity of the yielded findings, considering that data from different measures (also from different modalities) were gathered and compared before any conclusions could be drawn or assumed as truth. By using multiple measurements so that the data from different sources can be triangulated increases the overall credibility and reliability of the findings, particularly in this case where the validity of the timed writing is still questionable.

Apart from that, as suggested by Shintani and R. Ellis (2013), the validity of the timed writing as measure of implicit knowledge can be affirmed in two important ways: 1) by incorporating "a delayed writing task in the design of a study" (p. 291), and (2) by "obtaining evidence of how learners oriented to the task" (p.291). Shintani and R. Ellis (2013) explained that implicit knowledge, once acquired, will not be forgotten easily, unlike explicit knowledge. Thus, learners' improvement in their delayed writing (i.e., indicating that knowledge is sustained over time) can be considered as evidence of implicit knowledge development. Moreover, evidence that clearly demonstrates learners' orientation when performing the writing task or elicits whether learners pay conscious attention to the target form or meaning (Shintani & R. Ellis, 2013, p.291), could help affirm if the writing test biases the use of implicit knowledge as anticipated or not. Hence, in this study, the delayed writing task, exit questionnaire as well as qualitative interview were incorporated into the design, in part, to strengthen the validity of the timed writing task as measure of implicit knowledge.

During the timed argumentative writing test, similar to the writing treatment task, learners only had 30 minutes to finish a 300-word essay. The test was administered

during the pretest, posttest and delayed posttest. The writing topics varied on each test occasion. At the pretest, learners were asked to write a response to the topic, "University degree: is it necessary for success?", while at the post test, learners were asked if they should blame their teachers if they have low scores on the test (see Appendix D for the test instruction and topics).

With regard to the scoring of this task, the adapted version of Pica's (1991) obligatory occasion analysis was employed. In each analysis, all obligatory contexts for plural uses[35] and the overuse of plural forms in the non-obligatory contexts were first identified. The number of obligatory contexts and number of plural suppliance in non-obligatory contexts were then added into a total sum. Next, each required context was then examined if correct plural forms were supplied. The incorrect suppliance of plural forms was not counted towards the analysis. The incorrect suppliance included the omission of plural nouns in obligatory contexts, the use of singular in place of plural nouns, and the incorrect supply or oversupply of plural nouns (e.g., mans, womanes, childes). To control differences in text length, counts were first cut off at 300 words (in case some learners wrote more than 300 words). If learners' work did not reach 300-word requirement, their actual text length was used in the calculation. An additional equation (i.e., number of total words/300) was added to the original formula to approximate learners' actual accuracy if they could reach the 300-word requirement.

Learners' accuracy score/percentage was calculated based on the following adapted formula:

$$\frac{\text{Number of correct uses of plural nouns}}{\text{Number of obligatory contexts} + \text{Number of suppliance in non-obligatory contexts}} \times \frac{\text{Number of total words}}{300} \times 100$$

---

[35] Obligatory occasion was defined as the obligatory context where plural forms were required.

Example 1: A 300-word essay $= \dfrac{10}{20} \text{ X } \dfrac{300}{300} \text{ X } 100 = 50/100$

Example 2: A 180-word essay $= \dfrac{8}{18} \text{ X } \dfrac{180}{300} \text{ X } 100 = 26/100$

As elucidated, the additional equation (i.e., number of total words/300) was added to the original formula so that it was fair for all learners whose work had different length. It would be unjust if those who wrote shorter than 300 words and made fewer errors due to fewer obligatory contexts would get higher accuracy score/percentage when calculated with the original formula, than those who attempted to write as fast as they could within the 30 minute limit and, thus, were prone to make more errors. In other words, to be fair for all learners whose work had different text length, an additional equation was included in the original formula.

For reliability purposes, 19 of 75 essays from each test occasion (i.e., 57 of 225 essays or 25% of the overall essays collected from the pretest, posttest, and delayed posttest) were rescored by a trained and experienced native English speaker teacher. The two sets of scores of each test occasion (graded by the researcher and the native English speaker teacher) were then analyzed for inter-rater agreement using Intraclass Correlation Coefficient (ICC). ICC estimates and their 95% confident intervals were calculated based on a mean rating ($k = 2$), absolute-agreement, two-way random effects model. The obtained ICC value for the pretest scores was .994 with its 95% confidence interval ranging from .985 to .998, indicating excellent level of inter-rater reliability[36]. For the posttest scores, the obtained ICC value was .997 and its 95% confidence interval ranging from .991 to .999. For the delayed posttest scores, the obtained ICC value for the delayed posttest scores was .988 and its 95% confidence interval ranging from .924 to .997. The ICC estimates of the post and delayed posttest scores indicated excellent levels of

---

[36] The ICC value that is less than .50 indicates poor inter-rater reliability. The ICC values that range between .50 and .75, and between .75 and .90 indicate moderate and good levels of inter-rater reliability, respectively. ICC estimate that is greater than .90 is indicative of excellent level of inter-rater reliability (Koo & Li, 2016).

consistency between the two raters (ICC > .90). The summary of ICC estimates can be found in Appendix E.

3.1.2.5.3. Oral elicited imitation test (OEIT)

The OEIT test was designed to measure learners' implicit knowledge following R. Ellis et al.'s (2009) criteria for tests of implicit knowledge: the primary focus of the test was on meaning, learners were not invited to use their metalinguistic knowledge, and the task was timed pressure. The OEIT test was selected because a number of studies and meta-analyses provide empirical support of its reliability and validity as a test of implicit knowledge (see Bowles, 2011; Culbertson, Andersen & Christensen, 2020; R. Ellis, 2005; R. Ellis et al., 2009; Erlam, 2006; Kim & Nam, 2017; Sarandi, 2015; Spada, Shiu, & Tomita, 2015; Wu & Ortega, 2013; Yan, Maeda, Lv & Ginther, 2016; Zhang, 2015).

During the individual OEIT test, learners listened to 22 audio-recorded sentences in English (comprised both grammatical and ungrammatical sentences targeting plural noun usage and other grammatical features). After hearing each sentence once, learners decided whether the content of the sentence was a real fact or not. If the sentence constituted a real fact, learners said "True" in English and circled "True" on the OEIT answer sheet. If it was not true, they said "False" and circled "False" on the answer sheet. For example, after hearing the sentence, "America and England are neighboring countries", learners should say "False" and circle "False" on their answer sheets, as in reality America and England are not neighboring countries. This part of the test was designed to draw learners' attention to the sentence's meaning.

After judging the content of each sentence, learners repeated the sentence they heard in correct English within 5 seconds. Every time learners repeated the sentence back in correct English with the target form was correctly supplied and the original meaning of the sentence was reserved, they received one point for that sentence. Learners received zero if 1) they failed to imitate or reproduce the sentence back in correct English, 2) the target form was incorrectly supplied or omitted from the sentence, 3) the meaning of the sentence was completely deviant from the original, and/or 4) they failed to repeat the sentence back within 5 seconds. Learners' responses were audio recorded for scoring

purposes.

Six (3 grammatical and 3 ungrammatical sentences) out of 22 sentences targeted the use of plural nouns. Thus, six was the total score of this test (excluding other distractors' scores).

The internal consistency reliability of the two parallel versions of the OEIT tests (i.e., form A and form B) was assessed using Cronbach's alpha (as part of piloting) before administering at the pretest-posttest-delayed posttest. The internal consistency reliability coefficients of the two versions of the tests were at an acceptable level[37] ($\alpha = .78$ and $\alpha = .74$, respectively).

3.1.2.5.4. Untimed grammaticality judgment test (UGJT)

This test was designed to measure learners' explicit knowledge. In consistence with R. Ellis et al.'s (2009) criteria, the test required learners to (1) pay attention to forms, (2) deploy their metalinguistic knowledge to justify their answers, and (3) perform the test without any time pressure. As the test drew learners' attention to grammatical forms and there was no time limit for the test, learners had adequate time to deploy their metalinguistic rules when performing the test (R. Ellis et al., 2009). In other words, this test completely allowed learners to draw upon their explicit knowledge repertoire (R. Ellis, 2005; R. Ellis et al., 2009). The test is also widely accepted as reliable and valid measure of learners' explicit knowledge (Bowles, 2011; R. Ellis, 2005; R. Ellis et al., 2009; Godfroid et al., 2015; Gutiérrez, 2013; Zhang, 2015).

The UGJT test used in this study adopted the same test format as the explicit knowledge test employed in R. Ellis et al.'s (2009) study. It consisted of 40 decontextualized sentences, which were evenly divided into 20 grammatical and 20 ungrammatical sentences. Out of 40 sentences, 6 sentences targeted the use of plural nouns (3 correct and 3 incorrect usage), while 34 sentences concerned other grammatical features (i.e., distractors). Learners were required to identify each sentence as either grammatically correct or incorrect. If they believed that the sentence was grammatically

---

[37] The value of ($\alpha > .70$) is universally considered an acceptable level of internal consistency (Taber, 2018).

correct, they checked the "Correct" box and continued to the next sentence. If they believed that the sentence was grammatically incorrect, they checked the "Incorrect" box, underlined the incorrect part and substituted the incorrect part with the correct form.

Example: Mai completed her assignment and print it out.

☐ Correct   ☐ Incorrect

The test was scored on a discrete item basis. One point was given for each correct answer. A half point was given if learners can only judge the grammaticality or ungrammaticality of the sentence but cannot identify the ungrammatical part nor substitute the correct grammatical form to the ungrammatical sentence. For instance, learners received a half point if they could correctly judge that the example sentence was incorrect yet could not identify the ungrammatical part or substitute the error with the grammatically correct form (e.g., underlining the word "print" and write down "printed" above the underline). The total score of the UGJT test targeting plurals was six.

There were two parallel versions of the UGJT tests (i.e., form A and form B). The internal consistency of the tests was assessed using Cronbach's alpha (as part of piloting). The internal consistency reliability coefficients of the two versions of the tests were at an acceptable ($\alpha = .73$) and good ($\alpha = .88$) level, respectively. The two versions of the test can be found in Appendix D.

3.1 2.5.5. Metalinguistic knowledge test (MKT)

The MKT was designed to measure learners' explicit knowledge. It was an adapted version of the MKT originally devised by Alderson, Clapham, and Steel (1997) and R. Ellis et al. (2009). The test is also widely accepted as a well-established measure of learners' explicit knowledge (R. Ellis, 2005; R. Ellis et al., 2009; Gutiérrez, 2013; Vafaee, Suzuki, & Kachisnke, 2017; Zhang, 2015).

The test consisted of 10 ungrammatical sentences in which one out of these ten sentences targeted plural noun uses. All sentences were ungrammatical and one grammatical error in each sentence was underlined. Learners were required to 1) substitute the ungrammatical part with the correct grammatical form, and 2) explain the language rule that was being violated in that sentence.

Example: He saw <u>a elephant</u>.

In this example, learners should substitute the article "a" with "an" and explain why the use of article "a" is incorrect in this context.

Learners received one point if they successfully substituted the error with the correct linguistic form and explained the violation of the relevant language rule. A half point was given if learners can only explain the rule but cannot substitute the error with the correct grammatical form or vice versa. Learners received zero point if they could neither correct the error and substitute it with the correct form nor explicate the relevant metalinguistic rule. It should be noted that learners were allowed to explain the metalinguistic rule of the underlined error either in English, Thai or the mix of the two, so that they were able to fully and clearly explain the linguistic rule they knew and were not obstructed to explain the rule because of their limited linguistic knowledge.

There were 2 parallel versions of the MKT: form A and B. The internal consistency of the tests was assessed using Cronbach's alpha (as part of piloting). The internal consistency of the two versions of the tests was at a good and acceptable level

(α = .87 and α = .65, respectively)[38]. The two versions of the test can be found in Appendix D.

For inter-rater reliability purposes, 19 of 75 MKT from each test occasion (i.e., 57 of 225 MKT or 25% of the overall MKT collected from the pretest, posttest, and delayed posttest) were also rescored by a trained and experienced native English speaker teacher. The two sets of scores of each test occasion (graded by the researcher and the native English speaker teacher) were then analyzed for inter-rater agreement using Intraclass Correlation Coefficient (ICC). ICC estimates and their 95% confident intervals were calculated based on a mean rating (k = 2), absolute-agreement, two-way random effects model. The obtained ICC value for the pretest scores was .974 with its 95% confidence interval ranging from .933 to .990, indicating excellent level of inter-rater reliability. For the posttest scores, the obtained ICC value was .982 and its 95% confidence interval ranging from .954 to .993. For the delayed posttest scores, the obtained ICC value was .896 and its 95% confidence interval ranging from .532 to .967. The ICC estimates of the post and delayed posttest scores also indicated excellent levels of consistency between the two raters (ICC > .90). The summary of MKT ICC values can be found in Appendix E.

3.1.2.5.6. Backward digit span task

The backward digit span Task (see Wechsler, 1997) was one of two complex working memory measures (measuring processing and short-term storing components of working memory) used in this study. It was employed in this study because a large body of research has suggested that complex working memory is a stronger predictor of L2 learning outcomes than simple working memory that only measures storage component of working memory (Daneman & Merickle, 1996; Linck, et al., 2014; Tagarelli, Mota & Rebuschat, 2015). In addition, backward digit span task was selected because it is a non-verbal task, that is the processing part of the task does not involve the processing of linguistic materials but the nonlinguistic ones such as numbers, math problems, etc. (Juffs

---

[38] Despite an α estimate of the MKT Form B may appear questionable, the value (α = .65) is considered acceptable for tests with a small number of items (Tavakol & Dennick, 2011). As Field (2018) points out, the value of α also depends on the number of test items.

& Harrington, 2011; Linck et al., 2014); thus, reducing the possibility learners' prior L2 knowledge affects the test performance (Juffs & Harrington, 2011; Tagarelli, Mota & Rebuschat, 2015). In this study, the backward digit span task was also carried out in learners' L1 (Thai language) since it is posited that the task administered in an L2 may not purely measure learners' WMC but also their L2 proficiency (Grey, Serafini, Cox & Sanz, 2015; Juffs & Harrington, 2011; Linck et al., 2014; Sagarra, 2012; Serafini & Sanz, 2016; Wen et al., 2015) and "in the context of predicting L2 outcomes, this confound would inflate the WM outcome correlation estimate" (Linck et al., 2014, p. 872).

The Thai-version of the backward digit span task employed in this study was adopted from McDonough and Trofimovich's (2016) study. The test was already validated by McDonough and Trofimovich (2016). The reliability of the test was ($\alpha = .69$) indicating an acceptable level of internal consistency. In this task, twenty-digit sequences of increasing length (from 2 up to10 digits), with two sequences of each length, were recorded by a male Thai native speaker. Each digit sequence was presented to learners at the rate of one digit per second through the Power Point computer program. Learners listened to digit sequences and repeated the sequences back in reverse order after the sound signal. For instance, after hearing 6-2-2, learners repeated the sequence back in reverse order saying 2-2-6. Learners only had 5 seconds to repeat the sequence back.

Learners received one point for each correctly repeated sequence responding within 5 seconds. They received zero if they failed to repeat the sequence back within the time limit. The time limit was imposed in this task to prevent learners to employ processing strategies, e.g., covertly rehearsing the numbers, that may enhance their working memory test performance. The task was concluded when learners incorrectly repeated two sequences of the same length or correctly repeated two sequences of the length of 10 digits. The total score of this task was 20. Learners' responses were audio recorded. Note that, learners were tested individually in order to prevent them to develop some learning strategies to better recall the numbers[39].

---

[39] If the test was administered in group, it would be difficult to control learners' response time since some learners might attempt to answer before the sound signal while others may attempt to repeat the sequence back as late as possible, adding another confounding variable affecting the result.

3.1.2.5.7. Operation span task (OSpan)

The OSpan (see Turner and Engle, 1989) was another complex working memory measure used in this study. The task is non-verbal, and this reduces the possibility that learners' prior L2 knowledge impacts the test performance. The OSpan was employed together with the backward digit span task, as the literature has suggested that the results yielded from a single measure are usually inconsistent and therefore using more than one task is conjectured to bring out more reliable results (Waters and Caplan, 2003).

The OSpan used in this study was adapted from Tagarelli, Mota and Rebuschat's (2015) English-version OSpan. The task comprised the solving of simple math problems (involving addition, subtraction, multiplication and division), and remembering words in memory. The math processing part completely followed Tagarelli et al.'s (2015) version but for the storage component (i.e., remembering words), English words were replaced by Thai words. Following Tagarelli et al.'s (2015) English-version, all selected Thai words has one syllable and are frequently used in Thai daily life. The task was conducted in learners' L1 (Thai language) because as mentioned earlier, a number of WM studies have suggested that WMC may be subjected to learners' proficiency if tested in an L2 (see Grey et al., 2015; Juffs & Harrington, 2011; Linck et al., 2014; Sagarra, 2012; Serafini & Sanz, 2016; Wen et al., 2015). In contrast, as L1 ability is less variable across individuals than L2 ability, it can, thus, be more certain that by using an L1 version, individuals' WMC scores will not be influenced by language-related ability (Tagarelli et al., 2015).

The Thai-version OSpan test developed for this study was assessed by Cronbach's alpha for internal consistency reliability and Pearson Correlation Coefficient for test-retest reliability (as part of piloting) before administering for the pretest, posttest and delayed posttest. The internal consistency reliability coefficient of the test was found to be at an acceptable level ($\alpha = .71$). The Person Correlation coefficient estimate was ($r = .73$, $p < 0.01$) demonstrating a good level of test-retest reliability.

In this OSpan task, learners first (1) saw a math problem (e.g. (8 x 2) - 6 = 10?) on the computer screen, (2) read the math problem aloud in Thai and stated whether the answer provided was correct or incorrect, (e.g., saying " eight timed two minus six equal

ten…correct"), and (3) put either the check mark (if the answer was correct) or the cross symbol (if the answer was incorrect) on the answer sheet. After each math problem, learners (4) saw a to-be-remembered-word in Thai and (5) read the word out loud before advancing to the next math equation. Learners repeated the same steps (1-5) until they finished the whole set. At the end of each set, learners saw three question marks (???), which cued them to write down all the to-be-remembered-words in correct order on the answer sheet before moving to the next set. Learners were asked to read all stimuli out loud in order to prevent them to rehearse the to-be-remembered words, following the instructions from Unsworth, Heitz, Schrock, and Engle (2005).

Overall, the task comprised 12 sets, with 2–5 words in one set (also see Tagarelli et al., 2015). Three sets of each set length, from 2 to 5 words, were presented making it 42 trials in total. Note that within this task, the math part was processed under time limit (5 seconds for each math equation and 1 second for each to be remembered word before the computer screen disappeared) while the recall part was untimed (however, under appropriate time frame, not more than 30 seconds). Time limit for the response was set because self-pacing can potentially attribute to variability in the task. Learners were tested individually.

With regard to the scoring of this test, learners received one point for each correct math problem and correct word recall following its correct order appearing in a set. This made a total of 42 points for the math part (processing component) and 42 points for the word recall part (storage component), giving a composite score of 84. To ensure that all learners completely engaged in the task, learners had to reach 80% accuracy on the math part so that their scores could be counted for further analysis, following Unsworth et al.'s (2005) scoring criterion.

3.1.2.5.8. Qualitative interviews

The qualitative interview was employed in this study in order to acquire more in-depth information regarding learners' writing and revising processes and what they viewed as beneficial for their writing development. The qualitative interview was chosen as an additional qualitative research method for this study because it could serve as an effective introspective research tool that helps further explore learners' cognitive or thinking processes.

In this study, the qualitative interview was conducted 3 times; immediately after the 3rd treatment session, after the 6th (last) treatment session, and after the posttest, following Mackey and Gass's (2016) suggestion that interviews should be carried out as soon as the main task is completed to prevent learners from reporting inaccurate data resulting from learners' memory decay.

The interview was also conducted in Thai, learners' native language, in order to prevent miscommunication that may arise due to learners' limited L2 knowledge. In addition, it is possible that learners' limited L2 knowledge may hinder them from fully expressing their thoughts on the interview questions, impacting the quality of the interview data.

Six learners from the three participating groups (i.e., two learners from each group) were randomly selected to take part in the interview. During an individual interview, each learner was given a chance to look at his/her latest written feedback and his/her latest in-class practice written essay while answering interview questions such as "What did you do after seeing your feedback?", "What did you first do when the teacher said you could start writing your essay?", "Did you have any chance to monitor your writing?, What did you look for when monitoring your writing?", etc. The first two interviews conducted during the treatment sessions took approximately 5-10 minutes for each learner and learners were mainly inquired about their revision processes and writing progress. The last interview conducted after the posttest took approximately 10-20 minutes for each learner. Learners were mainly inquired about their writing processes, what they concerned about when writing, whether they monitored their work when writing and what they thought was the main goal of this study, etc. All interviews were

audio-recorded, transcribed and coded for subsequent analysis (see chapter 5 for interview coding procedures and its findings).

3.1.2.6. Procedure

In the first meeting (first week of the semester), volunteering learners were informed and provided with information sheets regarding the experiment details (see Appendix A). After reading the information sheets, learners were given opportunity to ask questions about the experiment. Learners who decided to take part in the experiment were asked to sign the consent forms (see Appendix A) and completed background questionnaires. A week after, learners who consented to participate in the study completed some pretests in class (comprising the timed writing, UGJT and MKT sequentially[40]) and scheduled individual test sessions for the working memory and OEIT tasks outside of class time.

After the completion of all group and individual pretests, learners were randomly assigned to different participating groups (either the direct focused WCF, direct unfocused WCF or control group). Next, learners attended the 6-week treatment sessions.

In each treatment session, learners received feedback on their previous writing (see figure 3.2. for a more specific treatment procedure), checked their feedback silently for 5 minutes and then wrote a new argumentative essay within 30 minutes (new topic every week). While practicing writing the essay, learners were allowed to look at their previous writing and feedback if they wanted to. However, learners were not allowed to take their previous writing back home. They needed to return their previous writing back to the researcher at the end of the new writing practice session. This was to prevent learners from reviewing the feedback or exposing themselves to the target linguistic feature outside the treatment sessions, adding confounding variables to the study's findings.

One week after the treatment sessions, learners completed the immediate posttests (comprising the timed writing, UGJT and MKT) together with the exit questionnaire and

---

[40] Timed writing was administered first so that learners did not become too focused on their grammatical usage (i.e., focus on forms).

scheduled for individual OEIT and working memory tests. The delayed posttest was administered three months after the immediate posttest to examine if the effect of WCF was sustained. Table. 3.2. summarizes the research procedure of the current study.

In an attempt to eliminate the chance that other independent variables might affect the outcomes of this study, the researcher administered all the tests, controlled the treatment procedures and provided feedback all by herself in order to decrease the variabilities in test administration, quality and quantity of feedback that might come into play. In addition, in an attempt to control learners' exposure to the target linguistic feature outside the experimental time, all English instructors responsible for the learners' English courses were asked not to review or emphasize the use of plural nouns in their classes.

Table. 3.3 Research procedure

	Time (min)	Direct Focused (n = 27)	Direct Unfocused (n = 25)	Control (n = 29)
Week 1	30	Consent Forms and Background Questionnaires		
Week 2 (Group Test)	90	**Pre-tests** (Timed Writing, UGJT and MKT)		
Week 3-6 (Individual Test)	30	OEIT and WM tests		
Week 7-12 (In class) **Treatment**	5	*Received direct focused WCF (targeting plural errors only)*	*Received direct unfocused WCF (targeting a wide range of grammatical errors)*	*Received content WCF*
	30	***Practice timed writing***		
Week 13 (Group Test)	90	**Post-tests** (Timed Writing, UGJT, MKT, and Exit Questionnaires)		
Week 14-19 (Individual Test)	30	OEIT and WM tests		
Week 30 (Group Test)	90	**Delayed Post-tests** (Timed Writing, UGJT and MKT)		
Week 31-35 (Individual Test)	15	OEIT		

Chapter IV
QUANTITATIVE RESULTS

This chapter presents the quantitative part of the current study. It first makes reference to the research questions of the current study. Then, the methods for the quantitative data analyses are presented, followed by the results and discussion sections. In the last section of this chapter, some limitations of this part of the study are offered.

4.1. Research questions

As mentioned in previous chapters, the main goals of the current study are (1) to examine if there are effects of WCF on the development of learners' L2 explicit and implicit knowledge, and (2) to explore if there is a relationship between learner differences in WMC and the potential benefits of WCF on learners' L2 learning outcomes.

The followings are research questions 1 to 4, which were formed in accordance with the main goals of this study and guided the subsequent analyses.

Research Questions

RQ1: Do direct focused and unfocused written corrective feedback have any effect on the development of L2 learners' explicit knowledge of English plurals?

RQ2: Do direct focused and unfocused written corrective feedback have any effect on the development of L2 learners' implicit knowledge of English plurals?

RQ3: Is there any difference in the effects of direct focused and unfocused written corrective feedback on the development of L2 learners' explicit and implicit knowledge of English plurals?

RQ4: Does working memory capacity moderate the extent to which L2 learners benefit from written corrective feedback?

4.2. Methods of data analyses

In this study, quantitative data were analyzed using IBM Statistical Package for Social Sciences (SPSS) version 25. The initial step required identifying and excluding influential outliers in the quantitative data since these extreme cases can significantly affect or bias the test results (Field, 2018). The extreme outliers in this study evidenced scores with standardized residuals (converted to z scores) greater than -3 or +3 and were identified as extreme cases in more than one set of test scores. Based on these criteria, three learners were identified as outliers and were excluded from the subsequent analyses. The control group was reduced from 26 to 24 and the direct focused WCF group, from 26 to 25. Consequently, the final sample size for the analyses comprised 72 learners (control [n = 24], direct focused WCF [n =25], and direct unfocused WCF [n = 23]).

After removing those extreme cases, to ensure that there were no inherent differences between the three participating groups at the onset of the study, the distribution of each set of pretest scores was examined for skewness and kurtosis. A Shapiro-Wilk test was also run to determine if the distribution satisfied the assumption of normality for parametric tests (see Appendix F for normality test results). All groups' pretest scores were then submitted to a series of one-way, between-groups ANOVAs to determine if there were any significant differences among the participating groups at the onset of the study. The test results indicated no statistically significant differences in the pretest scores among the three participating groups (see Appendix G). Therefore, the groups were determined to have equivalent knowledge of English plurals prior to the study. After completing the group equivalence tests, in all the analyses that follow, descriptive statistics were computed and reported first, followed by inferential statistics. The threshold for significance was set at $p < .05$[41].

---

[41] The significance level of $p < .05$ is the widely accepted level of significance for the Null Hypothesis Significance Testing (NHST) (Field, 2018). It is the rule-based framework used to determine if the null hypothesis is true. Anyhow, it is worth noting that there are still many misconceptions about statistical significance (*p value)*, namely, a non-significant result (*p >.05)* indicating that the null hypothesis is true or vice versa (see "problems with NHST" in Field, 2018, and Plonsky & Oswald, 2014).

To address RQ 1-3, i.e., investigating the effects of direct "*focused*" and "*unfocused*" WCF on the development of learners' explicit and implicit knowledge of English plurals, and examining if there were any comparative effect of both WCF on the development of learners' knowledge, a series of mixed design ANOVAs were conducted. Effect sizes were also computed for the analyses in order to measure the magnitude of the effects of different feedback types. Partial eta-squared (η_p^2) was used as an effect size measure for the mixed ANOVAs. Meanwhile, *d* value was used as an effect size indicator for the post hoc pairwise follow-up tests. Following Cohen's (1988, 1992) effect size index, partial eta-squared (η_p^2) values of .01, .06 and .14 were considered small, medium and large effect sizes, respectively. Meanwhile, the *d* values of .40, .70 and 1.00 represented small, medium, and large effect sizes sequentially, following Plonsky and Oswald's (2014) effect size index.

To address RQ 4, i.e., investigating whether learner differences in WMC moderated the extent to which learners can benefit from WCF, a series of moderated multiple regressions using Hayes' *PROCESS* tool were conducted.

4.3. Quantitative results

The following sections report quantitative results responding to the four research questions of the study. All results sections begin with descriptive statistics followed by inferential statistical analyses.

To address RQ 1-3, a series of 3 (feedback types: Direct Focused vs. Direct Unfocused vs. Control) x 3 (Time: Pretest vs. Posttest vs. Delayed Posttest) mixed design ANOVAs were carried out to examine the effects of different feedback types on four measures of plural learning outcomes. The Greenhouse-Geisser and Huynh-Feldt adjustment were applied to correct for violations of sphericity assumption when appropriate.

4.3.1.1. Explicit knowledge test results

Descriptive Statistics. Descriptive statistics for the pre, post and delayed posttest scores of all the explicit knowledge tests were computed, summarized and presented in Table 4.1., Figure 4.1., Figure 4.2., and Figure 4.3. below.

As Table 4.1. and Figures 4.1.to 4.3. demonstrate, the general pattern detected was that all groups' posttest scores (Time 2) on the UGJT, MKT and combined explicit knowledge posttest scores showed an increase from the pretest (Time 1), though the scores decreased at the delayed posttest (Time 3). When comparing the delayed posttest scores (Time 3) to the pretest scores (Time 1), only the MKT delayed posttest scores (Time 3) showed improvement from the pretest (Time 1). All groups' UGJT and combined explicit knowledge delayed posttest scores (Time 3) were lower than the pretest scores (Time 1), with the exception that the unfocused group maintained their combined explicit knowledge scores from the pretest (Time 1) to the delayed posttest (Time 3).

Table 4.1

Descriptive statistics of all the explicit knowledge tests

Treatment		N	Mean	Std. Deviation	Skewness Statistic	Std. Error	Kurtosis Statistic	Std. Error
Control	PreUGJT	24	3.45	1.141	-.654	.472	.085	.918
	PostUGJT	24	3.58	.974	.668	.472	.340	.918
	DelayedUGJT	24	2.95	.954	.416	.472	.513	.918
	PreMKT	24	.66	.318	-.408	.472	-.539	.918
	PostMKT	24	.91	.190	-1.910	.472	1.792	.918
	DelayedMKT	24	.70	.358	-.839	.472	-.485	.918
	PreExplicit	24	4.12	1.312	-.376	.472	-.734	.918
	PostExplicit	24	4.50	1.021	.567	.472	.571	.918
	DelayedExplicit	24	3.64	1.108	.408	.472	.404	.918
Unfocused	PreUGJT	23	3.47	.665	-.928	.481	-.124	.935
	PostUGJT	23	3.86	.868	.725	.481	-.074	.935
	DelayedUGJT	23	3.43	.843	-.025	.481	-.394	.935
	PreMKT	23	.78	.294	-1.002	.481	.157	.935
	PostMKT	23	.95	.144	-3.140	.481	8.605	.935
	DelayedMKT	23	.82	.323	-1.728	.481	1.945	.935
	PreExplicit	23	4.26	.810	-.675	.481	-.818	.935
	PostExplicit	23	4.82	.924	.515	.481	-.237	.935
	DelayedExplicit	23	4.26	.864	.076	.481	-.844	.935
Focused	PreUGJT	25	3.28	.678	-.410	.464	-.680	.902
	PostUGJT	25	3.48	.871	-.344	.464	-.532	.902
	DelayedUGJT	25	2.72	1.173	.261	.464	-.542	.902

Treatment		N	Mean	Std. Deviation	Skewness		Kurtosis	
					Statistic	Std. Error	Statistic	Std. Error
Focused	PreMKT	25	.76	.254	-.085	.464	-2.174	.902
	PostMKT	25	.88	.217	-1.297	.464	-.354	.902
	DelayedMKT	25	.88	.217	-1.297	.464	-.354	.902
	PreExplicit	25	4.04	.815	-.202	.464	-.996	.902
	PostExplicit	25	4.48	.929	-.275	.464	-.280	.902
	DelayedExplicit	25	3.62	1.227	.140	.464	-.293	.902

Figure 4.1

UGJT pre-post-delayed posttest scores across groups
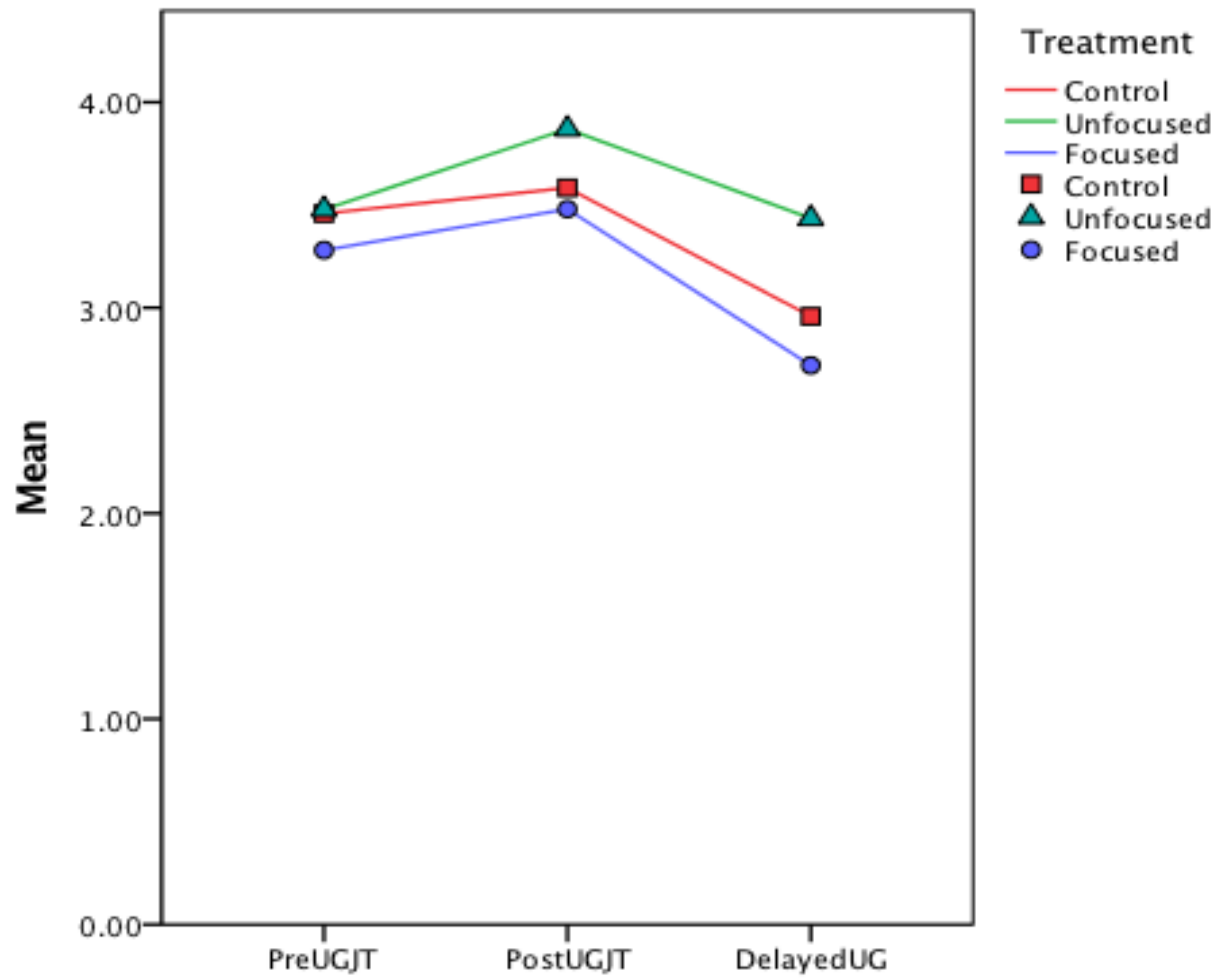
Figure 4.2

MKT pre-post-delayed posttest scores across groups
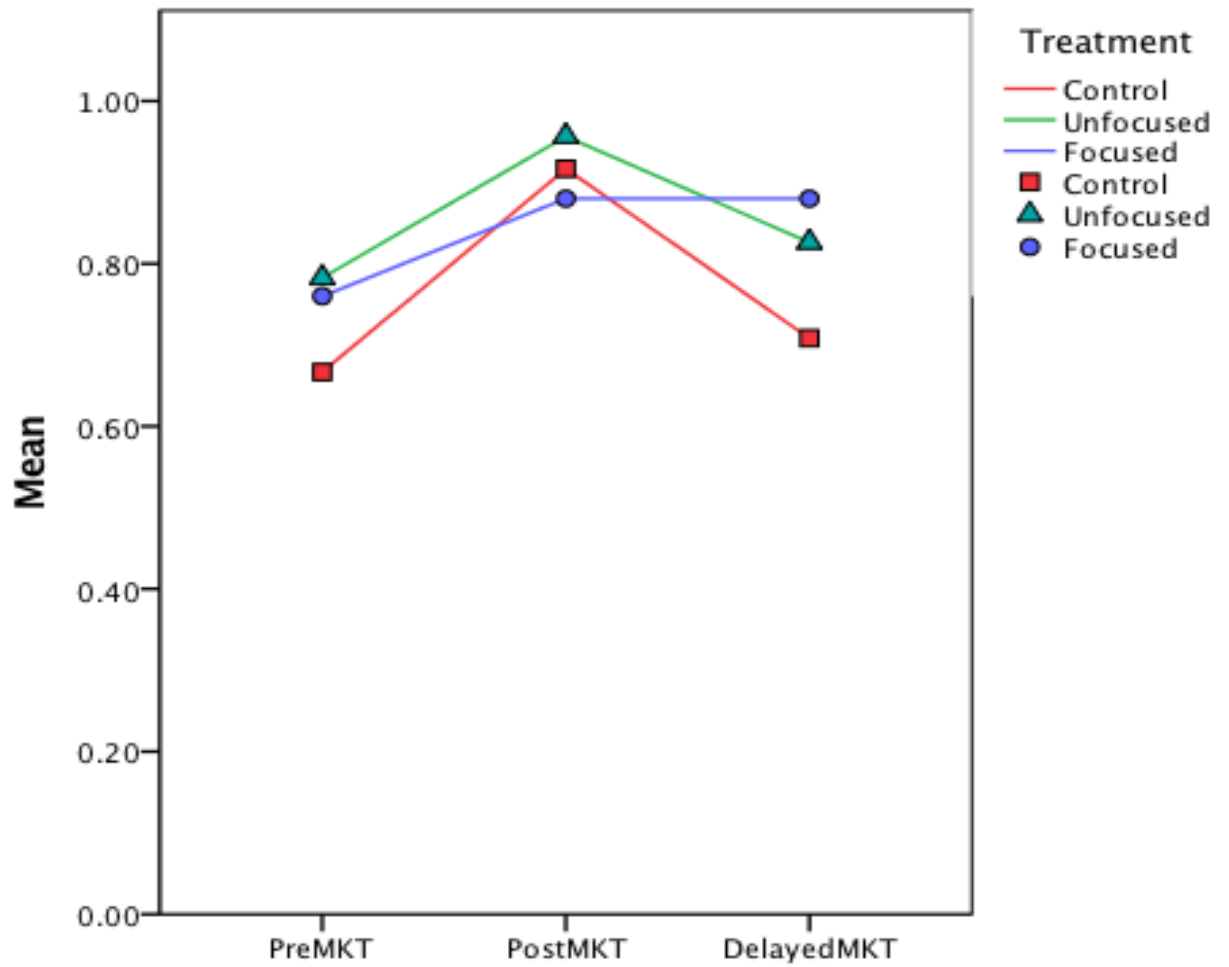
Figure 4.3

Combined explicit knowledge pre-post-delayed posttest scores across groups
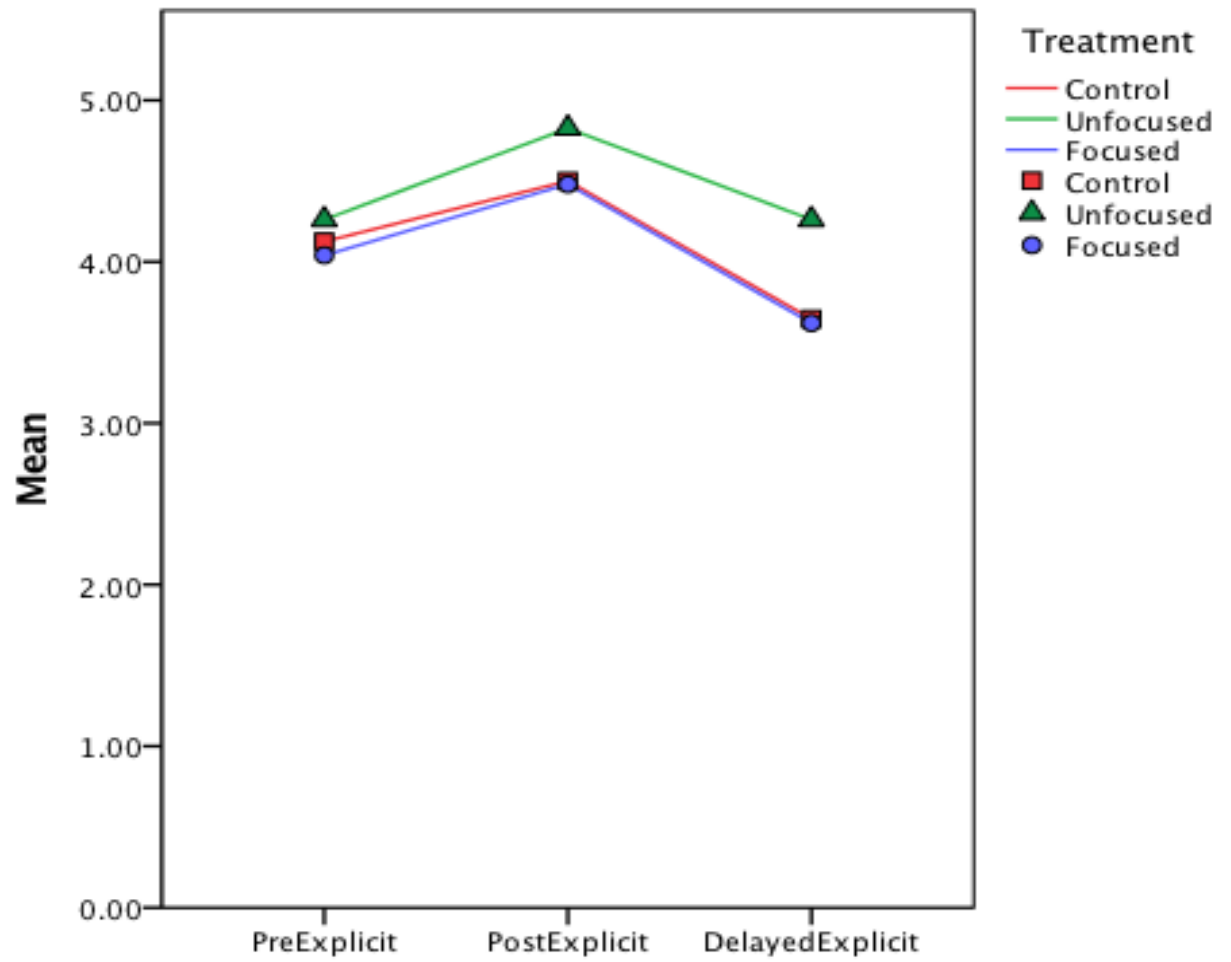
Untimed Grammaticality Judgement Test (UGJT). The results of the mixed design ANOVA with sphericity assumed ($\chi^2(2) = 5.423$, $p = .066$) revealed that although there were no significant group ($F(2, 69) = 2.805$, $p = .067$, partial eta squared (ηp^2) = .075, power =.534) and time-group interaction effects ($F(4, 138) = .731$, $p = .572$, partial eta squared (ηp^2) = .021, power =.231), there was a significant effect for time with a large effect size ($F(2, 138) = 9.963$, p < .001, partial eta squared (ηp^2) = .126, power =.983). The post hoc pairwise within group comparisons demonstrated that learners significantly improved their scores from pretest (Time 1) to posttest (Time 2) ($p = .045$) with a small effect size ($d = .270$). Nonetheless, their posttest gains (Time 2) markedly declined at the delayed posttest (Time 3) (p < .001), also with a small effect size ($d = .623$). Their delayed posttest scores (Time 3) were also significantly lowered than their pretest scores (Time1) ($p = .016$), with a small effect size ($d = .389$).

Table 4.2

UGJT results of the mixed design ANOVA

Variable	df	F	*p*-value.	ηp^2	Observed Power
Time	2,138	9.963	.000	.126	.983
Group	2, 69	2.805	.067	.075	.534
Time x Group	4,138	.731	.572	.021	.231

Metalinguistics Knowledge Test (MKT). The results of the mixed design ANOVA with sphericity assumed ($\chi^2(2) = 2.053$, $p = .358$) revealed that although there were no significant group ($F(2, 69) = 2.037$, $p = .138$, partial eta squared (ηp^2) = .056, power =.406) and time-group interaction effects ($F(4, 138) = 1.104$, $p = .357$, partial eta squared (ηp^2) = .031, power =.341), there was a significant time main effect with a large effect size ($F(2, 138) = 9.395$, p < .001, partial eta squared (ηp^2) = .120, power =.977). The post hoc pairwise within group comparisons demonstrated that learners significantly improved their scores from pretest (Time 1) to posttest (Time 2) (p < .001) with a medium effect size ($d = .744$). Nonetheless, their posttest gains (Time 2) declined at the delayed posttest (Time 3) (p = .015), with a small effect size ($d = .442$). Also, their delayed posttest scores (Time 3) were not significantly different from their pretest scores (Time 1) (p = .114).

Table 4.3

MKT results of the mixed design ANOVA

Variable	df	F	*p*-value	ηp^2	Observed Power
Time	2,138	9.395	.000	.120	.977
Group	2, 69	2.037	.138	.056	.406
Time x Group	4,138	1.104	.357	.031	.341

Combined Explicit Knowledge Test. The results of the mixed design ANOVA
using Huynh-Feldt estimates of sphericity ($\varepsilon = .94$) since the assumption of sphericity had
been violated ($\chi^2(2) = 8.961$, $p = .011$) indicated that although there were no significant
group (F (2, 69) = 2.166, $p = .122$, partial eta squared (ηp^2) = .059, power =.429) and
time-group interaction effects (F (3.754, 129.517) = .547, $p = .691$, partial eta squared
(ηp^2) = .016, power =.175), there was a significant effect for time with a large effect size
(F (1.877, 129.517) = 13.902, p < .001, partial eta squared (ηp^2) = .168, power =.997).
The post hoc pairwise within group comparisons demonstrated that learners significantly
improved their scores from pretest (Time 1) to posttest (Time 2) (p < .001) with a small
effect size ($d = .470$). Nonetheless, their posttest gains (Time 2) considerably declined at
the delayed posttest (Time 3) (p < .001), with a medium effect size ($d = .734$). Their
delayed posttest scores (Time 3) were not significantly different from their pretest scores
(Time 1) ($p = .078$).

Table 4.4

Combined Explicit Knowledge test results of the mixed design ANOVA

Variable	df	F	*p*-value	ηp^2	Observed Power
Time	1.877, 129.517	13.902	.000	.168	.997
Group	2, 69	2.166	.122	.059	.429
Time x Group	3.754, 129.517	.547	.691	.016	.175

Summary of Explicit Knowledge Test Results. The results of previous analyses showed statistically significant main effects of "time" on all explicit knowledge tests. This indicated that there were significant differences/changes in learners' explicit knowledge test scores over time. Given that the post hoc pairwise comparisons demonstrated that all participating groups performed significantly better in all their explicit knowledge measurements after treatment (Time 2), with small to medium effect sizes, the results suggested that, to some extent, the WCF provided for the treatment groups and the content feedback provided for the control group were effective in helping learners develop their explicit knowledge over time. Nonetheless, since no significant "group" and "time-group" interaction effects were observed, this suggested that different treatment conditions did not affect learners' learning differently. In other words, the non-significant "group" and "time-group" interaction effects suggested that both WCF (i.e., grammar feedback) and content feedback provided in this study were equally effective in helping learners improve their explicit knowledge of English plurals. Extended discussion about the results is provided in section 4.3.

4.3.1.2. Implicit knowledge test results

Descriptive Statistics. Descriptive statistics for the pre, post and delayed posttest scores of all the implicit knowledge tests were computed, summarized and presented in Table 4.5., Figure 4.4., Figure 4.5., and Figure 4.6. below.

As Table 4.5. and Figures 4.4. to 4.6. demonstrate, the general pattern detected was that all groups' implicit knowledge posttest (Time 2) and delayed posttest scores (Time 3) substantially increased from the pretest (Time 1). However, while the timed writing and combined implicit knowledge delayed posttest scores (Time 3) slightly decreased after the posttest (Time 2), the OEIT delayed posttest scores (Time 3) of all the participating groups remarkably increased from the posttest (Time 2), especially for the focused and control groups.

132

Table 4.5

Descriptive statistics of all the implicit knowledge tests

Treatment		N	Mean	Std. Deviation	Skewness		Kurtosis	
					Statistics	Std. Error	Statistics	Std. Error
Control	PreWriting	24	46.39	20.791	.015	.472	.218	.918
	PostWriting	24	61.02	17.771	.105	.472	-.406	.918
	DelayedWriting	24	54.27	18.351	.121	.472	.629	.918
	PreOEIT	24	.54	.883	1.513	.472	1.348	.918
	PostOEIT	24	.66	1.167	1.796	.472	2.319	.918
	DelayedOEIT	24	1.25	1.510	1.351	.472	.813	.918
	PreImplicit	24	49.10	22.111	-.036	.472	.342	.918
	PostImplicit	24	64.35	20.529	.324	.472	-.060	.918
	DelayedImplicit	24	60.52	24.224	.710	.472	.949	.918
Unfocused	PreWriting	23	48.08	17.392	.932	.481	1.392	.935
	PostWriting	23	60.89	16.080	-.435	.481	-.198	.935
	DelayedWriting	23	58.10	19.442	.113	.481	-.324	.935
	PreOEIT	23	.78	1.042	1.794	.481	3.504	.935
	PostOEIT	23	1.13	1.217	1.054	.481	.723	.935
	DelayedOEIT	23	1.26	1.483	1.611	.481	3.394	.935
	PreImplicit	23	52.00	20.446	.907	.481	.837	.935
	PostImplicit	23	66.54	17.597	.154	.481	-.377	.935
	DelayedImplicit	23	64.41	20.759	.678	.481	.408	.935

Treatment		N	Mean	Std. Deviation	Skewness		Kurtosis	
					Statistics	Std. Error	Statistics	Std. Error
Focused	PreWriting	25	38.02	14.485	.387	.464	-.569	.902
	PostWriting	25	62.62	17.706	-.330	.464	.188	.902
	DelayedWriting	25	53.20	14.218	.408	.464	-.295	.902
	PreOEIT	25	.36	.568	1.343	.464	1.036	.902
	PostOEIT	25	.68	.802	1.197	.464	1.462	.902
	DelayedOEIT	25	.92	.996	.994	.464	.167	.902
	PreImplicit	25	39.82	15.292	.347	.464	-.806	.902
	PostImplicit	25	66.02	19.235	-.145	.464	-.531	.902
	DelayedImplicit	25	57.80	15.814	.458	.464	-.377	.902

Figure 4.4

Timed writing pre-post-delayed posttest scores across groups
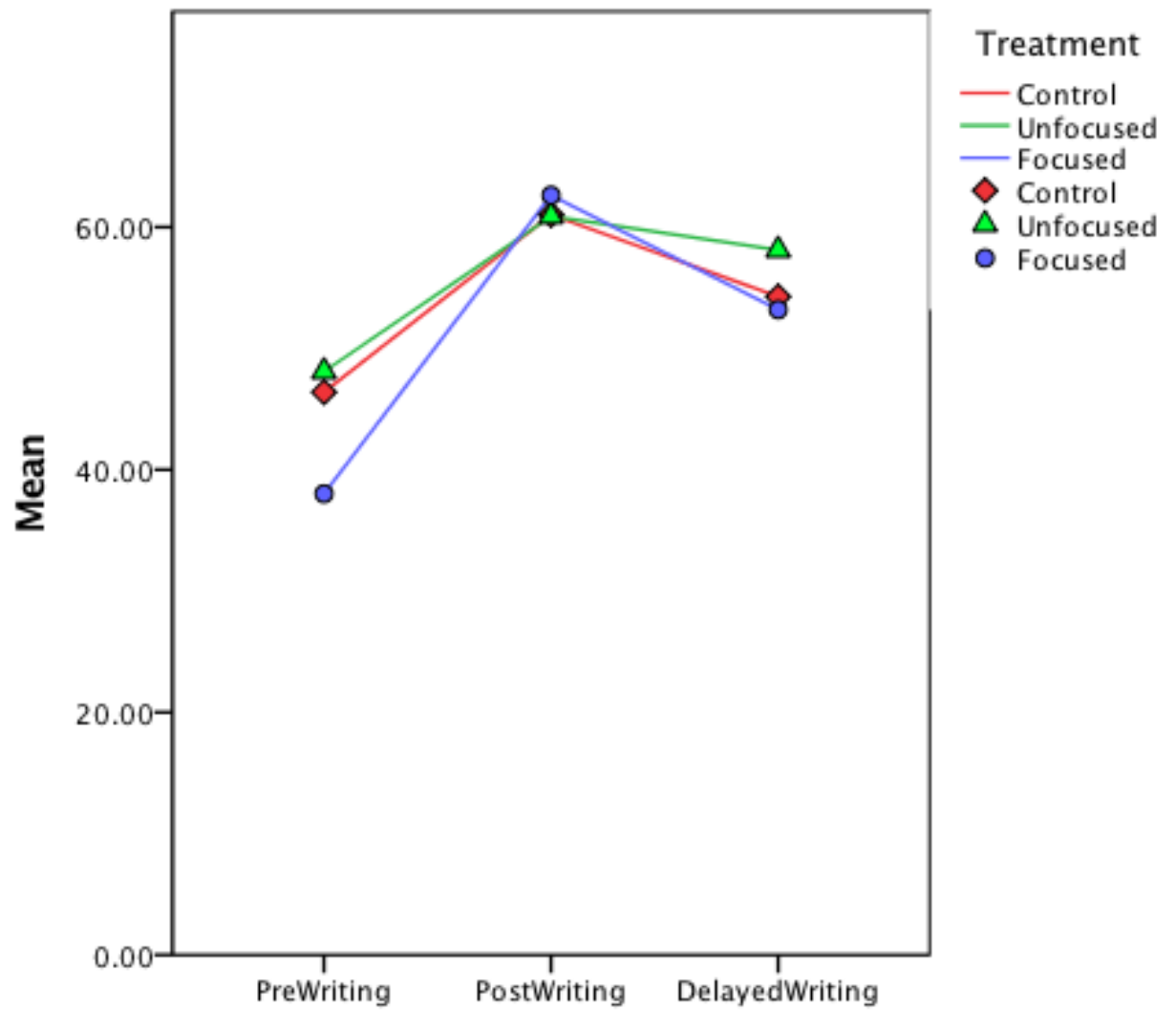
Figure 4.5

OEIT pre-post-delayed posttest scores across groups
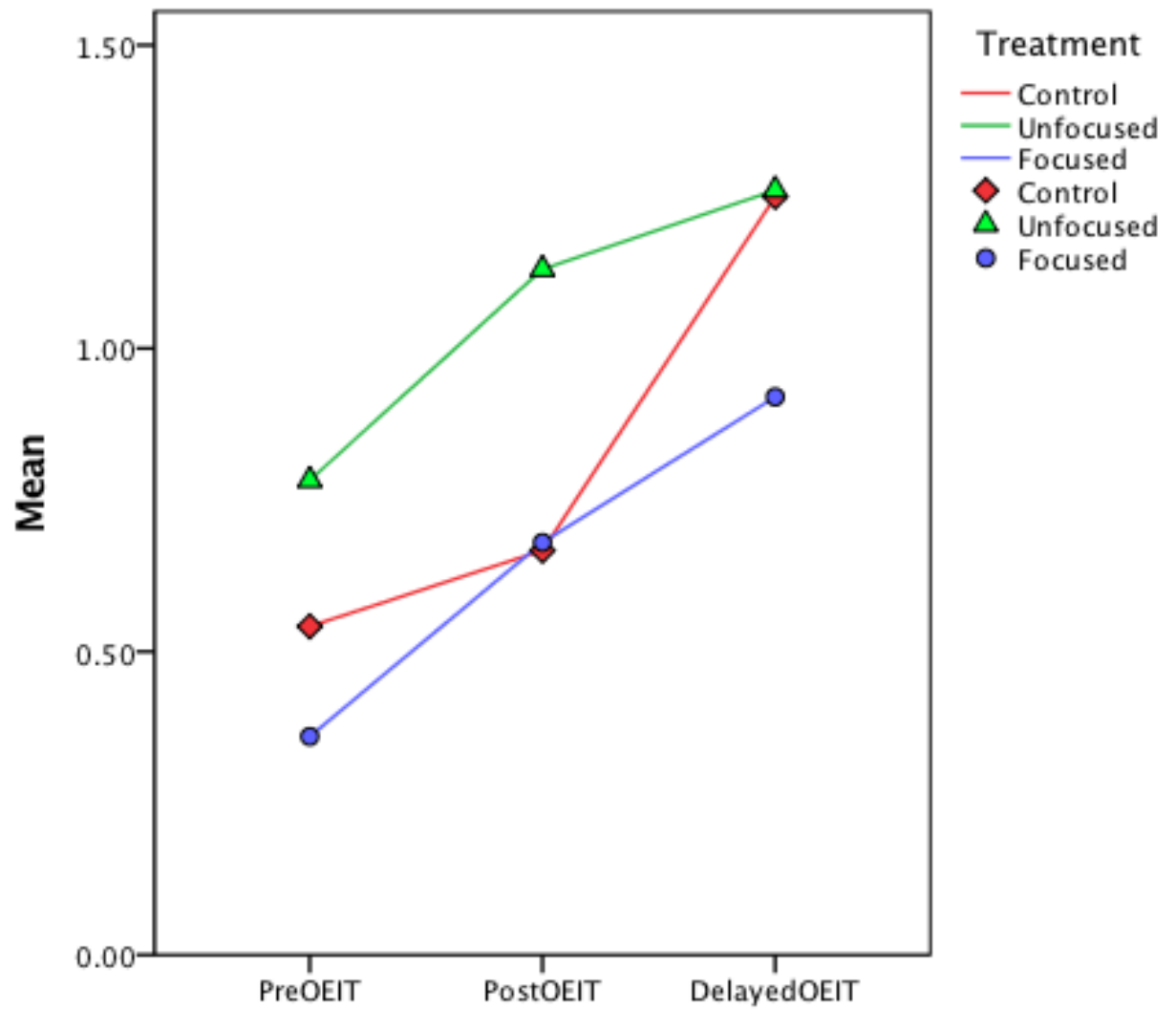
Figure 4.6

Combined implicit knowledge pre-post-delayed posttest scores across groups
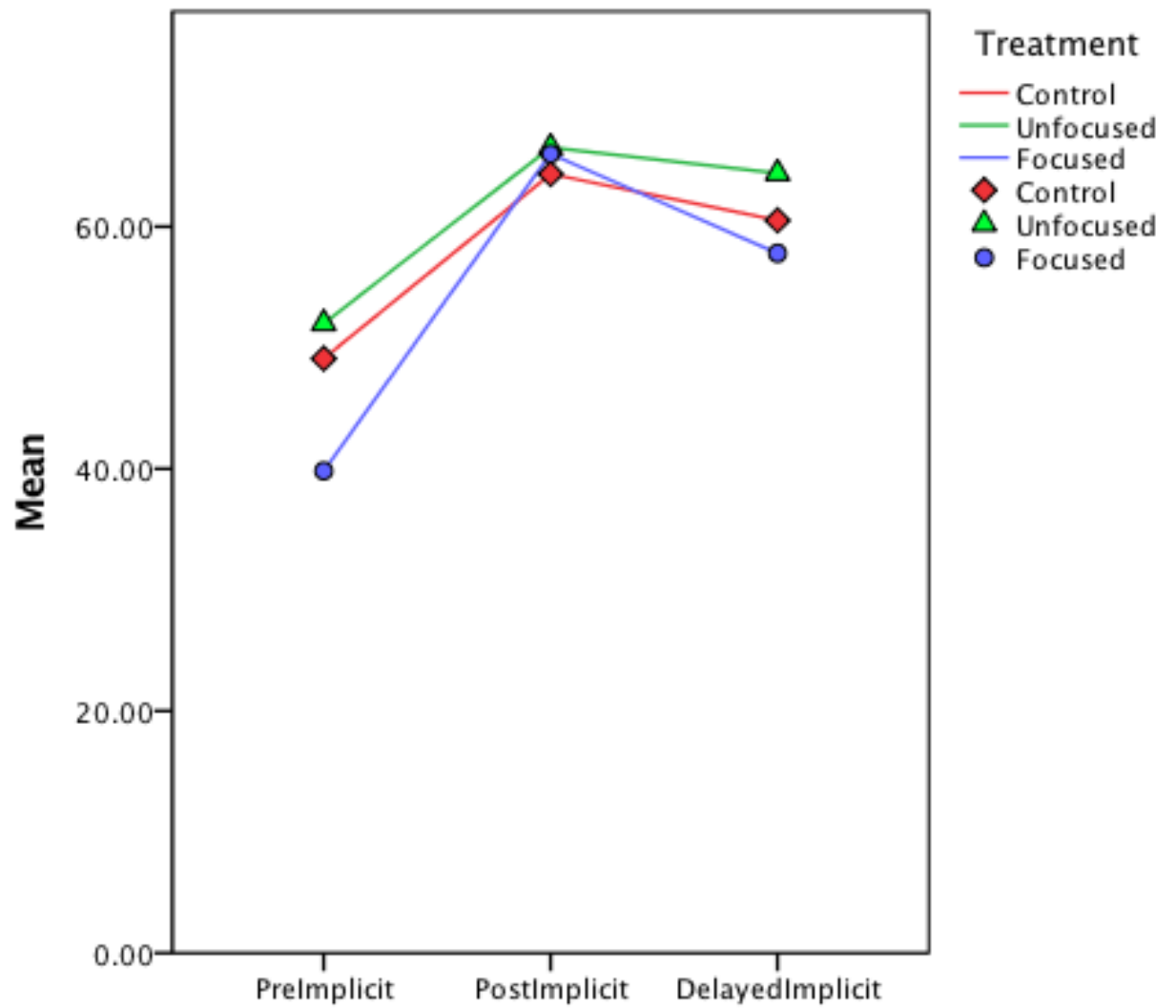
Timed Writing Task. The results of the mixed design ANOVA with sphericity assumed ($\chi^2(2) = 1.656$, $p = .437$) revealed that although there was no significant effect for group (F (2, 69) = .602, $p = .551$, partial eta squared (ηp^2) = .017, power =.146) and time-group interaction effect (F (4, 138) = 1.656, $p = .164$, partial eta squared (ηp^2) = .046, power =.499), there was a significant time effect with a large effect size (F (2, 138) = 34.503, p < .001, partial eta squared (ηp^2) = .333, power = 1.000). The post hoc pairwise within group comparisons demonstrated that learners significantly improved their scores from pretest (Time 1) to posttest (Time 2) (p < .001) with a large effect size (d = .990). Even though their posttest gains (Time 2) markedly declined at the delayed posttest (Time 3) (p = .002), with a small effect size (d = .368), their delayed posttest scores (Time 3) were significantly higher than their pretest scores (Time 1) (p < .001) with a small effect size (d = .624).

Table 4.6

Timed Writing results of the mixed design ANOVA

Variable	df	F	*p*-value	ηp^2	Observed Power
Time	2, 138	34.503	.000	.333	1.000
Group	2, 69	.602	.551	.017	.146
Time x Group	4, 138	1.656	.164	.046	.499

Oral Elicited Imitation Test (OEIT). The results of the mixed design ANOVA using Huynh-Feldt estimates of sphericity ($\varepsilon = .97$) since the assumption of sphericity had been violated ($\chi^2(2) = 6.285$, $p = .043$) indicated that although there were no significant effect for group ($F(2, 69) = 1.240$, $p = .296$, partial eta squared (ηp^2) = .035, power =.261) and time-group interaction effect ($F(3.881, 133.884) = .516$, $p = .718$, partial eta squared (ηp^2) = .015, power =.169), there was a significant effect for time with a large effect size ($F(1.940, 133.884) = 9.471$, $p < .001$, partial eta squared (ηp^2) = .121, power =.975). The post hoc pairwise within group comparisons demonstrated that learners significantly improved their scores from pretest (Time 1) to posttest (Time 2) ($p = .024$), and from posttest (Time 2) to delayed posttest (Time 3) ($p = .037$), all of which with small effect sizes ($d = .272$ and .261, respectively). Also, their delayed posttest scores (Time3) were significantly higher than their pretest scores ($p < .001$), with a small effect size ($d = .520$).

Table 4.7

OEIT results of the mixed design ANOVA

Variable	df	F	*p*-value	ηp^2	Observed Power
Time	1.940, 133.884	9.471	.000	.121	.975
Group	2, 69	1.240	.296	.035	.261
Time x Group	3.881, 133.884	.516	.718	.015	.169

Combined Implicit Knowledge Test. The results of the mixed design ANOVA with sphericity assumed ($\chi^2(2) = .044$, $p = .978$) revealed that although there were no significant group (F (2, 69) $= .938$, $p = .397$, partial eta squared (ηp^2) $= .026$, power $=.206$) and time-group interaction effects (F (4, 138) $= 1.442$, $p = .223$, partial eta squared (ηp^2) $= .040$, power $=.439$), there was a significant time effect with a large effect size (F (2, 138) $= 36.837$, p $< .001$, partial eta squared (ηp^2) $= .348$, power $= 1.000$). The post hoc pairwise within group comparisons demonstrated that learners significantly improved their scores from pretest (Time 1) to posttest (Time 2) (p $< .001$), with a large effect size ($d = .962$). Even though their posttest gains (Time 2) markedly declined at the delayed posttest (Time 3) ($p = .038$) with a small effect size ($d = .240$), their delayed posttest scores (Time 3) were significantly higher than their pretest scores (Time 1) (p $< .001$) with a medium effect size ($d = .692$).

Table 4.8

Combined Implicit Knowledge test results of the mixed design ANOVA

Variable	df	F	*p*-value	ηp^2	Observed Power
Time	2, 138	36.837	.000	.348	1.000
Group	2, 69	.938	.397	.026	.206
Time x Group	4, 138	1.442	.223	.040	.439

Summary of Implicit Knowledge Test Results. Similar to the previous explicit knowledge test results, only significant main effects of "time" were observed in these analyses, suggesting that there were significant changes in learners' implicit knowledge test scores over time. The results of the post hoc pairwise comparisons demonstrated that all participating groups performed significantly better in all their implicit knowledge measurements after treatment (Time 2), with small to large effect sizes. Moreover, all their delayed implicit knowledge test scores (Time 3) were significantly higher than their pretest scores (Time 1), implicating that the feedback may have a long-term effect on learners' L2 learning. Thus, the results suggested that, to some extent, the WCF provided for the treatment groups and the content feedback provided for the control group were effective in helping learners develop their implicit knowledge. However, considering that no significant "group" and "time-group interaction" effect was observed, this indicated that different types of feedback did not seem to affect learners' learning differently. In other words, both types of feedback, the WCF (i.e., grammar feedback) provided for the experimental groups and the content feedback provided for the comparison group, were equally effective in aiding the development of learners' implicit knowledge of English plurals. A more in-depth discussion regarding these results is offered in the subsequent discussion section of this chapter.

To address RQ4, a series of moderated multiple regressions were carried out to examine whether WMC moderated the effect of WCF on learners' plural learning outcomes. Bootstrapping analyses with 5,000 resamples using Hayes' PROCESS tool were performed to estimate the effects of two predictors: focused feedback (coded as 1) vs. unfocused feedback (coded as -1), and feedback (coded as 1) vs. no feedback (coded as -2). In each subsequent analysis, the dependent or outcome variable was the learners' posttest or delayed posttest scores, and the non-focal contrast was entered as a covariate of the focal contrast. Backward digit and OSpan scores (i.e., complex working memory scores) were entered as moderators in each model.

Prior to the analyses, the data of the learners whose OSpan processing scores did not reach the 80 % cut off point were removed[42]. Based on this criteria, five learners' data were taken out from the analyses. Consequently, only the data of 67 out of 72 learners were evaluated. In addition, the two sets of working memory test data (i.e., backward digit and OSpan scores) were coded as two separate moderators in order to avoid the method bias that could occur when one working memory test may be more challenging than the other for some learners[43] and therefore transforming the two sets of data into composite scores might potentially obscure the possibility to detect the effect of WMC on learners' performance.

To ensure that there were no significant differences between the WCF and control groups in terms of WMC scores, the WMC scores of the two groups were compared using

---

[42] To ensure that all learners completely engaged in the processing task, learners had to reach 80% accuracy on the processing task.
[43] Some learners were not good at making judgements about arithmetic problems, so they had difficult time performing OSpan task (under time pressure). However, the same group of learners seemed to perform reasonably well on the backward digit span task. Thus, by analyzing working memory test scores independently, there was a better chance to find WM effect if there was any, from either one of the two tests. As Sagarra (2017) suggests, the cognitively demanding level of the processing component of a working memory test (i.e., the level of difficulty of the processing task) seems to be one key factor accounting for the discrepant results found in working memory literature. Consequently, it might be more beneficial to employ more than one working memory measurement in a study as well as analyze them independently if possible.

a one-way between groups ANOVA. The results revealed that there were no observed differences between the two groups, indicating that both groups were equivalent at the onset of the study (see Appendix G). Descriptive statistics of the backward digit and OSpan scores, and the correlations between the two WMC measures and learners' test scores were then carried out to give a brief overview of the data. The results are summarized and displayed in Table 4.9., 4.10. and 4.11. below.

Table 4.9

Descriptive statistics of working memory tests

WM measure	Group	N	Mean	Std. Error	Std. Deviation
Backward	WCF	46	5.869	.261	1.771
	Control	21	5.714	.378	1.736
OSpan	WCF	46	67.434	.923	6.263
	Control	21	67.142	1.557	7.136

Table 4.10

Pearson Correlations of the WCF group

Test	Backward	OSpan	UGJT (T2)	UGJT (T3)	MKT (T2)	MKT (T3)	Writing (T2)	Writing (T3)	OEIT (T2)	OEIT (T3)
Backward	1.00	.440**	-.069	-.066	-.032	-.291*	.048	-.003	.125	.014
OSpan	.440**	1.00	-.30	-.320*	.137	-.121	.020	-.085	-.017	.094

Note: ** Correlation is significant at the .01 level (2 tailed).
     *. Correlation is significant at the .05 level (1 tailed).
     T2 = posttest and T3 = delayed posttest

Table 4.11

Pearson Correlations of the Control group

Test	Backward	OSpan	UGJT (T2)	UGJT (T3)	MKT (T2)	MKT (T3)	Writing (T2)	Writing (T3)	OEIT (T2)	OEIT (T3)
Backward	1.00	-.017	-.172	.133	-.082	-.121	.011	-.078	.225	.074
OSpan	-.017	1.00	.125	-.143	-.007	-.041	-.307	.340	.258	.081

Note: ** Correlation is significant at the .01 level (2 tailed).
    *. Correlation is significant at the .05 level (1 tailed).
    T2 = posttest and T3 = delayed posttest

As shown in Table 4.10., for the feedback group, backward digit and OSpan scores showed to have some significant correlations with learners' performance on the delayed MKT and UGJT. However, no significant correlation was observed between the two working memory measures and the timed writing task and OEIT scores. For the control group, as shown in Table 4.11., no significant correlation between the two working memory measures and learners' performance was detected. Based on the brief overview of the data presented above, WMC seemed to be associated with the performance of the learners who received WCF treatment but did not relate to the performance of the learners of the control group.

In what follows, the results of inferential statistical analyses (i.e., the moderated multiple regressions) are provided. The analyses were conducted to provide a more robust result as to whether WMC truly had a moderating effect on the performance of the learners who received WCF.

Untimed Grammaticality Judgment Test (UGJT). The results showed that the final model (i.e., focused vs. unfocused WCF, and backward digit and OSpan) did not account for significant variance in the learners' UGJT posttest scores of the feedback groups (R^2= .101, $F(6, 60)$ = 1.123, p = .360). The inspection of the coefficients showed that none of the predictors and their interactions were significant ($|b|$s < .001, ps > .977).

Nonetheless, the final model accounted for significant variance in the learners' UGJT delayed posttest scores of the feedback groups (R^2= .190, $F(6, 60)$ = 2.340, p = .043). The inspection of the coefficients showed that only the OSpan (the moderator) was negatively associated with the UGJT delayed posttest scores (b = -.056, t = -2.703, p = .009), indicating that the learners' gains decreased when the OSpan score increased or vice versa. However, the inspection of the coefficients showed that none of the predictors and their interactions were significant ($|b|$s < .051, ps > .605).

When examining if WMC had any moderating effect on the UGJT scores of the control group learners versus those of the feedback groups (i.e., focused and unfocused WCF groups), the results showed that the final model (i.e., WCF vs. Control, backward digit and OSpan) did not account for significant variance in the learners' posttest scores (R^2= .064, $F(6, 60)$ = .681, p = .666). The inspection of the coefficients also showed that none of the predictors and their interactions were significant ($|b|$s < .004, ps > .826). Similarly, the final model did not account for significant variance in the learners' delayed posttest scores (R^2= .180, $F(6, 60)$ = 2.197, p = .056). The inspection of the coefficients showed that none of the predictors and their interactions were significant ($|b|$s < -.011, ps > .835).

Metalinguistics Knowledge Test (MKT). The results of the logistic regression[44] showed that the final model did not account for significant variance in the learners' MKT posttest scores of the feedback groups (*Nagelkerke R^2* = .130, $\chi^2(6)$ = 54.574, p = .498). The inspection of the coefficients showed that none of the predictors and their interactions were significant (*Exp(B)*|s < -.119, ps > .617). Similarly, the final model did

---

[44] The MKT posttest scores only had two levels (i.e., the outcome variable was binary): 0.5 and 1, so the logistic regression was adopted in place of normal moderated multiple regression.

not account for significant variance in the learners' delayed posttest scores ($R^2 = .096$, F (6, 60) = 1.063, p = .395). The inspection of the coefficients showed that none of the predictors and their interactions were significant ($|b|$s < .000, ps > .998).

When examining if WMC had any moderating effect on the MKT scores of the control group learners versus those of the feedback groups, the results of the logistic regression showed that the final model did not account for significant variance in the learners' posttest scores (*Nagelkerke R^2* = .134, χ^2(6) = 54.296, p = .477). The inspection of the coefficients showed that none of the predictors and their interactions were significant (*Exp(B)*|s < -.047, ps > .733). In addition, the final model did not account for significant variance in the learners' delayed posttest scores (R^2= .085, F (6, 60) = .932, p = .479). The inspection of the coefficients showed that none of the predictors and their interactions were significant ($|b|$s < .01, ps > .008).

Timed Writing Task. The results showed that the final model did not account for significant variance in the learners' timed writing posttest scores of the feedback groups (R^2= .053, F (6, 60) = .560, p = .760). The inspection of the coefficients showed that none of the predictors and their interactions were significant ($|b|$s < .914, ps > .501). Similarly, the final model did not account for significant variance in the learners' delayed posttest scores (R^2= .038, F (6, 60) = .400, p = .876). The inspection of the coefficients showed that none of the predictors and their interactions were significant ($|b|$s < .287, ps > .867

When examining if WMC had any moderating effect on the timed writing scores of the control group learners versus those of the feedback groups, the results showed that the final model did not account for significant variance in the learners' posttest scores (R^2= .035, F (6, 60) = .362, p = .900). The inspection of the coefficients also showed that none of the predictors and their interactions were significant ($|b|$s < .141, ps > .882). Similarly, the final model did not account for significant variance in the learners' delayed posttest scores (R^2= .058, F (6, 60) = .615, p = .717). The inspection of the coefficients showed that none of the predictors and their interactions were significant ($|b|$s < -.017, ps > .989).

Oral Elicited Imitation Test (OEIT). The results demonstrated that the final model did not account for significant variance in the learners' OEIT posttest scores of the feedback groups ($R^2 = .076$, $F(6, 60) = .825$, $p = .555$). The inspection of the coefficients showed that none of the predictors and their interactions were significant ($|b|s < .002$, $ps > .922$). Likewise, the final model did not account for significant variance in the learners' delayed posttest scores ($R^2 = .035$, $F(6, 60) = .362$, $p = .900$). The inspection of the coefficients showed that none of the predictors and their interactions were significant ($|b|s < -.005$, $ps > .906$).

When examining if WMC had any moderating effect on the OEIT scores of the control group learners versus those of the feedback groups, the results showed that the final model did not account for significant variance in the learners' posttest scores ($R^2 = .100$, $F(6, 60) = 1.107$, $p = .369$). The inspection of the coefficients also showed that none of the predictors and their interactions were significant ($|b|s < .002$, $ps > .913$). Likewise, the final model did not account for significant variance in the learners' delayed posttest scores ($R^2 = .028$, $F(6, 60) = .285$, $p = .942$). The inspection of the coefficients showed that none of the predictors and their interactions were significant ($|b|s < .002$, $ps > .983$).

Summary of Research Question Four. The results of the moderated multiple regression analyses revealed that WMC did not moderate the extent to which learners benefited from WCF. Even though the OSpan scores of the WCF group showed to have significant negative association with the UGJT delayed posttest scores, this could be a result of other confounding factors, given that the OSpan scores only predicted one set of test scores and its interaction effect was not significant. A more in-depth discussion concerning these results is offered in the next section.

4.4. Discussion and conclusion

4.4.1. Discussion

Effects of WCF on the Development of Learners' Explicit and Implicit Knowledge

      The first and second research questions (RQ1 and 2) of this study explored whether WCF had any effect on the development of learners' explicit and implicit knowledge of English plurals. The results of the inferential statistics demonstrated that WCF treatment led to significant improvement in learners' explicit and implicit knowledge test performance as anticipated. The WCF effect on learners' implicit knowledge gains also appeared to be sustained over time (three months after treatment). The results, thus, suggest that direct *"focused"* and *"unfocused"* WCF provided in the current study were effective in helping learners develop their explicit and implicit knowledge of English plurals.

      These results concur with the findings of a number of previous WCF studies which also found "*direct"* WCF effective in helping L2 learners improve their grammatical accuracy (e.g., Benson & DeKeyser, 2019; Bitchener & Knoch, 2010b; Rummel, 2014; Shintani et al., 2014; Stefanou & Révész, 2015; Suzuki et al., 2019; Van Beuningen et al., 2012). Importantly, the results of the present study provide further evidence in support of previous research findings (e.g., Nemati et al., 2019; Rezazadeh et al., 2015; Shooshtari et al., 2019) which attest the effectiveness of WCF as a form-focused tool helping L2 learners develop their explicit and implicit knowledge. It is posited that the WCF provided in this study was effective in facilitating the development of learners' explicit and implicit knowledge because it provided learners more opportunities to notice the gaps in their interlanguage. These noticing opportunities helped these learners to review and fortify their pre-existing knowledge, leading to a better understanding of the underlying rules of the target structure. With more exposure and practicing opportunities, learners then integrated their newly developed knowledge with their pre-existing knowledge leading to the development of their explicit and implicit knowledge (also see N. Ellis, 2005, 2012, 2016a; Wulff & N. Ellis, 2018 for more details on how L2 knowledge is developed).

Notwithstanding, the results of this study appeared to diverge from those of Shintani and R. Ellis' (2013) pioneer study[45] which found the beneficial effect of the metalinguistic explanation (ME) on the development of learners' explicit knowledge but did not find any positive effect of ME nor "*direct*" WCF on the development of learners' implicit knowledge. There are, however, a number of factors contributing to the contradictory results between their study and the current study. One of those is the difference in the linguistic feature targeted in each study. To be more specific, while Shintani and R. Ellis (2013) investigated the effectiveness of WCF on the indefinite article, which is considered to be a more complex, idiosyncratic and, thus, a less treatable type of error[46] (Shintani & R. Ellis, 2013; Shintani et al., 2014), this study explored plural structure, a more treatable error. For this reason, it is probable that for WCF to have a positive effect on indefinite articles targeted in Shintani and R. Ellis' (2013) study, more time and intensive exposure to corrective feedback are needed, unlike a more treatable plural error targeting in this study which probably needs less time and exposure to the input to yield a positive result (also see Shintani et al., 2014, for similar discussion about the efficacy of WCF on learners' article usage). Given that the complexity and difficulty of the target grammatical structures could potentially affect the extent to which particular grammatical structures could be treated via feedback, that is, the effectiveness of feedback can vary depending on the nature of the target linguistic structure (Adams et al., 2011; Bitchener, Young & Cameron, 2005; Ferris, 1999; Frear, 2012; Mifka-Profozic, 2015; Li & Vuono, 2019; Nassaji, 2016; Shintani, et al., 2014), this may in part explain why the WCF provided in Shintani and R. Ellis'(2013) study was ineffective, but the WCF provided in this study was effective in helping learners develop both their explicit and implicit knowledge.

Another plausible explanation lies in the number of correction learners in each

---

[45] Shintani and R. Ellis' (2013) study is believed to be the first study within the WCF filed that investigated the effect of WCF on learners' explicit and implicit knowledge development.

[46] Treatable errors are errors that occur "in a patterned, rule-governed way" as opposed to untreatable errors in which "there is no handbook or set of rules students can consult" (Ferris, 1999, p. 6). However, Shintani et al. (2014) suggest that "what constitutes 'treatability' is not just a question of whether or not a feature is rule-based but also the complexity of the rule-based structure" (p. 12).

study received. While learners in Shintani and R. Ellis' (2013) study only received one correction on a single piece of writing, the learners in this study received six corrections on six pieces of writing during the 6-week treatment sessions (excluding the pretest, posttest and delayed posttest timed writing tasks). Providing that implicit knowledge needs an extensive amount of time and L2 exposure to develop (DeKeyser, 2015; N. Ellis, 2005; Paradis, 2009; Suzuki & DeKeyser; 2015, 2017), it is not unexpected that a one-shot feedback treatment provided in the previous study was unable to help L2 learners develop their implicit knowledge. Indeed, Shintani and R. Ellis (2013) acknowledged the possibility that they might not find a positive effect of WCF on the development of learners' implicit knowledge in their study as they noted:

> [I]mplicit knowledge is only developed gradually as learners are exposed to more L2 input and opportunities for production. This would suggest that one-shot written error feedback of the type investigated in many studies (including the one we report below) may contribute to explicit knowledge but will have no effect on their implicit knowledge. (p. 287-288)

As other researchers (e.g., Benson & DeKeyser, 2019; Shintani et al., 2014) also point out, for the formed-focused feedback to be fully effective in yielding a durable effect, that is, having a positive effect on the development of learners' implicit knowledge, more episodes of feedback treatment are necessary. Considering that the learners in Shintani and R. Ellis' (2013) study only received a one-shot treatment and limited production opportunity, in contrast to the number of corrections and writing practice opportunity provided in the current study, this also partially explains why the previous and current study yielded contradictory results.

Another factor accounting for the differing results concerns the different implicit knowledge measurements employed in each study. In Shintani and R. Ellis' (2013) study, only the timed narrative writing task was used to measure learners' implicit knowledge. In contrast, the current study employed both the timed argumentative writing task and oral elicited imitation test (OEIT) as implicit knowledge measurements. The OEIT was adopted for this study because a large number of researchers (e.g., R. Ellis, 2005a, 2009c; Kim & Nam, 2017; Sarandi, 2015; Tomita, Suzuki & Jessop, 2009; Yan, Maeda, Lv & Ginther, 2016) have warranted the OEIT as one of the most valid measures of implicit

knowledge available to date[47]. As mentioned earlier (see chapter 2 literature review), even though the timed writing task seems to have high face validity responding to the current research context (focusing on writing modality), its reliability and validity as an implicit knowledge measurement are still open for discussion (Polio, 2012; Shintani & R. Ellis, 2013). Hence, the difference in the choice of outcome measures used in each study may also influence the test results to some extent (Lyster & Saito, 2010; Mackey & Goo, 2007; Nassaji, 2016; Norris & Ortega, 2000; Révész, 2012).

Indeed, many SLA researchers (e.g., R. Ellis, 2004; R. Ellis, Loewen, & Erlam, 2006; Hama & Leow, 2010; Norris & Ortega, 2003; Rebuschat, Hamrick, Riestenberg, Sachs, & Ziegler, 2015; Révész, 2012) have endorsed and emphasized the need to include various types of assessment measures to elicit learners' behaviors, since different behaviors "may emerge in different contexts" (Révész, 2012, p. 98) and failing to do so might lead to the conclusion that is based on "a lack of evidence" instead of, "evidence for a lack of emerge" (Norris & Ortega, 2003, p.733). Further, given that until recently there is still no pure implicit knowledge measurement available[48] (R. Ellis, 2009c; Gutiérrez, 2013; Shintani & R. Ellis, 2013), it is even more necessary that multiple measures of implicit knowledge should be employed in order to obtain a more reliable outcome (e.g., Hama & Loew, 2010; Rebuschat et al., 2015; Suzuki & DeKeyser, 2017; Suzuki, Nakata & DeKeyser, 2019). Providing that the current study employed multiple measurements, instead of one measurement, which might increase a high probability of

---

[47] Apart from the OEIT task, RT (reaction time) method such as the word-monitoring task (see Suzuki & DeKeyser, 2015; Godfroid, 2016), the self-paced reading task (see Foote, 2011; Roberts & Liszka, 2013) and the visual-world task (see Grüter, Lew-Williams, & Fernald, 2012; Hopp, 2013) have begun to be well accepted among SLA researchers. However, in Suzuki and DeKeyser's (2017) recent study which attempted to test the validity of the three aforementioned measures, the authors concluded that "the validity of measures targeting implicit knowledge should be examined further. A weak convergence across the three measures of implicit knowledge is a serious limitation" (p. 779). This, thus, indicates that the validity and credibility of these newly developed measurements are still subjected to further examination.

[48] R. Ellis (2009c) points out that thus far there are no such "pure" measurements of implicit knowledge because when learners undertake the test, they will arbitrarily use any resources at their convenience. Yet, R. Ellis comments that albeit a lack of pure measurement, it is possible to design the test that biases learners to rely more on one type of knowledge than the other (R. Ellis, 2009c; Shintani & R. Ellis, 2013).

systematic errors or lead to a lack of evidence as elucidated, it could be presumed that the implicit knowledge test results of the current study reliably reflects the development of learners' implicit knowledge of English plurals at least within the context of this study.

To this point, one might argue that the current study also targeted different linguistic structures, provided different number of corrections and used different outcome measures than those previous studies (i.e., Nemati et al., 2019; Rezazadeh et al., 2015; Shooshtari et al., 2019), which also found the effect of WCF on learners' implicit knowledge development but why the finding of the current study appeared to be aligned with the findings of those studies. Two significant features that the current and those previous studies had in common (but Shintani and R. Ellis' study did not have) that might influence the results are 1) the multiple WCF treatment sessions our studies provided, and 2) the population targeted in our studies.

To be more specific, except for Rezazadeh et al.'s (2015) study, which provided two WCF treatment sessions, others (i.e., Nemati et al., 2019; Shooshtari et al., 2019) and the current study provided at least four WCF treatment sessions. As elucidated, time and intensive exposure to language input are essential for the development of implicit knowledge. It is, thus, logical that the multiple WCF treatment session design adopted in our studies could help learners develop their implicit knowledge. In contrast, the one-shot WCF treatment in Shintani and R. Ellis' (2013) study did not. In addition, while Shintani and R. Ellis' (2013) study targeted ESL learners, our studies' population were EFL learners living in their native countries. In this case, it can be argued that different educational settings may also account for the discrepant results. The findings of corrective feedback and meta-analysis studies on both oral and written corrective feedback (e.g., Kang & Han, 2015; Li, 2010; Li & Vuono, 2019; Loewen et al., 2009; Lyster & Mori, 2006; Sheen, 2004; Wang & Li, 2021) provide support of this speculation suggesting that educational setting (i.e., ESL vs. EFL) is one of the key moderators varying the degrees of effectiveness of feedback on learners' L2 learning outcomes.

For example, in Loewen et al.'s (2009) study, the authors found from the survey conducted on 724 L2 learners in the US that while ESL learners tended to focus on improving their oral communicative skill, EFL learners were more attentive to grammar instruction and error correction. In Li's (2010) meta-analysis, corrective feedback was

also found to be more effective in an EFL than ESL context. Loewen et al. (2009) and Li (2010) believe that error correction is more effective in an EFL than ESL setting because EFL learners may have more positive attitude towards error correction than those in the ESL context, and this may lead to more successful feedback uptake.

Likewise, in Biber et al.'s (2011) meta-analysis, the authors found that L1 English writers achieved higher gains on writing when receiving oral feedback and achieved no or small gains when receiving written feedback. The result implicates a high probability that learners in an ESL setting (who shared quite similar or identical learning environment to L1 English writers) might also share similar attitude on error correction with those native writers. That is, they might also prefer oral to written feedback, making WCF less effective in an ESL than in an EFL setting. Taken together, given that there is a possibility that ESL learners are likely to benefit less from WCF than those EFL learners, this may partly explain why Shintani and R. Ellis did not find the positive effect of WCF on learners' implicit knowledge while the opposite is true for the current and other aforementioned studies.

However, it is important to point out here that, despite the aforementioned, it is also possible that WCF might not be the main factor helping these learners develop their explicit and implicit knowledge but the opportunity for writing practice or task repetition. Given that there were no differential effects between the provided WCF and content feedback on learners' learning outcomes (see discussion of RQ3 in the subsequent section), the current finding, to some extent, can also be interpreted as feedback (either grammar or content feedback) might not be the main factor helping these learners develop their L2 knowledge. Instead, it could be the effect of writing practice opportunity or task repetition that enhanced learners' L2 development considering that it seemed to be the sole factor (apart from their background knowledge) that the three groups had in common. In other words, increases in learners' accuracy on explicit and implicit knowledge tests may not be the result of the effect of WCF or content feedback but possibly the effect of writing practice or task repetition. It is as Adams (2003) mentioned in her feedback (reformulation writing) study which found task repetition, not only the effect of feedback, contributed to learning, that it is essential for researchers to be conservative with the interpretation of the treatment effect as learners' improvement

153

could also be a result of other factors other than the provided treatment. Consequently, any interpretation drawn from the above findings may have to be exercised with caution, acknowledging the possibility that writing practice or task repetition could actually be the main factor facilitating the development of learners' L2 knowledge at least within the context of this study.

Differential Effects of Different Types of WCF on L2 Development

The third research question (RQ3) of the present study investigated whether there were any differential effects between different types of WCF provided in this study. The results of the inferential statistics revealed that there were no significant differences in learners' explicit and implicit test performance between those from the WCF groups and the control group, as all participating groups made remarkable progress after treatment. In other words, the result suggests that the WCF provided for the treatment groups and the content feedback (i.e., non-grammar feedback) provided for the control group were equally effective in helping learners improve their explicit and implicit knowledge of English plurals within this study context. The results were, however, quite unexpected given that it was predicted that focused WCF might be more effective than unfocused WCF. In addition, it was unanticipated that the content feedback provided for the control group could be equally effective as grammar feedback in helping learners improve their explicit and implicit knowledge. Several confounding factors are posited to influence the findings.

Regarding the non-significant difference between focused and unfocused WCF, it is possible that because of the meaning-focused instruction (i.e., implicit learning condition) this study adopted that did not emphasize the role of WCF during treatment, learners who received these two types of WCF may not fully benefit from the feedback they received (as they might not pay enough attention to the feedback); thus, the differential effect might not be evident to be detected (if there was any). In the same vein, the short treatment period (6 weeks) the study provided may not allow the effects of these two types of WCF to fully emerge; therefore, the differential effect could not be detected. With a longer treatment period, the differential effects may become apparent if there is any.

154

Regarding the fact that the control group learners who only received content feedback could significantly improve their plural knowledge similar to those who received WCF, three main factors are posited to have influenced the result. The first factor concerns learners' educational context (i.e., an EFL setting) that favors grammar learning. As previously mentioned, most Thai EFL learners are oriented to grammar learning because grammar teaching is the main teaching approach in Thailand. Thus, like most EFL learners in Asia, most Thai learners have been trained to be mindful of their grammar usage since an early age. Consequently, even though the control group learners only received content feedback, the feedback may not only help raise their awareness towards improving their content but also raise their awareness towards improving their grammar usage.[49] In addition, because of their educational context (i.e., a highly form-focused context) which values grammar learning, it is possible that these learners may have cultivated a high level of grammatical awareness or grammatical sensitivity aptitude which facilitates them to notice and use their pre-existing grammatical knowledge to self-correct their errors, especially during the absence of grammar feedback or form-focused instruction. In other words, it is posited that learners' high grammaticality sensitivity aptitude resulting from their educational context may have helped these control group learners become more oriented towards forms (i.e., become more aware of their grammar usage) and this led to subsequent improvement on their plural test scores.

According to Stefanou and Révész (2015), "Grammatical sensitivity is an ability…to recognize different syntactic patterns and grammatical functions of words in a given sentence structure, irrespective of knowledge of grammatical terminology" (p.265). It is one of the two components of language analytic ability aptitude: an ability to infer and generalize language rules (also see Carroll, 1981; Skehan, 2002). Grammatical sensitivity has also been posited in the literature that it is associated with learners' improved performance on the target linguistic features (Benson & DeKeyser, 2019; Erlam, 2005; Qi & Lapkin, 2001; Rosa & Leow, 2004; Sachs & Polio, 2007; Shintani & R. Ellis, 2013; Stefanou & Révész, 2015). Skehan (2015) comments that in the learning

---

[49] These learners had been trained to be mindful of their grammar usage for a long period of time (i.e., they were usually concerned about their grammar), so it is possible that content feedback may also remind and orient them to pay attention to their grammar usage apart from raising their awareness about their writing content.

context where the target grammatical feature is redundant and/or less salient, aptitude may compensate by making that grammatical point clearer so as to be noticed. Taken together, it could be conjectured that with the high level of grammatical sensitivity aptitude the control group learners may have possessed combined with the effect of feedback as a consciousness-raising tool, these learners may notice and become more aware of their grammatical usage leading to their grammatical improvement even in the absence of grammar feedback.

Learners' proficiency level is another moderating factor influencing the yielded result. Considering that all learners in this study were intermediate learners and all had been receiving instruction on English plurals since primary school (in other words, they were already familiar with the target structure rules), it is more likely that these learners had already established plural metalinguistic knowledge, either partially or fully. As a result, for learners at this proficiency level, it can be argued that even in the absence of grammar feedback, these learners should be able to correct their plural errors by themselves as long as they notice or are aware of the errors in their writing. Consequently, for the case of the control group, as previously elucidated, with their seemingly high grammatical sensitivity due to the influence of their educational context (i.e., previous learning experiences), when prompted by content feedback, these learners may become more aware and careful of their grammar usage. And with their current language ability or proficiency level (i.e., low-intermediate), these learners may then be able to notice and self-correct their own grammatical errors leading to some grammatical gains even in the absence of grammar feedback. This explains why the control group learners could make some advancement on their plural knowledge, similar to those who received grammar feedback. Notwithstanding, it is posited that if these learners were at a lower proficiency level (i.e., beginners) and did not have established metalinguistic knowledge of the target structure beforehand, receiving or not receiving grammar feedback may lead to differing results. For learners at a lower proficiency level, even though they may intuitively feel that there are errors in their writing or can actually detect the errors, they might not have adequate knowledge to self-correct those errors. In addition, providing that these learners were intermediate learners with some English plural background knowledge, which left them smaller room to grow when compared to

beginner learners, the short intervention period (6 weeks) they received might not be adequate for them to make observable gains on their test performance. This, thus, led to the small between-group discrepancy which made it more difficult for the employed statistical test to detect the differential effects between different feedback treatment conditions. In part, this explains why there were no significant differences in the effects of the three feedback types detected within the context of this study.

Apart from learners' educational context and proficiency level that seemed to have influenced the result, task instruction or instructional context may also have influenced the finding. As previously discussed, because Thai learners have a tendency to highly focus on forms as a result of their learning experiences or educational context, task instruction in this study was then designed to be more meaning-focused (i.e., implicit learning instruction) so that it could decrease learners' obsessive focus on forms. Consequently, during treatment, learners were instructed to only focus on developing their content. Even though some of them received grammar feedback (i.e., those in the focused and unfocused WCF groups), the role of feedback was not emphasized during treatment. Learners were only instructed to look at the feedback for five minutes and then wrote a new piece of essay. It is speculated that this instructional context (i.e., implicit learning condition) where the language forms (i.e., feedback) were not emphasized may have lessen the effectiveness of the given WCF. That is, the study's highly meaning-based instructional context where the role of form-focused instruction (i.e., feedback) was diminished might predispose learners more towards implicit learning and this learning condition may not allow learners to fully benefit from the given WCF (i.e., explicit or form-focused instruction). As Ullman (2016) points out, "the learning and/or retrieval of knowledge in declarative memory [i.e., explicit learning] may block (inhibit) the learning and/or retrieval of analogous knowledge in procedural memory [i.e., implicit learning]. The converse may hold as well" (p. 957) (also see competition hypothesis in Buffington & Morgan-Short, 2019, for relevant discussion). Hence, it is possible that if the role of WCF was more emphasized in this study, the results might be different.

In relation to the moderating effect the instructional context might have on the effectiveness of WCF, Erlam's (2005) finding on language aptitude and instruction effectiveness also suggests that under the more implicit instructional context (i.e., under

157

the implicit learning condition), learners' grammatical sensitivity becomes more facilitating to their learning (also see Stefanou & Révész, 2015). On the other hand, under the more explicit learning context where explicit instruction (i.e., form-focused instruction) is provided, the role of grammatical sensitivity in learning is minimized, i.e., learners with low or high grammaticality sensitivity levels equally benefit from explicit instruction. According to Erlam (2005), explicit instruction "minimize [s] or level [s] out any effect that individual differences in language aptitude may have with respect to instructional outcomes" (p.163).

Similar to Erlam's (2005) finding, in Li's (2013) investigation on the mediating effect of language analytic ability (i.e., grammatical sensitivity) on explicit and implicit feedback, the author also found that:

Other things being equal, language analytic ability is implicated in implicit conditions in the learning of easy, transparent structures that are within one's processing capacity, and in explicit conditions in the learning of hard, opaque structures where the internalization of available metalinguistic information sets heavy processing demands on internal cognitive resources (p.648).

Based on previous findings on grammatical sensitivity aptitude, it is highly probable that the seemingly high level of grammatical sensitivity of the control group learners may have helped them obtain more gains on their grammar usage, despite the lack of grammar feedback (i.e., the lack of explicit instruction) as previously speculated. In other words, it is probable that under the implicit learning context this current study provided, grammatical sensitivity of the control group learners may have compensated for their lack of grammar feedback helping these learners make some progress similar to those who received grammar feedback.

Working Memory Capacity and the Effectiveness of Written Corrective Feedback

The fourth and last research question (RQ4) of this study investigated whether WMC had any moderating effect on the effectiveness of WCF. The results of moderated multiple regression analyses revealed that WMC (i.e., complex working memory) did not significantly predict the extent to which learners benefited from WCF as hypothesized. That is, learners with high or low WMC did not benefit from WCF differently. Even though the OSpan scores of the WCF groups demonstrated to have a significant negative relationship with the UGJT delayed posttest scores, its interaction effect was not significant, indicating that OSpan did not moderate the relationship between WCF and UGJT scores. The negative correlation between OSpan and UGJT scores could be a result of other confounding factor(s) as the OSpan scores only predicted one out of eight sets of test scores. The current finding corroborates the result of Li and Roshan's (2019) study[50], which also found no significant association between complex working memory and the effect of direct WCF on learners' writing performance. Three main factors are posited to neutralize the facilitative effect of WMC on WCF within this study context: feedback type, instructional context and learners' proficiency level.

The first factor concerns the type of WCF used in this study. It is conjectured that the explicit nature and corrective force of direct WCF provided in this study may alleviate learners' feedback processing load operated through working memory leading to an absence of WMC moderating effect. It is as Li and Roshan (2019) explained in their study that given that direct WCF already helped locate learners' errors and provided them with corrections, learners were then alleviated from heavy cognitive processing load to notice and self-correct their errors (i.e., engaging in cognitive comparisons and hypothesis testing the errors), all of which involved working memory operations. Since learners' working memory resources were not heavily taxed on when processing direct WCF (i.e., explicit feedback), variations in WMC did not moderate the extent to which these learners benefited from the given WCF and, therefore, its moderating effect was not elicited. As Li and Roshan (2019) contend that "whether the [moderating] effect of working memory

---

[50] To the best of my knowledge, Li and Roshan's (2019) study is the first and only study to date that investigated the relationship between WMC, WCF and learners' writing performance. It is, thus, the only study that is comparable to the current study.

was evident had to do with the presence or absence of the demand for deep cognitive processing" (p. 10). This claim is in line with the generally accepted proposal of many cognitive psychologists indicating that cognitively demanding processing task is more likely to elicit working memory effect than the less demanding one (e.g., Conway, Cowan, Bunting, Therriault & Minkoff, 2002; Conway, Kane & Engle, 2003; Kane, Hambrick & Conway, 2005; Li, 2017; Linck et al., 2014; Ruiz, Rebuschat & Meurers, 2019).

In Goo's (2012) oral feedback study, a similar argument was put forward. Goo (2012) also found that individual differences in WMC did not mediate the effect of metalinguistic feedback (i.e., explicit feedback) but recasts (i.e., implicit feedback). The author also argued that because of the explicitness of metalinguistic feedback, learners did not need much effort to notice the feedback during interactions and their attention control mechanism in working memory was not heavily taxed on; therefore, the moderating effect of working memory was not evident. On the contrary, because of the implicit nature of recasts that often go unnoticed, learners needed to rely heavily on their attention control mechanism operated by working memory in order to notice and compare the given recasts with their previous output (i.e., engaging in cognitive comparisons); thus, the moderating effect of WMC was evident. In sum, the explicitness of direct WCF provided in this study seemed to help reduce learners' cognitive load alleviating their working memory operations. Since the processing of explicit type of feedback was not that cognitively demanding, this may in part explain why WMC did not moderate the effectiveness of WCF. As Li and Roshan (2019) assert, different types of feedback pose different levels of cognitive demands on learners' working memory. As a result, the moderating role of working memory may vary depending on the feedback type learners process.

Learners' instructional context is another confounding factor posited to neutralize the WMC effect. As previously illustrated, the instructional context of the current study was highly meaning-based (i.e., implicit learning) and, therefore, learners were directed to orient themselves more towards implicit learning. Accordingly, it is more likely that these learners may rely more on their implicit processing (i.e., learning without awareness) than explicit processing (i.e., learning with awareness) when processing feedback and

performing the tests. Since most scholars believe that implicit learning and processing do not significantly involve working memory operations, as working memory largely entails conscious operations implicated in explicit learning and processing (e.g., N. Ellis, 2005; R. Ellis, 2009b; Skehan, 2016, Tagarelli, Mota & Rebuschat, 2015), the implicit instructional context provided may reduce or inhibit learners' reliance on their explicit conscious learning (see Ullman, 2015, 2016 for competition hypothesis) which in turn reduces the extent to which these learners rely on their working memory resources. As learners' working memory might not be fully activated within the implicit learning condition as explicated, variations in WMC might not play a crucial role within this study context as well. Thus, the implicit instructional context might be another factor neutralizing the WMC effect within this study context.

The null WMC effect found in this study is in line with the results of previous studies, which also detected no WMC effect in implicit learning conditions (e.g., Ando, Fukunaga & Kurahachi, 1992; Hamrick, 2015; McDonough & Trofimovich, 2016; Ruiz et al., 2019; Tagarelli, Mota, & Rebuschat, 2015; Walker, Monaghan, Schoetensack, Rebuschat, 2020). In addition, the finding lends support to the claim that learners' instructional context mediates the WMC effect on learners' L2 performance (e.g., Ando et al., 1992; Faretta-Stutenberg & Morgan-Short, 2018; Ruiz et al., 2019; Suzuki & DeKeyser, 2017; Tagarelli et al., 2015; Walker et al., 2020; Yang, Shintani, Li, & Zhang 2017).

The last factor postulated to mediate the effect of WMC within this study context is learners' proficiency level. It is already well established in the field that learners' proficiency level influences the facilitative effect of WMC on learners' L2 learning outcomes (Adams & Mohammadtaghi, 2014; Alptekin & Erçetin, 2010; Sagarra, 2012, 2017; Serafini & Sanz, 2016). That is, learners at the beginning stage of development tend to rely more on their WMC (involving control processing) than those with more exposure to language input or at a more advanced level (leaning towards automatic processing) (Harrington & Sawyer, 1992; Mitchell, Jarvis, O'Malley, & Konstantinova, 2015; Roberts, 2012; Sagarra, 2012, 2017; Serafini & Sanz, 2016). In Serafini and Sanz's (2016) study, for example, complex working memory did not have any effect on learners'

morphosyntactic development within the advanced[51] but beginner groups. Other studies (e.g., Gilabert & Muñoz, 2010; Grey, Cox, Serafini & Sanz, 2015) also did not find complex working memory effect on intermediate learners' performance. Similar to those studies, as learners in this study were at the intermediate level, it is probable that they might rely on other cognitive resources or system (e.g., procedural memory) when processing feedback and when performing the tests instead of merely relying on their working memory resources. This, thus, explains why variations in WMC did not moderate the extent to which these learners benefited from WCF. Robinson (2002, 2005, 2012) also contends that, as adult learners progress in their development, they tend to rely more on other cognitive resources when performing language-related tasks.

Taken together, it can be concluded that learner differences in WMC did not moderate the extent to which the learners in this study benefited from the given WCF. It is posited that the explicit nature of the provided feedback, implicit instructional context and learners' current proficiency level are key factors mediating working memory effect.

---

[51] "Advanced learners" in Serafini and Sanz's (2016) study were described as learners who enrolled in the advanced Spanish course (6th semester) and they were equivalent to "intermediate learners" in this study.

4.4.2. Conclusion

The purpose of the current study was twofold: 1) to investigate the effects of direct *"focused"* and *"unfocused"* WCF and their differential effects on the development of L2 learners' explicit and implicit knowledge of English plurals, and 2) to examine the relationship between learners' WMC and the effectiveness of WCF. Regarding the effectiveness of WCF on learners' L2 development, the results of inferential statistics suggest that both direct *"focused"* and *"unfocused"* WCF were effective in helping learners develop their L2 explicit and implicit knowledge of English plurals. Nonetheless, there were no significant differences in the effects direct *"focused," "unfocused"* WCF, and content feedback had on learners' improved performance. In other words, direct *"focused" and "unfocused"* WCF provided for the WCF treatment groups, and content feedback provided for the control group were equally effective in assisting learners develop their English plural knowledge. The results provide further empirical evidence in support of the effectiveness of WCF on learners' L2 acquisition. Regarding the relationship between WMC and the effect of WCF, the results revealed that learners' variations in WMC did not influence the extent to which the learners in this study benefited from feedback. The finding sheds more light on the role WMC plays in different instructional contexts and how different types of feedback and different levels of L2 proficiency could potentially mediate its effect.

Notwithstanding, two limitations of this part of the study should be acknowledged and taken into consideration before generalizing the above findings. One limitation involves the study's small sample size due to a high attrition rate (45%, from 130 to 75 learners), leading to insufficient statistical power[52] to detect all the existing effects. Field (2018) points out that, the sample size is generally perceived to significantly associate with the ability of the statistical test to detect the effect and the larger sample size tends to have more statistical power to detect the effect than the smaller one (Field, 2018). In addition, the sample size and the statistical power can also influence the significance of the test statistics (i.e., *p* values) (Field, 2018). That is:

the sample size affects whether a difference between samples is deemed

---

[52] Statistical power refers to "the ability of a test to find an effect" (Field, 2018, p. 84).

significant or not. In large samples, small differences can be significant and in small samples large differences can be non-significant…even a difference of practically zero can be deemed significant if the sample size is big enough… [this is because] bigger samples have less noise, so even a tiny signal can be detected. (Field, 2018, p. 90)

In other words, significant effect can be easily detected within a large sample size study. Meanwhile, significant effect might not be detected, if at all, in a small sample size study. Taken together, given that the current study had a small sample size population, and this, to some extent, lowered the statistical power of the employed statistical tests, it is, thus, possible that the underpowered statistical tests may not be able to detect all the existing true effects that occurred in this study. Accordingly, the current findings should be interpreted and exercised with caution.

Another limitation is the study's short intervention period which may attribute to small differences between the participating groups. The small differences could potentially make it more difficult, especially for the underpowered statistical tests to detect all the existing effects, even if there was any. Indeed, given that the learners in this study were at intermediate level and already had some established plural knowledge, the 6-week treatment period can be viewed as minimal or even insufficient for learners at this level to make drastic progress compared to beginners who have much room to grow. Further, the fact that these learners already had plural instruction since primary school but still have difficulty acquiring this linguistic feature suggests that intensive focus on this feature is needed. Even though six weeks may be adequate to make remarkable progress for L2 learners from other L1 backgrounds, for this group of learners whose L1 (Thai) does not have plural inflection morphemes, a longer treatment period, i.e., more feedback treatment episodes, may be more appropriate as the effect of WCF might be more evident and this could potentially lead to greater learning outcomes allowing the difference between groups to be more discernable.

Taken together, future replication of this study should attempt to include more participants to increase the statistical power of the statistical tests. Power analysis is needed to be carried out, especially before longitudinal research is conducted to avoid an unexpected loss of statistical power caused by high attrition. Moreover, longer

intervention periods should be included in the future design to discover whether more WCF episodes could actually effectively help learners at this level achieve greater learning outcomes.

Chapter V
QUALITATIVE RESULTS

This chapter presents the results of the qualitative data analysis. It first introduces the qualitative data analysis method followed by the results. The chapter ends with the discussion and conclusion of the findings.

5.1. Method of qualitative data analysis

The qualitative analysis part of this study aimed to complement the main quantitative findings presented in chapter 4. The qualitative analysis was conducted to examine if there were other individual differences or contextual factors that might have influenced 1) learners' responses to WCF on their writing, and 2) learners' abilities to benefit from WCF, apart from the predetermined variables.

In this qualitative data analysis section, qualitative interview and exit questionnaire data were analyzed. Regarding the qualitative interview, the first step of the analysis involved transcribing and translating the interview from learners' L1 (Thai) to English by the researcher. Twenty-five percent of all translated interviews were reevaluated by another experienced Thai native English university instructor to confirm the reliability of the translation. Any discrepancies in translation arising between the researcher and the other instructor were resolved through discussion. After completing all transcribing and translating processes, the interview data were then coded and assigned into categories (see section 5.2.). With regard to the exit questionnaire data, the data were arranged in the same vein: translating (where necessary) and then categorizing.

5.2. Qualitative results

5.2.1. Qualitative interviews

To analyze the qualitative interviews, the initial stage was to transcribe the interviews. Overall, 18 interview records from six learners were transcribed (two learners were randomly selected from each participating group). Among these six learners, one learner did not complete the delayed posttest, and, therefore, her test data could not be used for quantitative analyses. Nevertheless, her interview records were included in this qualitative data analysis because this learner had completed all required interview sessions and her interview information contributed to the deeper understanding of learner factors that might affect the current research findings. Thus, all 18 interview records were transcribed as planned.

After completing the transcripts, the transcripts were then translated from Thai to English by the researcher. An experienced Thai native English university instructor rechecked 25 percent of the translation to ensure the reliability of this version of the interview. Any discrepancies in translation between the researcher and the Thai instructor were resolved through discussion until they reached a full agreement. After establishing the reliability of the translation, the interview data were coded and assigned into categories. Categories were not predetermined. They were assigned based on factual information obtained from the interviews. Based on the interview data, six main emergent categories were formed: Feedback Noticing, Feedback Application, Motivation, Feedback Usefulness, Self-Rating Progress, and Writing Concerns. Table 5.1. exhibits all the categories exemplified by some extracts from the interviews.

Table 5.1

Categories of the qualitative interview data

Categories	Examples
1) Feedback Noticing	"I tried to look at the errors I might have made through the feedback you provided because I am not confident about my own grammar knowledge, like how to correctly use the pronouns. I usually misuse the pronouns or do not use them at all. Some errors I never realized they were errors because I was so used to them, but when I got feedback and had a chance to check them out, I then realized that they were really errors." (Jackson, pseudonym)
2) Feedback Application	"I did recheck if my writing covered all the ideas I wanted to write about. Like from the feedback I received last time that I separated one main idea into two separate main ideas, so this time I tried to make sure that that kind of thing didn't happen again." (Ross, pseudonym)
3) Motivation	"I still hit the panic mode when I could not think of anything to write about or I usually realized later that I should have added particular information when I already finished the whole thing. I need to work harder." (Amelia, pseudonym).
	"It was fun thinking about ways to deliver the message out of my mind" and "I think it is an opportunity for us to bring out what we have in our head, conveying our thoughts through writing." (Monica, pseudonym)

4) Feedback Usefulness	"I prefer feedback because without it, I don't know whether I have made some errors or not, and, as a result, I will continue making the same errors. As I mentioned earlier, I can't judge my own grammar." (Tony, pseudonym)
5) Self-Rating Progress	"Better in the way that my writing is not too broad anymore. It became more precise like more focused." (Amelia, pseudonym) "I wrote better because I believe I produced much clearer content." (Meredith, pseudonym)
6) Writing Concerns	"Like for some topics, I don't have any background knowledge about them at all and that can be difficult to write about, given the topic that we wrote about smoking lounges should be banned from airports for example. It is the most difficult topic for me because I don't know anything about this topic. I am also concerned about some vocabularies. I know how to call things in Thai but not all in English." (Jackson, pseudonym)

The following section presents the findings of the qualitative interviews.

5.2.1.1. Feedback noticing

Learners' interview data were assigned to this category if the data showed that learners noticed the feedback they obtained at the level of understanding (i.e., comprehending the significance of noticed objects and attempting to solve problems based on what has been observed [Schmidt, 1990]). In this context, Feedback Noticing was, thus, evidenced when 1) learners were able to specify the feedback they received (i.e., they reported whether it was a pronoun/tense error or content feedback), and 2) they knew the cause of the error or were able to elaborate on how to correct it (e.g., Jackson, from the unfocused WCF group: ***"I usually misuse the pronouns or do not use them at all***.", Ross, from the control group:

> ***"I knew that I didn't do my best on that. My thought was all over the place****. **My thought was not arranged in systematic manner, like I talked about something here and suddenly switched to talk about another thing there… I have learnt that I should finish one main idea with all the supporting details before moving on to the next main idea."***

Based on this criterion, four out of six learners showed evidence of Feedback Noticing during the first interview session (2nd week of the treatment phase). For example, Amelia, from the focused WCF group, reported that:

> ***"The first thing I did was to check the feedback and my errors. I looked at them to check whether I did make the same errors again and whether I had improved from last time. From my recent feedback, I could see my progress. Only that some types of grammatical errors like plural errors just kept reappearing. I think I need more time to write. Honestly, I didn't have time to recheck and correct my errors."***

However, two learners, Monica from the control group and Tony from the focused WCF group, despite acknowledging the given feedback, could not give specific details about the feedback they received. Tony, for instance, gave a vague statement about the feedback he received: ***"I looked at the feedback. Honestly, sometimes I knew that I used it wrong, but I was in a hurry."***

170

Even though a lack of feedback specification in Tony's case as exemplified above, did not necessarily mean he was completely unaware of the feedback, this might indicate that Tony only paid peripheral attention to the feedback he received and, therefore, his noticing may not be at the level of understanding when compared to others.

During the second interview session (4th week of the treatment phase, two weeks after the first interview session), five learners including Monica showed evidence of Feedback Noticing.

In her second interview, Monica from the control group reported that:

"***This time I got the feedback about how I should give examples to support my main idea. In my recent work, I just wrote about what to do in general, but I didn't give specific details about how to do it. I admit that I should have provided more specific examples or evidence to support my claims. Also, you had pointed out that my first and second reasons were actually the same reason. That was because I didn't plan in advance. Just kept writing and this was the result,***"

whereas in her first interview, Monica just gave a short answer regarding feedback: "***I just looked at the feedback and my errors.***"

While Monica demonstrated evidence of Feedback Noticing in her second interview, Tony, on the other hand, was confused about the feedback he received (he did not know the cause of his informed errors), and, therefore, his account was not considered as evidence of Feedback Noticing.

Tony: *I received feedback about subject verb agreement…that I forgot to add an "s" or about tenses.*

Researcher: *Was it plural error?* (Tony was from the focused WCF group which only received corrections on plural errors)

Tony: *I don't know…there were many "s" in my feedback.*

Researcher: Did you carefully check them?

Tony: *Maybe not that much.*

Based on the above results, it showed that most or nearly all learners noticed the given feedback at the level of understanding as they recognized the feedback and were aware of how to improve their work based on the feedback they received. This, however,

was not unexpected given that they were intermediate learners, who might already have some established grammatical knowledge (i.e., metalinguistic knowledge) to help them comprehend the cause of their informed errors. Nonetheless, Tony from the focused WCF group was an exception, since he "noticed" the feedback he received yet did not recognize it as feedback on plural errors. His lack of noticing at the level of understanding (e.g., could not identify if his errors were subject-verb agreement, tense or plural errors) might be a result of his inadequate attention paid to the feedback he received, his lack of interest or motivation towards the feedback (he mentioned he was in a hurry), insufficient time to articulate the feedback, his limited explicit grammar knowledge, or the input flooding effect (too much information that overwhelmed his limited cognitive ability). These said factors were also speculated to have influence on how learners could successfully notice the target structure at the level of understanding.

Another point worth mentioning was the fact that WCF, when providing consistently, seemed to help promote "noticing" at the level of understanding among feedback receivers. This claim was supported by the above finding which exhibited that one learner, Monica from the control group, who did not show evidence of Feedback Noticing in her first interview (two weeks after receiving feedback), demonstrated evidence of Feedback Noticing later on (four weeks after receiving feedback). Nonetheless, this claim was based on the observation of one learner; further evidence is needed in order to gain fuller understanding of this association.

5.2.1.2. Feedback application

Concerning Feedback Application, this category involved learners reported using feedback to improve their work or to resolve their errors (e.g., Ross from the control group: "***Like from the feedback I received last time*** that I separated one main idea into two separate main ideas***, so this time I tried to make sure that that kind of thing didn't happen again***."). This also included learners' comments on reexamining certain writing aspects or grammatical structures relating to the feedback they received (e.g., Amelia from the focused WCF group: ***"I tried to recheck my grammar more often. I tried to tell myself not to forget to add an "s"***).

Based on this criterion, the results of the first interview session showed that only Ross from the control group, reported clear evidence of Feedback Application:

> "***I knew that I didn't do my best on that. My thought was all over the place***. ***My thought was not arranged in systematic manner, like I talked about something else here and suddenly switched to talk about another thing there...So I tried to be more focused, paying attention to just one topic/reason each time. I have learnt that I should finish one main idea with all the supporting details before moving on to the next main idea.***"

Other learners, however, showed vague comments on Feedback Application. Jackson from the unfocused WCF group, for instance, just stated that "***Yes, I did use some of the feedback*** in my new piece of writing but still made the same errors", or Tony from the focused WCF group, who simply said *"**I used it sometimes.**"*

For the second interview session, three learners showed evidence of Feedback Application. One was Ross from the control group who also reported evidence of Feedback Application in his first interview. The other two were Monica from the control group and Amelia from the focused WCF group, who both gave a clear statement of how they thought about the feedback they received, and how they used the feedback to improve their writing. Three other learners, Jackson and Meredith from the unfocused WCF group and Tony from the focused WCF group, showed ambiguous evidence of Feedback Application.

The results of the Feedback Application above, thus, suggested that even though WCF increased chances of noticing among feedback receivers as previously mentioned in the Feedback Noticing section, not all that were noticed would be internalized and become an uptake, assisting learners to properly correct their informed errors. This was based on the fact that most learners in this study noticed the given feedback, yet only half of them showed evidence of Feedback Application. Learners' lack of interest or motivation to improve their work may account for lack of Feedback Application. It should also be highlighted here that successful Feedback Application or uptake (i.e., correcting errors or making fewer errors as a result of feedback) has been regarded among scholars (e.g., Afitska, 2015; Beuningen, De Jong & Kuiken, 2012; R. Ellis & Sheen, 2006; Lightbown, 1998; Loewen, 2004, 2005; Lyster, 1998; McDonough, 2005; Sheen, 2004) as a key predictive indicator of learners' successful acquisition of the target linguistic features. Consequently, a lack of Feedback Application or uptake might indicate otherwise.

All in all, noticing is inevitably associated with the effective use of feedback or feedback uptake, given Monica's case, for instance. In her case, Monica from the control group did not report evidence of either Feedback Noticing or Feedback Application in her first interview session, but she did report both in her second interview session. This implicates that noticing has to occur first before the efficient use of feedback can take place. The discussion and conclusion section of this chapter provides an extended discussion regarding this issue.

Learners' accounts were assigned to this category if the learners showed their effort to learn by applying what they had observed for their better learning outcomes or demonstrated their intention to improve themselves. For example, Amelia from the focused WCF group said that "I still hit the panic mode when I could not think of anything to write about or I usually realized later that I should have added particular information when I already finished the whole thing. ***I need to work harder***."

The above criteria were developed based on the ground that motivated language learners should, in general, have positive attitudes towards their language learning (Dörnyei, 1990, 2003, 2005; Gardner, 1985, 2000, 2001); thus, making an effort to work harder on their language training (Dörnyei, 2010; Dörnyei & Ushioda, 2009, 2011; Gardner, 1985, 2000, 2001; MacIntyre, Baker, Clement & Conrod, 2001). Following this guideline, learners who demonstrated their willingness to learn or improve their writing during the time of the study were henceforth regarded as motivated language learners.

Based on the above criteria, Ross from the control group was regarded as a motivated learner in this study since he demonstrated his strong motivation towards his writing development. According to Ross, "***I just need to work harder if I want to improve***…***Honestly, I want to improve every aspect of my writing because writing is actually important for my field of study***." In all interviews, Ross always stated that he needed to work harder and he preferred to have lots of feedback because he believed that ***"It is good to receive feedback regularly, so we know when we make any mistakes. Sometimes we do not notice our own mistakes. Others may be able to see our work (mistakes) more thoroughly."***

Monica from the control group also showed her unwavering motivation to learn during the time of the study. She demonstrated what she had learned from the given feedback and attempted to use it to improve her work from time to time. Her views towards writing practice and feedback were:

> ***"It was fun thinking about ways to deliver the message out of my mind"*** **and "*I think it is an opportunity for us to bring out what we have in our head, conveying our thoughts through writing. If you give us phrases to remember, we***

sure can remember but it will be short lived. Practicing is different. But I don't think many students would think like I do. They are probably lazy to do a lot of work and they would blame teachers for giving them a lot of work."

Apart from Ross and Monica, Amelia from the focused WCF group also exhibited strong motivation towards her writing development during treatment.

While the above accounts showed strong motivation towards learning, Jackson, Meredith and Tony's reports were indistinct, given Meredith's report, for instance. Even though Meredith from the unfocused WCF group stated that she wanted to improve her writing content, she did not explain what she would do in particular to improve her work or express her intention to work harder. In fact, in her last interview, Meredith did mention that she was absent minded sometimes when she wrote and that also occurred while she was taking her posttest. Similar to Meredith, Jackson's motivation was ambiguous, and it was speculated that even if he was motivated, his level of motivation might fluctuate a lot during participation. For example, Jackson stated earlier in his second interview that "*I have tried my best to make fewer errors. I try to be more careful not to make errors.*" However, when he was asked if he carefully looked at the feedback he received, his reply was, "*I didn't look at it thoroughly today because I felt dizzy and sleepy. I nearly fell asleep before your session.*" Also, in his last interview, when being asked if he always checked the feedback, he said, "*Sometimes I didn't have a chance to look through all the feedback I received, but most of the time I did.*"

Nonetheless, unlike Meredith and Jackson from the unfocused WCF group, Tony's motivation seemed to be quite low or even non-existent especially during participation.

> "*If it is not for the test, I might not pay full attention to or focus on my study*. *It is like if we haven't had midterm exams just yet, we don't have to study hard, but when we are about to have the exams, we have to read so that we can do well on the tests*" and "*I made some improvements but not much as I didn't really go back and think about what I could have done better,*" said Tony.

From the above accounts, three learners seemed to be motivated towards learning because of the given feedback and they showed their strong intention to improve accordingly. In contrast, the rest showed unclear messages towards improvement and,

therefore, indicating a low level of motivation or a lack thereof. However, an interesting observation here would be that the three learners who were motivated towards learning and self-improvement were the same learners who reported evidence of Feedback Application. Meanwhile, the other three learners who lacked evidence of Motivation also lacked evidence of Feedback Application. Thus, the findings implicated that a low motivation level or a lack thereof might potentially affect learners' abilities to "uptake" or apply new input/knowledge (corrections) to improve their work. The findings also highlighted the significance of motivation as a key moderator influencing the extent to which learners apply what they learn or acquire into practice.

5.2.1.4. Feedback usefulness

Learners' comments were coded into this category when they stated the usefulness of feedback towards any of their development, for example, Ross from the control group:

> "***Without feedback, I might think I already did everything right when it was not. Feedback helps me realize my mistakes or aspects that I need to improve so that I can improve***.", Jackson from the unfocused WCF group: "***I become more aware of my grammatical errors, more careful about them. Though I still can't make it 100% correct, I have tried my best to make fewer errors.***"

Based on this criterion, all learners reported on feedback usefulness in all their interview sessions, including the last interview after the posttest. This also included Jackson and Meredith from the unfocused WCF group and Tony from the focused WCF group, who had not shown any evidence of Feedback Application but viewed feedback as a useful tool. Jackson, for instance, commented that because of the feedback, he was more careful with his grammar and he preferred to receive teacher's corrections since it helped him to remember his errors better. Tony also asserted that he too preferred feedback because he believed that "***without it, we don't know whether what we do is right or wrong and because of that, it is more likely that we will continue making an error.***" In line with Tony's comment, Meredith admitted that receiving feedback was better "***Because we can't be sure. Sometimes we might think we have done it right but actually it is not.***"

In conclusion, all learners viewed feedback as a useful pedagogical tool that can help them recognize their own errors or writing aspects they needed to improve. Even though some learners did not show evidence of Feedback Application, all learners were well-aware of how feedback can actually help them and preferred to receive it.

5.2.1.5. Self-rating progress

Learners' comments regarding their writing progress during the time of the study were assigned to this category. Table 5.2. displays each learner's self-rating progress report.

Table 5.2

Self-Rating progress reports

Names	Progress
Control Group (content feedback)	
1) Ross	**Content** "My thinking skill. I feel like I think more systematically. In the past, I just wrote whatever came to my mind, never planned in advance, so my thought was all over the place. Even for me, when I read my own work, I get confused sometimes."
2) Monica	**Content** "I could write faster. I realized that I could write better under time limit. It is quite challenging trying to convey what is in my mind out in words and to make sure that the selected words represent the exact message I want to convey."

Unfocused WCF Group

(feedback on multiple grammatical types of errors)

1) Jackson

Grammar "I am more careful with grammar. It has been improved a little bit."

Content "I am more careful about my arguments. I realize I should have strong evidence to support my claims, instead of saying something is good or bad without any strong evidence to support my claims…Maybe my general knowledge has been increased a little bit too."

Vocabulary "When we practiced writing, I had a chance to learn new vocabulary I never ever used before."

2) Meredith

Content "I am better at organizing my ideas."

Grammar (no further explanation)

Focused WCF Group

(feedback on plural errors only)

1) Amelia

Content "Better in the way that my writing is not too broad anymore. It becomes more precise like more focused."

2) Tony

Grammar "Possibly grammar. I usually get confused because of grammar. It is because sometimes you can use more than one tense to express your ideas. So I think writing practice helps me on this."

Content "It helps me think in English faster because usually it takes time for me to think in English. It is like we already have something ready to be delivered out of our heads."

Based on the results above, all learners believed they had made improvement in their content development. Ross and Monica from the control group, who received only content feedback, commented that their written content had been improved as a result of writing practice and the feedback they received. However, all the others from the focused and unfocused WCF groups (except Amelia from the focused WCF group), who received only grammar feedback, believed they had improved both their content and grammar knowledge. Two explanations for these results would be: 1) writing practice in general helped with content development, regardless of the feedback type learners received (whether it was content or grammar feedback), and 2) the progress on written content was more evident than the grammatical one, suggesting that the improvement of grammar knowledge might take a longer time to develop than it does to the written content.

5.2.1.6. Writing concerns

Learners' writing concerns and the writing aspects they wanted to improve were assigned to this category. Writing concerns involved learners' concerns on particular writing aspects they believed they were still not good at and these writing aspects made them worried when writing. Consequently, learners may attempt to monitor or recheck these writing aspects more often than others. On the contrary, their needs for writing improvement involved the aspects of writing the learners believed they should improve (i.e., they wished they could improve these writing aspects) to make their overall writing better (i.e., they believed these writing aspects attributed to good essay writing). Table 5.3. illustrates each learner's writing concerns and his/her needs for writing improvement.

Table 5.3

Writing concern reports

Names	Concerns	Need for Improvement
Control Group (content feedback)		
1) Ross	**Vocabulary** "I was afraid I couldn't think of the right vocabulary to use in my writing." **Grammar (Tense)** "I was afraid I would use wrong grammar."	**Content** "Ideas are important and thereby we should focus more on organizing our ideas." **Grammar (Tense)**
2) Monica	**Content** "I was afraid I would write off- topic. I wanted to have strong reasons to support my argument…I like to think about how I could deliver my message differently. I meant keeping the meaning but using more formal vocabulary to elaborate for example. I just want my writing to be better."	**Content** "I would like to be more argumentative like I can strongly support my claims." **Grammar** "I want to improve my grammar too. My grammar is still not perfect."
Unfocused WCF Group (feedback on multiple grammatical types of errors)		
1) Jackson	**Content** "Like for some topics, I don't have any background knowledge about them at all and that can be difficult to write about, given the topic that we wrote about smoking lounges should be banned from airports for example. It	**Vocabulary** "Word choice. I think mine is not that good just yet, kind of strange and redundant. I want it to be more precise. I think I don't have much chance to practice."

is the most difficult topic for me because I don't know anything about this topic."

Vocabulary "I was also concerned about some vocabularies. I know how to call things in Thai but not all in English."

2) Meredith	**Content** "I was concerned whether I would have any ideas or strong reasons to support my argument. I was afraid I might not have any interesting ideas or strong arguments."	**Vocabulary** "Language use because Thai learners usually think or draft their writing in Thai first." **Content** "How to write a good content essay. Most people like to put lots of irrelevant information in their writing making their arguments either confusing or weak. I want my writing to be direct and concise."

Focused WCF Group
(feedback on plural errors only)

1) Amelia	**Content** "I was concerned about the topic and worried if I would have enough background knowledge about the topic…like when I had to write about immigrants and smoking ban, these are topics I don't really have much knowledge about so it was difficult to find reasons to support my arguments."	**Content** "creativity" **Grammar (Tense)** "I think most learners are still confused about grammar aspect. I don't like that we learn grammar through random sentences from textbooks because I am not sure if we can really use that in our daily life. I think I want real life examples."

2) Tony	**Grammar (Sentence Structure)**	**Content** "How to develop ideas"
	Tony: "*I am concerned about my sentence structure*"	
	Researcher: "What do you mean by sentence structure? Please elaborate."	
	Tony: "*Like grammar.* Sometimes I don't know how to use it correctly so I have to cut it off from my writing because if I write it wrong, that part will look strange. I usually try to find a different way to create the sentence instead then."	

Based on the results above, most learners were concerned about their written content. Some were worried about their choice of vocabulary and grammar usage. Interestingly, while learners from the control and focused WCF groups wanted to improve both their content and grammar aspects, learners from the unfocused WCF group seemed to pay their exclusive attention to their content and vocabulary development.

Before analyzing the exit questionnaire data, all 75 exit questionnaires were first translated from Thai to English where necessary (most questionnaires were already written in English and did not require translation). After the translation, all questionnaire data were categorized.

Based on the questionnaire data, three main categories were established: Self-Rating Progress, Writing Concerns and Research Goal. In line with the categories of the qualitative interviews, all learners' writing-related concerns and writing aspects they wanted to improve were assigned to the Writing Concern category, whereas learners' writing progress reports were assigned to the Self-Rating Progress category. With regard to the Research Goal, this category was further subdivided into a "Yes" and "No" subcategories. If the learners' replies indicated accurate acknowledgement of the research goal (i.e., to learn from feedback and use feedback to improve their work), their replies were assigned to the "Yes" subcategory, otherwise they were assigned to the "No" (including the ones commenting that the goal was to improve learners' L2 writing skill).

5.2.2.1. Self-rating progress

As mentioned earlier, learners' writing progress reports were assigned to this category. Based on the exit questionnaire data collected from 75 learners, 73 of them believed that they had made some progress in their writing abilities after participating in this study, while two learners reported otherwise. Of these two learners, one from the unfocused WCF group was uncertain about his progress. Meanwhile, the other from the focused WCF group believed she did not make any progress at all, stating that it was because she was only allowed to practice writing once a week and for less than an hour each time (reviewing feedback for 5 minutes and writing for 30 minutes).

Despite these two learners' doubtfulness on their progress, the majority of the learners believed they had made some progress in their writing. Of these 73 learners, 43%

reported that they had made progress in their writing in general[53] (i.e., other writing aspects apart from content, grammar and vocabulary), followed by written content and content organization (35%), grammar (16%) and vocabulary (6%), respectively.

5.2.2.2. Writing concerns

Based on the 75 exit questionnaires, learners from all groups were most concerned about their content development (37%), followed by their correct use of English grammar (30%), vocabulary (21%) and other writing aspects (12%) successively. These learners also indicated that the writing aspects they aimed to improve the most included: written content (40.5%), followed by grammar (36%), vocabulary (15%) and writing in general (8.5%), all of which was in line with previous interview results.

When analyzing each group's data separately, the results were also in accordance with the overall findings presented above. The only exception was that of the focused WCF group learners who showed their most concern over content development (similar to the unfocused WCF and control groups) yet expressed the desire to improve their correct use of English grammar the most. Table 5.4. below summarizes each group's writing concerns and writing aspects they aimed to improve during the time of the study.

---

[53] This category included, for example, the ability to write faster within the time limit, spelling and punctuation usage.

Table 5.4

All groups' writing concerns and goals for improvement

Names	Concerns (out of 100%)	Need for Improvement (out of 100%)
Control Group	Content development: 38%	Content Development: 45%
	Grammar: 33%	Grammar: 37%
	Vocabulary: 18%	Vocabulary: 16%
	Others: 11%	Others: 2%
Unfocused WCF Group	Content Development: 36%	Content Development: 41%
	Grammar: 32%	Grammar: 27%
	Vocabulary: 19%	Vocabulary: 16%
	Others: 13%	Others: 16%
Focused WCF Group	Content Development: 37%	Content Development: 36%
	Grammar: 25.5%	Grammar: 43%
	Vocabulary: 25.5%	Vocabulary: 14%
	Others: 12%	Others: 7%

The results illustrated above suggested that most learners seemed to extend themselves more towards "meaning", that is, focusing on the development of "meaning" rather than "forms." Even though the focused WCF group exhibited their highest interest in improving their grammar usage (43%), when viewing their overall results (considering content + vocabulary[54] development = 50%), it was clear that they still paid more attention to their "meaning" rather than "forms." Additionally, even though the control group placed most of their worries and attention on content development as anticipated (since they only received content feedback), the control group learners were still tremendously concerned about their grammar usage (33%) (despite not receiving any grammar feedback). This remark is further discussed in the next chapter.

5.2.2.3. Research goal

The last category identified from the exit questionnaire data was the Research Goal. In this category, learners' opinions towards the goal of the study were divided into a "Yes" and "No" subcategories. If the learners' replies indicated accurate recognition of the research goal (i.e., to learn from feedback and use feedback to improve their work), their replies were assigned to the "Yes" subcategory, otherwise they were assigned to the "No" (including comments claiming that the goal was to improve learners' L2 writing or English skills in general).

Based on the exit questionnaire data, only three learners' comments were assigned to the Yes subcategory. These three comments clearly demonstrated that the learners accurately recognized the goal of this study[55], which was to evaluate whether learners were able to employ the feedback they received to improve their work or not.

---

[54] In this study, vocabulary was considered an aspect contributing to meaning-based learning. In other words, when learners placed their focuses on vocabulary development, it simply meant their focuses were placed on "meaning" instead of "forms."

[55] Among these three learners, two were from the control group and one from the unfocused WCF group. These learners stated that the true purpose of the study was to explore the extent to which the learners can improve their writing as a result of feedback. However, none of them acknowledged that the target linguistic structure of the study was English plurals.

In contrast, 72 comments were assigned to the No subcategory since their responses did not reflect accurate acknowledgement of the goal of this study. These included responses claiming that the study was designed to help improve learners' English writing skill, to assess learners' current English writing abilities, to examine if writing can help enhance other English language skills or to investigate learners' short-term memory, for instance. Nevertheless, it is worth noting here that none of the 75 learners, including the three learners in the Yes subcategory, mentioned English plurals as the target linguistic feature of this study.

Taken together the results above suggested that despite the explicit nature of WCF that inevitably induces noticing, its explicitness could be lessen when the feedback was introduced in a more "meaning" focused task where the feedback role was not emphasized. In this case, the effectiveness of WCF might be mitigated, limiting its capacity to help learners improve their knowledge. However, this might not always be the case especially for those motivated learners.

In conclusion, based on the results of the questionnaire data that only three learners were able to recognize the true purpose of the study while the other 72 did not, despite two-thirds of them received extensive grammar feedback every week, it can only be assumed that 1) most learners did not pay much attention to the feedback they received, 2) they did not think it was essential to attend to the provided feedback earnestly, and, therefore, 3) it was plausible that most of them might not even try to apply the received feedback to their work as hypothesized. Providing that the learners showed little attention towards the feedback they received (less focused on "forms"), but how to write a good content argumentative essay (focused more on "meaning"), this after all might influence the overall results of the present study to a greater extent. The next section offers a more in-depth discussion regarding these qualitative results.

5.3. Discussion and conclusion

As stated earlier, the qualitative analysis part of this study was conducted in order to explore other internal and/or external factors that could have influenced learners' responses to the provided pedagogical treatment. This investigation was aimed to serve as a complement to the main quantitative findings (chapter 4). Qualitative interviews and exit questionnaires were the primary measures employed for the qualitative data collection. The collected data were then coded and assigned into six main overarching categories, namely Feedback Noticing, Feedback Application, Motivation, Feedback Usefulness, Self-Rating Progress, and Writing Concerns, as previously illustrated.

In what follows, the discussion and conclusion of the results are offered.

5.3.1. Discussion
5.3.1.1. Feedback noticing

Previous interview data on Feedback Noticing demonstrated that most learners showed evidence of Feedback Noticing at the level of understanding throughout the time of the study. This finding, thus, corroborates previous claim stating that, while serving as an input providing tool, WCF helps inducing and promoting "noticing" (Bitchener & Storch, 2016; Polio, 2012; Sheen & R. Ellis, 2011; Van Beuningen, 2010, Williams, 2012), the cognitive condition significant for successful second language acquisition (Gass & Mackey, 2015; Long, 1996, 2007, 2014; Schmidt, 1990, 1995, 2001; Swain, 2005).

In the SLA field, it is well accepted that noticing is the crucial cognitive condition underlying acquisition. According to Schmidt (1990, 1995, 2001), there are two levels of noticing: noticing without metalinguistic understanding (henceforth, "noticing at the level of attention") and noticing with metalinguistic understanding (henceforth, "noticing at the level of understanding"). While "noticing at the attention level" is viewed by many scholars as a requisite for learning, "noticing at the level of understanding," which is when learners recognize the mismatch between their current knowledge and the target knowledge they aim to accomplish, as well as realize what needs to be done to acquire the

target-like knowledge, is facilitative to successful L2 acquisition. As Shintani and R. Ellis (2013) point out, "what seems to have been important is not the 'noticing' of errors…but their 'understanding' of the rule and their ability to use this to correct their errors" (p. 300).

Hence, it is crucial to discover that the WCF provided in this study not only helped induce "noticing at the attention level" (as most noticed the given WCF and acknowledged that they had made some errors) but also "noticing at the understanding level." This indicates that providing WCF to L2 learners is beneficial. The frequency and saliency of the target linguistic features that the WCF provides help learners notice the gaps in their learning and thereby uncover the underlying causes of the errors faster; consequently, accelerating their acquisition processes.

Even though the level of noticing may vary among learners as shown in this study, as a result of learners' disparate background knowledge, as well as individual differences, it is indisputable that WCF is an effective consciousness-raising tool (Bitchener, 2012; Bitchener & Storch, 2016; Kuiken & Vedder, 2011; Polio, 2012; Sheen & R. Ellis, 2011, Williams, 2012) helping learners notice the gaps in their learning: most WCF rarely goes unnoticed, unlike oral feedback that learners sometimes fail to attend to (Bitchener, 2012; Sheen, 2010).

As already mentioned, it should be emphasized that a lack of evidence of noticing at the level of understanding shown in the interview data does not mean that there was a lack of noticing at the attention level. Even though one interview learner, Tony from the focused WCF group, did not show evidence of Feedback Noticing at the level of understanding since he seemed to be confused about the cause of his errors, he did notice that he had made some errors (i.e., noticed at the attention level). Tony's lack of feedback understanding could be a result of his inadequate attention paid to the feedback he received or his lack of interest or motivation towards the feedback, and/or inadequate time to articulate the feedback he received as he may need more time than others to reflect on his errors. It is also probable that Tony might let some feedback go unnoticed because the feedback he received (involving with many "inflection morpheme "s") had "exceeded the subjective threshold" (Schmidt, 2001) and therefore exceeding his cognitive ability to cope with.

Another possibility could be that because of his limited grammatical knowledge about the structure, he was unable to form an accurate understanding of the nature of his errors. However, what is important here is that Tony, like others, noticed that he had made some errors and acknowledged that there was something wrong with his current grammar. Even though Tony only noticed without an accurate understanding of the causes or the rules underlying his errors, the fact that he noticed the errors at all mattered, as this gave him an opportunity to recognize the gap in his learning and that he needed to work on that gap if he wanted to achieve the target-like knowledge. Even if it might take him longer than others to come to understand and be able to solve the gap in his learning, it is better than not knowing that he has made any errors at all, assuming everything is all right. If Tony did not notice the gap in his learning, it is likely that he may get stuck in the same state for quite some time, and this may eventually lead him to the state of fossilization. Consequently, it might as well say that noticing, with or without understanding, increases learners' chances of success in acquiring a second language as it helps them to acknowledge their lacking and, as a result, encouraging them to work on fixing that hole in their knowledge.

All in all, based on the result of Feedback Noticing, WCF provided in this study was effective in inducing learners' noticing. The WCF provided did not only prompt the "noticing at the level of attention" but promoted the "noticing at the level of understanding," making learners recognizing the gaps in their knowledge; thus, finding ways to improve their current states. This, thus, suggests that by providing L2 learners with WCF, learners would have a greater chance to improve themselves in the areas they are lacking leading to successful mastery of the target features sooner than later.

While noticing is a crucial step underlying acquisition (Long, 1996, 2007, 2014; Schmidt, 1990, 1995, 2001), learners' uptake has been regarded as a pivotal predictor of successful L2 learning (Beuningen, De Jong & Kuiken, 2012; R. Ellis & Sheen, 2006; Loewen, 2004, 2005; Loewen & Philp, 2006; McDonough, 2005; Sachs & Polio, 2007; Sheen, 2010). Moreover, since learners need to notice first before they could proceed to uptake, noticing and uptake are inevitably associated. In Sachs and Polio's (2007) study, "noticing during the processing of written corrective feedback" was reported to be related to "subsequent revision changes" (p.85). Likewise, in Storch and Wigglesworth's (2010) study, a close relationship between learners' engagement with feedback and their corresponding uptake was also identified.

Aligned with the above studies, a strong affiliation between learners' noticing at the understanding level and their subsequent uptake (henceforth, Feedback Application) was also found in this present study. Ross, Monica and Tony's cases were clear examples representing the intertwined relationship between noticing and uptake. Based on these observations, it could be concluded that evidence of Feedback Application was frequently found to be reported together with evidence of Feedback Noticing like in Ross's case, where he reported both Feedback Noticing and Feedback Application evidence in both of his interviews. In addition, when noticing evidence was missing, no evidence of Feedback Application or uptake was reported as well, given Tony's case where there was no evidence of Feedback Noticing nor Feedback Application found in any of his interview sessions, for example.

Monica's case was the clearest example of this coexisting relationship, as while Monica did not report evidence of either Feedback Noticing or Feedback Application in her first interview session, she did report both Feedback Noticing as well as Feedback Application evidence later in her second interview session. Monica's case indicated that in the absence of noticing evidence (i.e., noticing at the understanding level), a process of uptake might not occur.

As stated earlier, even though Feedback Application or learners' uptake was mostly found to coexist with Feedback Noticing in this study, Feedback Noticing did not

always co-occur with Feedback Application. It was probable that not all that was noticed were proceeded to uptake or put into usage. The current findings support this speculation exhibiting that while most learners noticed the feedback at the level of understanding, only half of them demonstrated evidence of Feedback Application. In part, learners' lack of interest or motivation to improve their written work may attribute to why not all that were noticed were internalized into an uptake in this study.

It is undeniable that motivation as an individual difference factor plays a significant role in many aspects of second language acquisition. Like in other areas of learning, in the best light, motivation can unquestionably help promote acquisition processes. However, just as there are two sides to every coin, motivation can also impede or block the processes completely. In this study, motivation seemed to affect learners' level of noticing as well as Feedback Application to some extent. Tony's case is a clear example of how instrumental motivation may affect learners' Feedback Noticing and Application. The followings were Tony's views towards his learning and the feedback he received during the experiment:

> ***"If it is not for the test, I might not pay full attention to or focus on my study***. *It is like if we haven't had midterm exams just yet, we don't have to study hard, but when we are about to have the exams, we have to read so that we can do well on the tests,"* "*I made some improvements but not much as I didn't really go back and think about what I could have done better,* and ***"I didn't carefully check my feedback, maybe not that much."***

> From his comments, it could be assumed that Tony from the focused WCF group seemed to have a low level of motivation towards his learning in general, and that his motivation could easily drop down, especially when learning does not involve testing or grades. As a consequence, it was not unexpected that there was neither evidence of Feedback Noticing at the level of understanding nor Feedback Application found in any of his verbal reports. In stark contrast, Ross from the control group, who showed strong motivation towards learning during the time of the study, also demonstrated evidence of both Feedback Noticing at the level of understanding as well as Feedback Application in all his interview sessions. From Tony's to Ross's cases, motivation seemed to be accompanied by Feedback

Noticing and Application.

Unlike Tony's and Ross' cases, Amelia from the focused WCF group and Monica from the control group did not show clear evidence of Feedback Noticing (for Monica) and Feedback Application (for both of them) in their first interview, despite their strong motivation recounts. However, the two of them demonstrated evidence of both in their second interview. An absence of Feedback Noticing and Feedback Application evidence in their first interview did not implicate that motivation did not have any effect on Amelia's and Monica's noticing or feedback application. Instead, this suggested that there might be other factors besides motivation moderating the results. Monica, for example, might need more time to articulate and understand her content feedback as it constantly changed depending on each week's assigned topic, and/or she was overwhelmed by the amount of feedback she received. The same applied to Amelia, who might need more time to uptake all the feedback she received.

Another observation worth mentioning here would be that learners' Feedback Application reports appeared to be incongruent with learners' test scores. While Ross, Monica, Jackson and Amelia's Feedback Application reports seemed to be congruent with their quantitative test results, Meredith and Tony's lack of Feedback Application evidence did not seem to be in tandem with their test scores. This is because albeit an absence of Feedback Application evidence, Meredith from the unfocused WCF group and Tony from the focused WCF group still showed observable improvement on their test performance after treatment. That is, Meredith and Tony's UGJT and timed writing scores increased after treatment. Their delayed writing scores were also noticeably higher than their pretest scores. This, thus, suggests that, in reality, Meredith and Tony might apply what they noticed from the given feedback to improve their work, but because they may not articulate well during the interviews leading to a lack of their Feedback Application recounts after all.

In addition, the incongruence between the qualitative reports and quantitative test data pointed to the fact that while the interview method can unveil more in-depth information that quantitative measurements cannot, the method has its own limitation. The main shortcoming of the method is that it is difficult to verify the veracity of some, if not all, of the interview or self-reported data. Still, the interview method is the widely

most accepted method used to acquire more thorough information on learners' thoughts and their cognitive processes. Acknowledging this shortcoming, caveats should then be exercised when attempting to interpret the information obtained from this subjective measure.

In essence, interview results suggest that Feedback Application could not occur in the absence of Feedback Noticing, though the opposite is true for Feedback Noticing. The results also provide support to the claim that WCF is beneficial for L2 learning, given that WCF helped learners notice the gaps in their learning and facilitated learners' uptake processes in this study. Learners' motivation level is also regarded as a key moderator affecting learners' propensity to notice and fully benefit from the feedback they receive.

5.3.1.3. Motivation

Motivation is a powerful predictor of learners' successful second language acquisition (Csizér & Dörnyei, 2005; Dörnyei & Ryan, 2015; Gardner & MacIntyre, 1991; Noels, Clément, & Pelletier, 1999; Schmidt & Watanabe, 2001). Motivation explains why learners choose to invest in particular learning activities while disregarding others, how they are willing to persevere with or give up on certain respects of learning, and how much effort they are going to invest in each given task (Dörnyei, 2010; Dörnyei & Ushioda, 2009, 2011; Gardner, 1985, 2000, 2001; MacIntyre et al., 2001). Based on this exposition, it was logical to assume that in this study motivated learners were likely to pay more attention to the feedback they received (showed more evidence of Feedback Noticing) and made more effort to learn and improved their work (showed more evidence of Feedback Application) when compared to the unmotivated ones.

Previous results on Feedback Noticing and Feedback Application provide support to the above supposition as interview learners who demonstrated strong motivation recounts seemed to render more evidence of Feedback Noticing and Feedback Application than those with a lower level of motivation or a lack thereof. To be more precise, motivation results exhibited that Ross and Monica from the control group, and Amelia from the focused WCF group, who demonstrated clear motivation evidence, were

the same learners who reported evidence of Feedback Noticing as well as Feedback Application.

In contrast, Jackson and Meredith from the unfocused WCF group, and Tony from the focused WCF group, who lacked clear evidence of Motivation also lacked evidence of either Feedback Noticing, Feedback Application, or both. Even though Monica from the control group, a motivated one, did not demonstrate Feedback Application evidence in her first interview session, similar to Jackson, Meredith, and Tony, with her seemingly high motivation level, she eventually succeeded in using feedback for the benefit of her work later on. In contrast, Jackson, Meredith, and Tony who did not show any motivation recounts also did not show any evidence of Feedback Application in any interview sessions. Likewise, with Amelia's unwavering motivation to learn, despite a lack of Feedback uptake or application at the beginning of the study, Amelia from the focused WCF group later showed evidence of Feedback Application in her second interview session.

Monica and Amelia's cases are good examples exemplifying Dörnyei and Ushioda's (2011) postulation stating that motivation can lead to successful acquisition only when it is continuously sustained over time. Anyhow, it should also be noted that even though half of the learners in this study did not show clear motivation recounts as elucidated above, it does not necessarily mean that these learners were completely unmotivated. Instead, there is a high probability that these learners might possess a lower level of motivation compared to the ones with distinct evidence.

Another observation worth mentioning here is the fact that motivation seems to associate with Feedback Application than Feedback Noticing. This is because even though Jackson and Meredith from the unfocused WCF group did not show any clear motivation traits during the time of study, both displayed clear evidence of Feedback Noticing. This in part distributed to the saliency and frequency nature of WCF that it was hard even for unmotivated learners to let it pass unnoticed. Nonetheless, while Jackson and Meredith were able to notice WCF at the level of understanding, without much motivation, both did not exhibit any Feedback uptake evidence later on. Noticing could indeed lead to uptake, but it is not always the case, part of that pertains to this motivation

factor. Given that the learner was not motivated to invest in the provided learning activity (like in this case, assuming that they were not interested in applying the knowledge acquired from feedback to improve their work), it was unlikely that he/she would attempt to advance their work further. The observation above hence suggests that motivation might play a more crucial role in learners' Feedback uptake or application than Feedback Noticing. It is speculated that motivation might affect the amount of input learners notice, with the more motivated ones are able to notice more.

As aforementioned, motivation is the drive behind the success or failure of second language learning since it is a key factor influencing the extent to which the learners notice and apply what they acquire from feedback into practice. Although motivation is likely to have more effect on learners' Feedback Application than Feedback Noticing, it is speculated that learners with higher levels of motivation had a tendency to notice more input than those with the lower-level ones.

5.3.1.4. Feedback usefulness

Motivation is a dynamic, multifaceted, and complex phenomenon (Csizér & Dörnyei 2005; Dörnyei, 2010). With its dynamic nature in conjunction with recent effect of globalization, motivation has thus been reconceptualized by different groups of scholars for the past decades. Apart from the well-respected Gardner and Lambert's (1959) L2 motivation model of integrative (the desire to be a "passable" member of the target L2 community) and instrumental motivation (utilitarian value of learning the target language for better grades, future career, a pay raise and so forth), many theories have emerged and among them are the well-accepted "self-efficacy beliefs" model by Bandura (1986) and "motivation self-system" theory by Dörnyei (2005).

Bandura (1986) elucidates self-efficacy as "people's judgments of their capabilities to organize and execute courses of action required and designated types of performances" (p. 391). To simply put, self-efficacy concerns the person's belief whether he/she is capable of performing the given tasks and this belief after all affects his/her achievement in performance. It is proposed that if the person has a positive self-efficacy

belief that he/she has an ability to perform the task, this view could enhance his/her performance leading to the successful accomplishment of his/her determined goal.

The L2 motivation self-system proposed by Dörnyei (2005) is quite similar to the self-efficacy beliefs since it also associates with the way learners perceive themselves. In this theory, there are two components of "self": the ideal and ought-to self. The ideal self involves the learner's projected future self-image of becoming a competent user of the target language. In contrast, the ought-to self is related to the learner's belief to behave in a certain way as a result of his/her responsibility or obligation so that he/she could avoid the negative consequences such as bad grades or parents' punishment (Dörnyei, 2005). Regardless of what motivation really comprises of, however, one thing is clear: motivated behavior is directed by positive views towards the goal or what the learners are doing. If the learners are not satisfied, do not foresee the benefit of doing something, or do not believe that the goal is after all achievable (thus, inducing negative views towards what they are doing), it is unlikely that they would pursue the course and invest in that action or activity. In short, motivation is primarily established based on a person's positive view and confidence towards his/her ability to achieve. Consequently, to motivate the learners is to make them feel positive about themselves or the activities they are engaged in.

Grounded on the above models, the study proposes that learners' positive views towards the usefulness of feedback on their writing advancement, one way or another, attributed to the increase in their learning motivation. This is because learners' positive views towards feedback could help them to feel more reassured and confident towards their writing abilities; thus, inducing their positive self-efficacy beliefs, ideal self, and their motivation to learn. Given that all learners viewed feedback as a useful pedagogical tool that helped them recognize their errors or aspects they needed to improve and improved their writing abilities (i.e., had positive views towards the provision of feedback), feedback can thus be viewed as an effective motivator helping learners become motivated.

Some might argue that if all learners were somehow motivated by feedback, why only half, not all of them, showed motivation evidence. A plausible explanation could be that even though all learners agreed that feedback was a useful tool and they preferred it

to no feedback, not all shared the same level of expectation towards what feedback can actually provide. In other words, while some learners considered feedback as "essential" and it was important for their writing development, some might view it as an "additional" tool, better than receiving nothing. The varying degree of necessity and usefulness of feedback each learner viewed clearly affected the level of motivation he/she developed. This might reflect on the absence of some motivation recounts accordingly. That is, it is probable that the learner who viewed feedback as "essential" may possess a higher level of motivation, and because of that, his/her motivation trait was reflected through his/her verbal account. In contrast, the learner who did not have strong view towards feedback may have a lower level of motivation and, as such, did not show motivation trait through his/her recount.

Another possibility could be that even though these learners valued the provided feedback and they were indeed motivated by it, it was still possible that they might not elaborate this during the interviews. From the interview observation field notes, some learners like Meredith from the unfocused WCF group and Tony from the focused WCF group appeared to be more reserved than others, and they tended not to explain their answers in detail unless required. This might in part explain an absence of motivation evidence.

Nonetheless, caution needed to be exercised when interpreting this result. It might be true that all interview learners had positive attitudes towards feedback provision and were somehow motivated. However, the data did not warrant that all learners were able to maintain their motivation from fluctuation during the 9-month study period. In fact, the data from observation field notes showed that learners' motivation seemed to dramatically drop during the delayed posttest.

In conclusion, interview data on Feedback Usefulness demonstrated that all learners viewed feedback as an effective tool and preferred to have it in their writing classes. The data also suggests that since feedback gave learners a sense of controlling and confidence towards writing achievement, feedback might be an effective motivator motivating learners to be more engaged in and make more effort on their learning. The data, however, did not suggest that because of the positive attitudes towards feedback, all learners might possess a high level of motivation during the time of study. As motivation

199

is one of the decisive factors accounting for differences in second language achievement, future research on how teachers' feedback or a lack thereof could have an effect on learners' motivation is thus beneficial.

5.3.1.5. Self-rating progress

Previous motivation results had shown that half of all learners exhibited clear motivation evidence while the rest did not. The results implicated that the ones without motivation evidence might possess a low level or a lack of motivation during the time of the study. However, when scrutinizing the motivation results in combination with the Feedback Usefulness findings, the analysis suggested, it was more probable that the interview learners possessed a varying degree of motivation rather than a lack thereof. This conclusion was based on the premise that 1) a lack of clear evidence does not necessarily mean that the opposite is true as there might be other factors at play and more empirical evidence is required, and 2) the fact that all learners had positive views towards the feedback believing it could help them to write better reflected their "positive attitudes towards learning activity" which in Dörnyei's (1990, 2003, 2005, 2010) and Gardner's (1985, 2000, 2001) terms, denote motivated behavior. Nonetheless, as noted earlier, even though feedback had shown to give learners confidence towards writing and boost their motivation levels, the data did not suggest that because of it, all interview learners had developed a high motivation level at the time of the study.

In addition, the self-rating progress results showed that 71 out of 73 learners believed they had made progress during the experiment. Learners' positive attitudes and high confidence towards their abilities or progress were postulated by Dörnyei (1990, 2003, 2005) and Gardner's (1985, 2000, 2001) to characterize motivation. This, thus, corroborates the above assumption that most learners in this study possessed varying degree of motivation even though some of them might not demonstrate motivation evidence. Several other factors may have an impact on their motivation levels such as learners' expectations towards the experiment, their learning goals, and so on.

For example, during the individual tests, some learners from the control group commented that they preferred grammar to content feedback and asked why they only

received content feedback. From this instance, these learners might not feel that their needs or goals were met, and this might somehow lower their motivation levels. Another instance was that as the experiment progressed, many learners started to complain about the fact that they had to retake the tests (pre-post-delayed) several times both individually and in group as the tests were administered outside of their regular course schedule. In the delayed posttest (three months after treatment), it was obvious that most learners spent half the time they spent on their pretest and posttest to complete the delayed posttest. Some even withdrew from the study. Some of these learners mentioned that if they were given some credits in return, they would be more willing to invest in the experiment. This incident evidently indicated that an absence of instrumental motivation provided in this study (course credits/grades or other incentives) can affect learners' motivation to a great extent.

Another equally noteworthy observation that should not be disregarded would be how most learners from all groups, regardless of the feedback types they received, unanimously agreed that they had made an improvement in their written content. Even those from the focused and unfocused WCF groups who received only grammar feedback believed their written content was also improved. This result implicated that 1) writing practice in general helped improve learners' written content, even in an absence of feedback (however, the ones who received content feedback improved their content better), and 2) progress on written content was more evident than grammatical progress, and that was probably because it took longer to develop the grammatical knowledge. While learners from all participating groups including those who received only grammar feedback, believed that they had made noticeable progress on their written content, some control group learners who only received content feedback also believed that their grammar had been improved as well. Interestingly, in line with their Self-Rating Progress reports, most control group learners substantially improved their plural scores from pretest to posttest. This implicated that even though these learners only received content feedback, they were still attentive to their grammatical usage.

At this point, some might argue that if receiving or not receiving grammar feedback did not lead to disparate learning outcomes, it would then be better and more

logical for teachers to stop giving learners WCF (i.e., grammar feedback) and invest their time on other effective methods. Notwithstanding, care should be exercised before drawing such generalization based on this idiosyncratic context. This was because, first of all, these learners were at the intermediate level, implicating that all might already have some established knowledge of English plurals (Thai learners received plural instruction since primary school). In other words, the plural structure was not new knowledge for them, and as a result, albeit the absence of plural feedback, the control group learners were still able to correct their own plural errors to some extent, if they paid close attention to their plural usage.

Second, grammar is always a primary concern of Thai learners, for they have been trained to focus on their grammatical accuracy since an early age and this is speculated to have helped them cultivate strong grammar sensitivity aptitude. It is thus not unexpected that even in the absence of grammar feedback, learners in the control group still pretty much attended to their grammar usage and could make some grammatical progress. With all these factors, it is then plausible that learners in the control group, even without grammar corrections, can markedly increase their plural scores. However, it is important to note that because of the distinctive characteristics of this group of learners and context as noted above, it is then premature or even detrimental to conclude that content feedback is equally effective as grammar feedback in helping learners improve their grammatical accuracy. Differences in educational settings (ESL vs. EFL), proficiency levels (beginners, advanced learners, etc.), target linguistics features (simple vs. complex), instructional contexts (explicit vs. implicit learning), might interact distinctively with different types of WCF, leading to differences in learning outcomes.

All in all, the self-rating progress indicated that most learners seemed to have positive views towards their writing progress, signifying that they might have a certain level of motivation while participating in the study, though the levels of their motivation may vary.

The previous Self-Rating Progress report revealed that most learners believed they had made progress on their written content and grammar, respectively. Aligning with the results of Self-Rating Progress, the Writing Concern data demonstrated that most learners had their utmost concern over their written content, then grammar and wanted to improve these respects accordingly. This somehow suggests that there were correlations between learners' concerns, desire for improvement, and their self-reported progression, together with their actual development (i.e., test results).

Further analysis on each group's data demonstrated the same results, as all groups were concerned over and attempted to improve their written content the most. This suggests that most learners in this study might orient themselves more towards "meaning" rather than "forms" when practicing writing or taking the tests. This implicates that learners might resort to their implicit knowledge rather than their explicit knowledge during the experiment, for their focus was on meaning instead of forms.

For the control group, it was as anticipated that the learners were most worried about their written content, considering that they only received content feedback for the whole six-week treatment. However, it was interesting to mention that while 38% of their concerns was on written content development, 33% of their concerns was on grammar, even though they did not receive any grammar feedback during the time of the study. The data further revealed that, indeed, the control group had the greatest concern over grammar usage (33%) among the three participating groups. The data, thus, corroborates the previous supposition proposing that in part, it could be the learners' grammatical sensitivity, which was cultivated through years of exposure to the form-focused instruction, that helped the control group learners to be more aware of their grammatical errors in their writing contributing to their grammar progress.

In conclusion, content and grammar were writing aspects the learners from all participating groups were concerned about and wanted to improve the most. In addition, it is conjectured that the more the learners focused on their written content (meaning-focused), there is a greater possibility that the learners might rely more on their implicit rather than their explicit knowledge. In the same vein, the more the learners focused on

their grammatical aspects (form-focused), the greater the chance they might resort more to their explicit knowledge when performing the tasks.

5.3.1.7. Research goal

Previous Writing Concern results suggest the possibility that most learners in this study might take the meaning-based approach during the experiment, considering that their main aim was to develop their written content. The Research Goal results from 75 exit questionnaires provided evidence in support of the above claim that most learners in this study focused more on "meaning" rather than "forms". Given that only 3 out of 75 learners recognized the role of WCF played in this present study (the true purpose of the study), while the rest (72 learners) did not[56], this indicated that the majority of them might not pay enough attention to the feedback provided. The fact that they paid very little attention to the given WCF, while paying most of their attention to content development, as previously illustrated, can only point to the assumption that most learners seemed to orient themselves more towards meaning-based learning approach rather than the form-focused one during the investigation.

There are two main factors attributing to why learners extended themselves more towards the meaning-based approach. The first pertained to the way the writing practice task was introduced in class. In this study, the writing practice task was always introduced within a meaningful context (through small talk and discussion about learners' topics of interests), where learners were encouraged to mainly focus on developing good content essays. That is, learners were instructed to pay attention to their written content but not grammar. Even in the WCF groups, the role of WCF was not emphasized. Learners were only informed to check their feedback but were not asked to do anything with it. Developing a strong argumentative essay appeared to be the only goal of the practice. In part, this is to prevent the learners from becoming exceedingly focused on their grammar and from reviewing their grammar outside treatment sessions, adding another intervening

---

[56] Seventy-two out of seventy-five learners believed that this study either aimed to assess their writing ability, how writing practice helped enhance their writing skill, or how writing might affect other English skills. Few learners conjectured that this study was about the examination of learners' short-term memory capacity.

variable to the study. Besides the meaning-focused instruction, every week, a few rewards were given to learners who wrote the best argumentative essays. The essay was awarded if it had strong argumentative content. The learners were also informed that grammar was not the focal point of this essay competition as long as the meaning was intelligible. This explained why the learners became more focused on their content rather than grammar development.

While the learners were directed to concentrate on developing strong argumentative essays in hope that it could refrain them from being overly focused on their grammar usage, the role of WCF was not adequately emphasized as already delineated. As a result, the learners might not think that they needed to earnestly attend to the given feedback. Considering that the learners were instructed as such during the experiment, it was not unanticipated that the learners eventually turned to adopt the meaning-based learning approach.

All in all, it could be concluded that the learners in this study regardless of the feedback types they were provided had a high propensity to endorse the meaning-based approach rather than the formed-focused one. This probably is a result of the meaning-focused instruction during the experiment, combined with the less emphasis on the role of feedback, a form-focused tool. The fact that the learners extended themselves more towards a meaning-based approach as such might, to some extent, have an influence on the quantitative test results.

5.3.2. Conclusion

The qualitative part of this study was developed in hope that it could cast additional light on the main quantitative findings uncovering factors that might have an impact on learners' learning outcomes. Based on qualitative interview and exit questionnaire analysis, the findings showed that WCF effectively helped induce noticing at the level of understanding, especially for motivated learners, and that feedback application or feedback uptake was unlikely to occur when there was a lack of noticing at this level. It was also evident in this qualitative part of this study that learners' motivation, their background knowledge on the target feature, grammatical sensitivity aptitude and instructional context (meaning or form-focused) seemed to moderate the effectiveness of the given WCF to some extent. Notwithstanding, the interpretation or generalization of these findings has to be executed with caution since, despite the careful design, this part of the study still has its own limitations.

The first limitation of the qualitative part of this current study pertains to a small number of learners taking part in the qualitative interviews. Since only two learners were drawn to represent each group's responses on feedback or the overall treatment, it might be quite challenging to generalize what had been found in this study to the wider population, especially those from diverse backgrounds or learning experiences. In addition, there is also an incongruence of data between the data derived from the interviews and exit questionnaires. Ideally, specific questions regarding Feedback Noticing, Feedback Application, and Feedback Usefulness, as seen in the qualitative interviews, should have been added in the exit questionnaires as well since the data could strengthen the above findings.

Another limitation involves the ecological validity of the study. Despite the benefits of the experimental design this study adopted, whether it is the high reliability of the outcomes or the better generalizability of the findings, it is undeniable that to some extent, the results obtained from this design do not always well reflect or predict real-world behaviors. This is because while it is possible to control all variables within the experiment, it is more challenging to control or isolate those variables in real-life settings. Acknowledging its shortcomings, this study attempted to resemble the real learning

setting as much as possible. However, admittedly, this unreal situation, after all, seems to have an impact on learners' motivation. It could be assumed that since learners received nothing from participating in the project (except for the small writing rewards), their motivation during the experiment might be different from their normal motivation levels when they are in the real classroom setting, and their grades are at stake. That is, in the real classroom setting, the learners might be more motivated to perform the tests. Nonetheless, even though the current findings might not well reflect learners' real-world behaviors or motivation levels, the present investigation might not be able to unequivocally verify the effectiveness of WCF in a less controlled environment either.

Finally, the last shortcoming of the qualitative analysis concerns the writing task and instruction given during treatment. As mentioned earlier, to prevent learners from exceedingly focused on grammar, more meaning-focused instruction was implemented while less emphasis was placed on WCF. While this might not seem to pose any threatening problems towards learning in general, the imbalance between meaning and forms might lead to distinctive learning outcomes and implications. As when "meaning" is excessively focused, it is also likely that "forms" will be neglected or vice versa, and this could have biased the overall test results.

In fact, theoretically, most SLA scholars seem to endorse the proportional combination of meaning and formed-focused instruction rather than the extreme focus on one's end. This is because even though fluency (meaning-focused) should be developed first following L1 acquisition route, accuracy or form-focused instruction is deemed necessary, especially for L2 learners since a lack thereof could lead to learners' language fossilization. Nonetheless, to reiterate the point already mentioned, because most Thai learners are already exceedingly concerned over their grammar usage, to emphasize a form-focused instruction during treatment might direct them to become more focused on forms adding another intervening variable to the study (if the learners reviewed their grammar for the better test scores). Further, given that these learners already have some pre-existing plural knowledge, it was expected that despite receiving the more meaning-focused instruction, they would not completely disregard their grammar. Data on Writing Concerns corroborates this speculation. Further discussion regarding how the more

meaning-focused approach provided in this study impacted the quantitative results is offered in the next chapter.

Despite the above limitations, the qualitative findings presented in this chapter revealed learners' thoughts and some other factors that might influence their learning outcomes. The qualitative findings also appeared to be consistent with the quantitative test results of the previous chapter, and this is assuredly beneficial for the in-depth analysis of the quantitative results. In the next chapter, the combined discussion of the quantitative and qualitative results is offered, followed by theoretical and pedagogical implications. The chapter ends with the limitations of the study and recommendations for future research.

Chapter VI

DISCUSSION and CONCLUSION

In this final chapter, results from the quantitative and qualitative parts of the study are discussed together in a combined analysis. For the credibility and validity of the yielded results, data obtained from both parts of the study were cross-examined. Both theoretical and pedagogical implications of the study are also presented. In the last section, limitations of the current investigation are offered together with avenues for future research.

6.1. Discussion

The current study had two main purposes: 1) to investigate the effectiveness of WCF on the development of L2 learners' explicit and implicit knowledge of English plurals, and 2) to determine the extent to which L2 learners' WMC mediated the effectiveness of WCF on learners' learning outcomes. The following discussion drew upon the findings obtained from both quantitative and qualitative parts of the present study and is contended in relation to relevant theories and previous research investigations.

Effects of Written Corrective Feedback on the Development of Learners' Explicit and Implicit Knowledge of English Plurals

The first and second research questions (RQ 1 and 2) asked whether the two types of WCF, direct *"focused"* and *"unfocused"* WCF, provided in this study had any effect on learners' acquisition of explicit and implicit knowledge of English plurals. The results of statistical analyses suggest that both types of WCF had positive effect on the development of learners' explicit and implicit knowledge, given that the WCF groups performed significantly better on both their explicit and implicit knowledge tests after treatment. Additionally, the WCF effect on learners' implicit knowledge had shown to be sustained over time providing that learners' delayed posttest scores on the timed writing, OEIT, and combined implicit knowledge tests were significantly higher than their pretest scores. Learners' significant improvement in all implicit knowledge measurements immediately after treatment and three months later during the delayed posttest affirms the effectiveness of WCF in aiding learners' implicit knowledge development. As Shintani and R. Ellis (2013) contend that "We argued that if the feedback had had an effect on implicit knowledge, this would have been evident in the delayed as well as the immediate writing task" (p. 301).

The findings, thus, corroborate previous claims on the effectiveness of WCF in promoting learners' grammatical accuracy, and the increment in accuracy denotes the development of explicit knowledge (Benson & DeKeyser, 2018; Bitchener & Knoch, 2009a, 2009b, 2010; Bitchener & Ferris, 2012; Bitchener & Storch, 2016; Shintani & R. Ellis, 2013; Shintani et al., 2014; Stefanou & Révész, 2015; Van Beuningen, De Jong & Kuiken, 2012). Most importantly, the results provide evidence to support previous findings (e.g., Nemati et al., 2019; Rezazadeh et al., 2015; Shooshtari et al., 2019) which found the positive effect of WCF on the development of learners' explicit and implicit knowledge. Moreover, the findings negate Truscott's (1996, 1999, 2004, 2007, 2010) argument contending that WCF is futile and that it has no effect on the development of learners' L2 knowledge.

However, the current findings differ from those of Shintani and R. Ellis' (2013) study which also investigated the effectiveness of WCF on the development of learners' implicit knowledge but found null results. Nonetheless, the differing findings are

speculated to result from differences in grammatical structures targeted in each study, the amount of correction provided, opportunity for language production, and implicit knowledge measurements each study adopted (see chapter 4 for detailed discussion).

In essence, as already elucidated in previous chapter (chapter 4), WCF was effective in helping learners develop their explicit and implicit knowledge of English plurals because it provided learners with more opportunities to notice the gaps in their learning, especially giving them opportunities to notice the less salient plural errors potentially blocked by their L1 cues. For learners, noticing the forms is a significant stage. Only when learners notice the forms and acknowledge the gaps in their learning can they begin to fine-tune their current knowledge leading to increments in their L2 knowledge. As many SLA researchers (Bitchener & Storch, 2016; DeKeyser, 2007, 2015; Gass, 1997, 2003; Gass & Mackey, 2015; Long, 1996, 2007, 2014; Polio, 2012; Schmidt, 1990, 1995, 2001; Swain, 1985, 1993, 1995, 1998 2005; Van Beuningen, 2010; Williams, 2012) have pointed out, feedback, as a consciousness-raising or form-focused device, can effectively help learners to notice the mismatch between their output and the target-like input, prompting the destabilization and reconstruction of learners' interlanguage grammar. This premise is also substantiated by this study's qualitative interview data.

Based on interview data, while four out of six learners displayed evidence of noticing at the understanding level during the first interview session, five learners demonstrated evidence of noticing in their second interview. This suggests that the explicit nature of WCF seemed to effectively help promote noticing among feedback receivers. Significantly, the noticing at the understanding level most learners displayed is posited to facilitate subsequent L2 acquisition processes (Schmidt, 1990, 1995, 2001). Taken together, it could be conjectured that the WCF the learners received during the 6-week treatment sessions seemed to effectively help induce learners' noticing, making them become more conscious of the gaps in their learning, resulting in significant improvement in their test performance. Besides inducing more noticing among feedback receivers, the provided WCF seemed to significantly raise learners' conscious awareness towards the target form.

Jackson from the unfocused WCF group commented in one of his interviews that because of the feedback he received, "***I become more aware of my grammatical errors,***

more careful about them. Similarly, Amelia from the focused WCF group, reported that ***"From my recent feedback, I could see my progress. Only that some types of grammatical errors like plural errors just kept reappearing. I think I need more time to write. Honestly, I didn't have time to recheck and correct my errors."*** Amelia seemed to be highly aware of her persistent plural errors as a result of feedback. In an exit questionnaire survey, Suzanne from the focused WCF group wrote on her exit questionnaire when asked if her writing had been improved during the past months at all: ***"Yes, it has, especially my grammar. Normally when I write, I just write without thinking about grammar. These past months had taught me and reminded me to write more grammatically".*** Patricia, another learner from the unfocused WCF group, also wrote that ***"My writing has improved for the past months. Participating the project helped me become more aware of some grammatical aspects as well as my writing organization".*** The reports from these learners emphasize the role of WCF as a consciousness-raising tool helping feedback receivers develop more awareness towards the target forms and this prompted them to compare their pre-existing knowledge about the target structure with the new knowledge brought by feedback. Learners were then able to form deeper understanding about the underlying rules of the target structure, resulting in an increment in their knowledge.

As N. Ellis (2016a) comments that feedback as a formed-focused and consciousness-raising tool is beneficial for learners' explicit and implicit knowledge development because it helps induce learners' noticing or learned attention and raise learners' metalinguistic awareness to the less salient and redundant linguistic forms blocked by learners' L1 cues. This noticing facilitates the consolidation of form-function mappings of new constructions (i.e., explicit knowledge) and "Once a construction has been represented in this way, its use in subsequent implicit processing can update the statistical tallying of its frequency of usage and probabilities of form-function mapping" (Wulff & N. Ellis, 2018, p. 51). "With enough usage examples, implicit learning [and implicit knowledge] might get there in the end" (N. Ellis, 2005, p. 316).

Based on the above exposition of L2 development processes, the primary functions of feedback are to help induce learners' noticing and raise learner's awareness towards the target form facilitating the form-function mapping processes. Nonetheless,

this study further argues that the type of WCF that provides learners with both positive (correct forms) and negative evidence (error indication) like direct WCF provided in this study may not only help inducing noticing but also help promote implicit learning processes by providing learners with adequate target exemplars (i.e., plural forms) allowing the priming and implicit tallying of the target exemplars facilitating the development of implicit knowledge. In other words, it is posited that after frequent encounters with target exemplars introduced by direct WCF through practice for a prolonged period, even though the learners' focus was not oriented to forms (this study adopted a highly implicit learning approach), the learners might begin to implicitly tally the target exemplars provided by feedback and subsequently integrating them with their existing knowledge. When the implicit learning system has sufficient exemplars, it begins to create knowledge representation (i.e., prototype), contributing to learners' implicit knowledge development (see N. Ellis, 2005, 2011, 2012, 2015b, 2016a; Wulff & N. Ellis, 2018). The processes illustrated follow the connectionist model (see Christiansen & Chater, 2001; MacWhinney, 1999) and usage-based framework (Goldberg, 1995, 2003; N. Ellis, 2005, 2012, 2015b; 2016a; Wulff & N. Ellis, 2018), which elucidate the process of language acquisition as the frequency of input occurrence leads to input processing, and once the input is integrated into the system as a result of frequent exposure, it stimulates other already existing or relevant input to work together and yields the linguistic prototype and category for future usage.

In sum, the direct WCF provided in this study can effectively enhance learners' explicit and implicit knowledge development for two main reasons. First, it helped learners notice the gaps in their learning (i.e., their persistent plural errors) and become more aware of the gaps or the less salient forms, deepening their understanding of the target structure's underlying rules; thus, consolidating the form-function mapping processes. Second, it provided learners with sufficient target exemplars (i.e., correct plural forms) to be implicitly tallied, leading to the successful acquisition of the target linguistic feature.

Three observations concerning the above results are worth discussing. The first concerns the role of noticing played in these learners' learning. As mentioned earlier, five out of six interview learners showed evidence of noticing at the understanding level. The

noticing at this level is posited in the field to be significantly linked to the extent to which learners are able to uptake and apply the feedback they received to improve their work[57]. However, the interview results had shown that even though Tony from the focused WCF group did not show any evidence of noticing at the understanding level similar to his peers, he was still able to make noticeable improvement on his explicit and implicit knowledge. Based on his interview account, Tony only demonstrated evidence of noticing at the attention level providing that he acknowledged the feedback, yet he did not seem to understand the feedback he received (i.e., he did not understand the nature of his errors).

> **Tony: *I received feedback about subject-verb agreement…that I forgot to add "s" or about tenses.***
>
> **Researcher: *Was it about plural?*** (Tony was from the focused WCF group, which only received corrections on plural errors)
>
> **Tony: *I don't know…there were many "s" in my feedback.***
>
> **Researcher:** Did you carefully check them?
>
> **Tony: *Maybe not that much.***

From the extract, Tony did not seem to understand the feedback he received and assumingly may not have accurate understanding of the target form. However, his test performance showed that his MKT, UGJT, and timed writing posttreatment scores substantially improved after treatment. His timed writing delayed posttest score was also superior than his pretest score (23 vs. 38). Thus, it is posited that even though Tony did not seem to develop accurate metalinguistic knowledge of the target form (i.e., notice at the understanding level), the noticing at the attention level is adequate for him to implicitly tally target exemplars (i.e., plural forms) provided by feedback into his implicit learning system. Through frequent exposure and usage experiences, Tony eventually successfully established the form-function mapping and language representation making progression on his explicit and implicit knowledge test performance. Tony's case indicates that noticing at the attention level is sufficient for L2 learning to take place. His case also provides evidence in support of Schmidt's (1990, 1995, 2001) and N. Ellis'

---

[57] It is posited in the field that noticing at the understanding level is significantly correlated with learners' subsequent uptake. That is, the more the learners notice at the understanding level, the more likely they are able to uptake the input or feedback (see chapter 5).

(2005, 2015a, 2016a) proposition suggesting that it is the noticing at the attention level[58] that is essential for the development of learners' L2 knowledge.

Another issue that deserves further attention is that the learners in this study made greater improvements on their implicit knowledge tests than their explicit knowledge tests. In addition, the effect of WCF on implicit knowledge tests had shown to be sustained over time, considering that learners' OEIT and timed writing delayed posttest scores were significantly greater than their pretest scores. In part, learners' greater improvements on their implicit knowledge tests is perceived to be a result of Transfer Appropriate Processing (TAP), which suggests that learners can better remember and transfer "what [they] have learned if the cognitive processes that are active during learning are similar to those that are active during retrieval (Lightbown, 2007, p. 27). That is, learners' memory could be best recalled when the learning and production (or in this case, testing) shared similar conditions, for example, "learning to use language in a communicative context may improve the ability to retrieve it in such contexts" (Lightbown, 2007, p. 27). Consequently, given that the learners in this study were oriented more towards implicit learning (meaning-based approach) during treatment and the implicit knowledge tests were also meaning-focused, suggesting that learners' instructional context and assessment tasks probably share the same encoding and retrieving cognitive processes, this, to a greater extent, explains why learners performed significantly better on implicit knowledge tests, as opposed to the explicit knowledge (i.e., form-focused) tests.

The tenet of TAP also explains why the learners obtained higher gains on their timed writing task than the OEIT. While the timed writing task and the OEIT were both meaning-focused, learners performed better on the timed writing task as they were trained on the writing task similar to the writing test during treatment, while they were not trained

---

[58] In Schmidt's Noticing Hypothesis (1990, 1995, 2001), noticing at the attention level (i.e., awareness) is claimed to be requisite for L2 learning. Without it, learning cannot take place. In contrast, noticing at the understanding level is facilitative but not necessarily essential. In the same vein, N. Ellis (2005) contends that "Although noticing is not necessary for priming and tallying, attention is" (p. 311). Note that, noticing in N. Ellis' term refers to "noticing at the understanding level" and "attention" in his term refers to "noticing at the attention level" in Schmidt's terms.

on their oral production before. As a result, learners can better transfer the knowledge they gained during writing practice to be used during the timed writing test, helping them gain high scores on their time writing test. Meanwhile, it may be more challenging to retrieve and transfer what they have learnt (i.e., encoded) in written context to be used in oral context, such as when they performed the OEIT. In addition, since both the practicing writing task and the timed writing test shared identical encoding and retrieving processes which facilitated knowledge transfer, this attributed to why learners performed better on their timed writing test.

Taking this line of thought further, learners' superior performance on the implicit knowledge tests as opposed to explicit knowledge tests also indicates that implicit learning (i.e., meaning-focused) might predominantly promote the development of implicit knowledge and that explicit and implicit knowledge might have distinct routes of development as many researchers have previously postulated (Andringa & Rebuschat, 2015; N. Ellis, 2005, 2012, 2015b; 2016a; Suzuki & DeKeyser, 2017; Paradis, 2009; Wulff & N. Ellis, 2018; Ullman, 2004, 2005, 2015; 2016; Ullman & Lovelett, 2018).

According to usage-based theory, explicit and implicit knowledge are distinctively developed (N. Ellis, 2005, 2012; 2015b, 2016a). In this respect, N. Ellis (2005, 2012; 2015b, 2016a) explains that explicit learning involves conscious registration of linguistic features or patterns that are later bound together to establish a coherent knowledge representation. Conscious awareness, thus, plays a crucial role in explicit learning and explicit knowledge development. However, the development of implicit knowledge is different. It begins with the integration of "new information with existing knowledge, using overlapping distributed representations to extract the general statistical structure of the environment" (p.319). In other words, through the high frequency of input exposure or usage in meaningful contexts, implicit knowledge is established and developed. Conscious awareness[59] is no longer necessary for implicit learning or implicit knowledge development. Based on the usage-based theory framework, explicit and implicit

---

[59] Even though N. Ellis (2005) contends that conscious awareness is not necessary for implicit knowledge development, he asserts that attention (i.e., noticing at the attention level) is needed for successful tallying and priming processes required for implicit knowledge development.

knowledge are acquired differently through different developmental routes, and conscious awareness is necessary for explicit but not necessarily for implicit learning.

Ullmann's (2016) declarative/procedural (DP) model seems to share similar views with the usage-based proposition on how explicit and implicit knowledge are distinctively developed. According to Ullmann's (2016) declarative/procedural (DP) model, the two most important memory systems that play a crucial role in language learning and language use are the declarative and procedural memory systems. The declarative memory system is specifically responsible for learning and consolidating new knowledge (i.e., new stimulus) and binding arbitrary bits of these new information together. Learning in this system can occur very rapidly "with as little as a single exposure to the stimulus, although additional exposures strengthen memories" (Ullman, 2016, p. 956). The declarative memory system underlies conscious learning, and explicit knowledge is developed within this memory system (though some implicit knowledge can also be acquired in this system, see the redundancy hypothesis[60] in Ullman, 2016).

On the contrary, the procedural memory system is responsible for learning language rules and categories as well as learning to predict the probabilistic outcomes, "the next item in a sequence or the output of a rule" (Ullman, 2016, p. 956). "Learning in the system requires practice, and thus is slower than learning in declarative memory-though what is eventually learned seems to be processed more rapidly and automatically than knowledge in declarative memory" (Ullman, 2016, p. 956). The procedural memory system underlies unconscious learning (i.e., without awareness), and implicit knowledge is believed to be primarily acquired within this memory system. Based on Ullman's (2016) premise, different types of knowledge are developed distinctively in different memory systems, and while explicit knowledge implicates conscious learning, implicit knowledge can be acquired without awareness.

---

[60] According to Ullman's (2016) redundancy hypothesis, the two memory systems can "acquire the same or analogous knowledge or skills, that is, to play at least partly redundant roles" (p.960), especially for grammar as opposed to lexical or semantics. More specifically, "Given the learning power of declarative memory, and the fact that it can underlie implicit as well as explicit knowledge, it may be able to at least partly support most if not all aspects of language sub-served by procedural memory, including grammar and word segmentation" (p. 960).

Ullman (2016) further suggests that explicit or form-focused instruction (i.e., learners are directed to pay attention to forms or underlying rules and patterns) promotes learning in the declarative memory system. On the other hand,

> "a lack of explicit instruction, as well as manipulations that reduce attention to the stimuli (e.g., in dual-task paradigms), or a high level of complexity of rules or patterns (thus decreasing the subject's ability to explicitly detect patterns) may all shift learning toward procedural memory." (p. 957)

Given that the learners in this study were oriented their attention towards meaning (a highly meaning-focused learning condition) and the feedback provided was not emphasized (i.e., a lack of explicit or form-focused instruction), it was more likely that the implemented meaning-based instructional context helped promote the development of implicit knowledge in learners' procedural memory, rather than explicit knowledge development in their declarative memory, resulting in learners' superior performance on their implicit knowledge tests.

The exit questionnaire data provided evidence in support of the view that the learners in this study may orient themselves more towards meaning (i.e., implicit learning) rather than forms (i.e., explicit learning). Based on the exit questionnaire data, 58% of all learners showed concerns about their content and vocabulary development, while 30% was concerned about the correct use of English grammar. Further, while 56 % of all learners wanted to improve their content and vocabulary, 36% wanted to improve their grammar. Of significance here would be the Research Goal results which showed that only 3 out of 75 learners acknowledged the role of WCF in this present study (the true purpose of the study)[61]. In contrast, 72 others did not seem to acknowledge the given feedback. Most learners seemed to pay little attention to the given WCF while paying most of their attention to develop their written content. The results, thus, implicated that the learners in this study might extend themselves more towards meaning-based learning as claimed.

---

[61] Only three learners stated that they needed to attend to feedback and used feedback to improve their work while the rest believed that the study generally helped them improve their argumentative essay writing.

In addition, the results of the exit questionnaire data also provided further evidence indicating that as a result of implicit learning condition (i.e., meaning-based context), most learners in this study seemed to improve their grammar knowledge without awareness, and this lack of awareness suggests that the knowledge they predominantly acquired may likely be implicit (i.e., unconscious knowledge[62]), rather than explicit knowledge.

Based on the exit questionnaire data, among 73 out of 75 learners who believed they had made some progress during the experiment (two learners did not think they had made any progress), 84% of them reported their progress on written content, or writing in general (e.g., writing faster) including vocabulary usage. Only 16% of the 73 learners reported their progress on grammar even though their grammar test scores indicated significant improvement. Sarah's case is a good example showcasing how learners in this study may be unaware of the knowledge they acquired due to the instructional context. To be more specific, in her response to the question asking if she had made any progress during treatment at all, Sarah from the focused WCF group replied that she did not think she had made any progress during treatment because of the limited exposure to the treatment and limited opportunity for practice. She contended that the 6-week treatment sessions were insufficient for her to make any progress. However, when looking at her test scores over time, the data showed that her OEIT delayed posttest scores (pre: 0; post: 0; delayed: 1), and timed writing delayed posttest scores (pre: 41.5; post: 50; delayed posttest: 57) significantly improved from the pretests. Her explicit knowledge performance, on the other hand, did not exhibit any advancement. In fact, her explicit knowledge test scores showed a slight decrease over time. Given that Sarah did not acknowledge the true purpose of the current study (based on her exit questionnaire response), which targeted English plural usage, it could be assumed that Sarah might somehow improve her plural knowledge unintentionally (i.e., implicitly or incidentally)

---

[62] Implicit learning can also be defined as learning without awareness of the knowledge being acquired and implicit knowledge as the knowledge that learners are not aware of having and use it without awareness (Godfroid, Loewen, Jung, Park, Gass, & R. Ellis, 2015; Reber, 1967; Williams, 2009)

and that explains why she did not recognize that she had made some progress during the experiment. In sum, the results of learners' progress report suggest that most learners were not aware of the knowledge they acquired (they did not think their grammar had improved), and this points to the possibility that the knowledge these learners predominantly acquired was implicit knowledge (i.e., unconscious knowledge).

However, it is also important to note that even though most learners may appear to extend themselves more towards implicit learning (i.e., meaning-focused), it is probable that explicit learning (i.e., form-focused) may also co-occur in this highly meaning-focused context, providing that the learners in this study also made improvements on their explicit knowledge tests. Given that the learners in this study already had some pre-existing explicit knowledge of the target structure, it is possible that while exposing to the given feedback (i.e., target exemplars) and implicitly developed their implicit knowledge, at time the provided WCF may direct learners' attention and raise their conscious awareness towards the target form strengthening their pre-existing explicit knowledge leading to the increment of their explicit knowledge. A number of studies on incidental learning (i.e., implicit learning) also found that implicit learning condition can facilitate learners to develop both their explicit and implicit knowledge of the target linguistic structures (e.g., Rebuschat, 2008; Rebuschat & Williams, 2012; Rogers, Révész, Rebuschat, 2016; Ruiz, Tagarelli, Rebuschat, 2018).

Nonetheless, as emphasized earlier in chapter 4, despite the aforementioned, it is important to acknowledge that it might not be the effect of WCF that helped these learners develop their explicit and implicit knowledge but the opportunity for writing practice or task repetition, at least within this study context. As already elucidated, given that there were no significant differences in the effects of the WCF and content feedback on learners' L2 development, the current finding, to some extent, can also be interpreted as feedback may not be the main factor facilitating learners' L2 knowledge development, but the opportunity for writing practice or task repetition, considering that it seemed to be the only factor (apart from their background knowledge) that the three groups had in common during the experiment. Consequently, any interpretation drawn from the current findings should be exercised with caution, bearing this possibility in mind.

Differential Effects of Different Types of Written Corrective Feedback on L2
Development

The third research question (RQ 3) investigated whether different types of WCF had any differential effects on learners' L2 explicit and implicit knowledge of English plurals. The results of inferential statistic demonstrated no significant differences between the WCF provided for the WCF groups and the content feedback provided for the control group, as all groups made significant progress after treatment. This suggests that both grammar (i.e., direct focused and unfocused WCF) and content feedback provided in this study were equally effective in helping learners improve their explicit and implicit knowledge of English plurals. Three key factors are posited to influence the finding.

The first factor concerns learners' educational context (i.e., learning experiences influenced by educational context). As explained earlier, it is posited that learners' educational context that favors form-focused instruction may facilitate this group of learners to develop a higher level of grammatical sensitivity aptitude compared to learners from other educational setting (EFL vs. ESL). With such a high level of grammatical sensitivity aptitude, even though the control group learners received only content feedback, the effect of feedback combined with their high level of grammatical sensitivity aptitude may help raise their awareness not only towards content development (i.e., meaning) but grammar usage (i.e., forms). The awareness then prompted them to notice the gaps, analyze, and self-correct their own grammatical errors. Given that these learners were intermediate learners and may already have some established English plural knowledge, it was not unexpected that even in the absence of grammar feedback, they were able to self-correct their grammatical errors as long as they noticed and were aware of the errors.

According to Skehan (1998), grammatical sensitivity (i.e., language analytic ability[63]) refers to "the capacity to infer rules of language and make linguistic generalizations and extrapolations" (p. 204). Stefanou and Révész (2015) also define it as an ability "to recognize different syntactic patterns and grammatical functions of words in a given sentence structure, irrespective of knowledge of grammatical terminology"

---

[63] Carroll (1962) combined grammatical sensitivity and inductive learning ability together to form the notion of "language analytic ability".

(p.265). Prior studies (e.g., Benson & DeKeyser, 2019; Erlam, 2005; Sachs, 2012; Skehan 2015, 2016; Stefanou, 2014; Stefanou & Révész, 2015; Yilmaz, 2013) have attested the role of grammar sensitivity in enhancing learners' development of knowledge in different learning contexts. For example, Erlam's (2005) findings on the role of grammatical sensitivity revealed that a high grammatical sensitivity level is especially beneficial in inductive learning (i.e., implicit learning) context where information about the target structure is not explicitly provided or explained, as opposed to deductive learning (i.e., explicit learning) condition where the effect of grammatical sensitivity is neutralized.

In Stefanou and Révész's (2015) study, grammatical sensitivity was a significant moderator of learners' gains within the group that received the less explicit type of WCF. That is, variations in learners' grammatical sensitivity led to variations in performance within the group that received less explicit WCF as opposed to the group that received explicit feedback: learners with higher grammatical sensitivity performed better than those with lower levels of grammatical sensitivity, only when they were given the less explicit type of instruction. The results of Benson and DeKeyser's (2019) study also support these previous findings suggesting that learners' aptitude becomes more important when learners have to rely on themselves to work out grammatical rules (i.e., within implicit learning context). In Skehan's (2015) review of the relationship between aptitude and the efficacy of instruction, Skehan (2015) also remarks that in the learning context where the target grammatical feature is redundant and/or less salient, aptitude may compensate by making that grammatical point become clearer to be noticed. Taken together, it could be hypothesized that the control group learners in this study may benefit probably more from their grammatical sensitivity aptitude than those of the WCF groups whose aptitude might be neutralized by feedback provision[64]. In other words, the control group's grammatical sensitivity may compensate for the lack of explicit grammar feedback. In part, this may explain why the control group learners who received only content feedback could make significant improvement similar to those receiving grammar feedback.

---

[64] The WCF group learners did not have to engage in cognitive comparisons inferring grammatical rules to find corrections for their errors by themselves, and this is posited to reduce their reliance on their grammatical sensitivity aptitude.

The results of the exit questionnaire data provided further evidence in support of the above supposition, revealing that even in the absence of grammar feedback, most control group learners were still concerned about their grammar usage. In fact, the control group's concern over grammar usage was unexpectedly high as while 38% of them worried about their written content, 33% of them demonstrated concern over grammar usage. Precisely, among the three participating groups, the control group demonstrated to have the greatest concern over grammar usage (33%), followed by the unfocused WCF group (32%) and focused WCF group (25.5%), respectively. The control group's greatest concern over grammar usage among the three groups emphasized the fact that despite a lack of grammar feedback provision, grammar remained one of their main focuses. It is also interesting to see that the provision of grammar feedback seemed to help decrease learners' grammatical concern, considering that learners in the WCF treatment groups demonstrated less grammatical concern than those from the control group. In this respect, it is speculated that grammar feedback may make these learners less concerned over their grammar usage because they were assured that they would receive corrections if they made any errors, and that their grammatical errors would be accurately corrected. On the contrary, without grammar feedback, the control group learners concerned if they could accurately correct their own grammatical errors or whether their hypotheses about the target forms were correct and, thus, they developed the sense of insecurity towards their grammar usage.

All in all, it is postulated that the control group learners' high grammaticality sensitivity aptitude assumingly influenced by their educational context might compensate for their lack of grammar feedback and help them achieve higher grammatical scores, similar to those from the WCF treatment groups. However, it should be noted that this proposition is speculative at best since learners' grammatical sensitivity aptitude was not measured in this study. Future research should robustly investigate whether or not grammatical sensitivity plays a more crucial role than other aptitude components in the EFL setting, as opposed to the ESL setting.

In addition, it is worth noting that it is probable that these learners (i.e., English major students) might already have a high level of grammatical sensitivity aptitude to begin with and because of that these learners chose to major in English (they might feel

that they could perform well on the subject). In other words, there is a possibility that educational context might not have any influence on or shape these learners' grammatical sensitivity aptitude as hypothesized. This proposition, whether or not educational context could have shaped learners' grammatical sensitivity aptitude, is, however, still subjected to further investigation.

Apart from learners' educational setting, learners' proficiency level could be another confounding factor moderating the outcome. Generally, considering that the learners in this study were intermediate learners who assumingly already had partial knowledge of English plurals, it could be argued that even in the absence of grammar feedback, they should still be able to notice and self-correct their plural errors to a certain extent. However, since plural errors persist among these learners due to their L1 interference, it may be more difficult for these learners to notice the error in the first place. Without noticing and conscious awareness of the form or error, although the learners have sufficient knowledge to self-correct the error, they might not be able to do so. Therefore, grammar feedback is more helpful in this respect as it could help accentuate the less salient form directing learners' attention and raising their awareness towards the form or persistent error leading to better performance.

Nonetheless, as previously elucidated, for the control group, content feedback might also help raise these learners' grammatical awareness. In part, it was due to their educational setting that trained them to be sensitive towards grammar usage, so when they were prompted by content feedback, instead of focusing only on improving their content, they were reminded to look at their grammar usage as well. Because of the effect of feedback combined with their assumingly high grammatical sensitivity (which might become more useful in the absence of grammar feedback/form-focused instruction), the control group learners were postulated to become more aware of their grammatical errors and usage similar to those who received grammar feedback. With their background knowledge, these learners were then able to self-correct their errors making some grammatical gains, despite a lack of grammar feedback. Thus, learners' proficiency level is another factor attributing to why the control group learners who received only content feedback could make plural improvement similar to those who received grammar feedback. This also explains why content feedback provided in this study can be equally

224

effective as grammar feedback. However, it is speculated that for lower proficient learners and those who do not have some established knowledge of the target structure, receiving or not receiving grammar feedback may lead to discrepant results because without grammar feedback, it is more likely that lower proficient learners might not be able to recognize their own grammatical errors and/or have adequate knowledge to self-correct the errors.

Besides learners' educational setting and proficiency level, task instruction or instructional context is another factor that might influence the current finding. Considering that the study adopted the meaning-based approach to learning (i.e., implicit learning) in which learners were directed to pay attention to meaning rather than forms, and the role of feedback (i.e., form-focused instruction) was minimized or not emphasized, this learning condition might not allow learners from the WCF groups to fully benefit from the WCF provided (i.e., explicit instruction). As Ullman (2016) suggests, "the learning and/or retrieval of knowledge in declarative memory [i.e., through explicit learning] may block (inhibit) the learning and/or retrieval of analogous knowledge in procedural memory [i.e., through implicit learning]. The converse may hold as well" (p. 957). Providing that the learners from the WCF groups did not fully benefit from the given WCF as such, whereas the learners from the control group benefited more from their grammatical sensitivity aptitude due to the instructional context, this may be another reason why grammar and content feedback appeared to be equally effective in helping learners improve their plural knowledge within this study context.

Working Memory Capacity and the Effectiveness of Written Corrective Feedback

The fourth and last research question (RQ 4) of this study explored whether learner differences in WMC mediated the effect of WCF on learners' L2 development. The findings of moderated multiple regression analyses revealed that WMC did not moderate the extent to which the learners in this study benefited from WCF, even though a negative correlation was found between OSpan and UGJT delayed posttest scores of the feedback groups. As previously noted, given that the OSpan scores only predicted one out of eight sets of test scores and the interaction effect was non-significant, the definitive answer regarding the relationship between WMC and WCF could not be borne out based solely on this correlation. Moreover, the qualitative data provide evidence in support of the possibility that learners' low motivation during the delayed posttest might be a key factor influencing the yielded result.

Based on the interview records, only three out of six learners demonstrated motivation evidence during treatment. The other three learners who did not exhibit distinct motivation recounts were speculated to either have fluctuating or seemingly low motivation. Tony from the focused WCF group, for example, commented in one of his interview sessions that "***I had made some improvements but not much as I didn't go back and think about what I could have done better.***" Another example is Jackson from the unfocused WCF group, who stated that "***Sometimes I didn't have a chance to look through all the feedback I received but most of the time I did.***" This implicated that some learners may have possessed a low level of motivation since the beginning of the study.

In addition to the interview data, the data from observation field notes provide further evidence suggesting that learners' motivation may be at the lowest level during the delayed posttest. Considering that up to 15 learners withdrew from the study during the delayed posttest, the high attrition, to some extent, indicated that learners' motivation might plunge dramatically during the delayed posttest. In part, learners' low motivation during the delayed posttest could be a result of the longitudinal nature of the study (9-month period) that required them to undertake the same test battery repeatedly causing boredom among them. Further, the fact that the test results, either good or bad, would not

affect their course grades (i.e., the lack of extrinsic motivation), might, after all, lower learners' motivation to perform their best when they were asked to do these same tests for the third time. The learners who withdrew from the study explained that they would have continued to participate in the project if they could still gain some extra credits from their instructors.[65]

Apart from the high attrition which occurred during the delayed posttest, the field note data also suggested the possibility that the UGJT delayed posttest test scores might be affected by learners' low motivation the most. Based on the field note data, while most learners generally took approximately 1 to 1.30 hours to complete the UGJT pretest and posttest, the learners only took 30-40 minutes to finish the UGJT delayed posttest. In contrast, learners still took approximately the same amount of time when performing other tests. This and the fact that the learners continued to perform significantly better on their OEIT and timed writing delayed posttests, their MKT delayed posttest scores were maintained from the pretest, and only their UGJT delayed posttest scores showed a significant decline from the pretest, suggest a high probability that most learners might pay the least attention when taking the UGJT delayed posttest. Consequently, it is possible that learners' low motivation, especially during the UGJT may affect their UGJT delayed posttest scores; therefore, their UGJT scores may not well reflect the true ability of the learners. Since the UGJT delayed posttest scores might be more deviant than other sets of test scores due to learners' low motivation, this poses questions about the reliability and legitimacy of the detected negative association. This relationship, thus, remains open for further investigation.

Regarding the null relationship between WMC and WCF, three factors are posited to neutralize the WMC effect within this study context (see the discussion section of chapter 4). The first factor concerns the type of WCF provided in this study. As previously discussed, because of the explicit nature of direct WCF, learners can notice and internalize the feedback without much effort as they did not have to engage in the cognitive comparisons to find solutions to their errors. Thus, noticing and processing this

---

[65] During the first 5 months of the study (from pretest to posttest), learners received extra credits from the instructors if they could complete all the required experiment sessions. Nonetheless, the learners did not receive any extra credits during the delayed posttest.

type of feedback were not cognitively demanding, and learners' working memory (e.g., attentional control function, phonological short-term memory) was not heavily taxed on. Accordingly, learners with high or lower WMC can equally benefit from this type of explicit feedback. On the contrary, it is hypothesized that the processing of implicit feedback[66] may be more cognitively demanding since learners have to find the solutions to their errors by themselves; thus, learners have to rely heavily on their working memory resources. In this case, learners with higher WMC (i.e., having higher ability to control and regulate their attentional resources or having larger phonological memory space) may benefit more from implicit type of feedback, as they may be better at allocating their attention to the feedback and holding the feedback (long enough) for their hypothesis testing as well as for the development of their metalinguistic understanding. Thus, it could be concluded that the explicit nature of the WCF provided in this study (i.e., direct WCF) might neutralize the effect of WMC on WCF.

To this point, one might raise the question as to why WMC did not have any effect on WCF but oral feedback, given that both are form-focused tools that help raise learners' attention and awareness towards forms. A number of oral feedback studies also found that learners' WMC moderated the extent to which learners can benefit from oral feedback. In part, this could attribute to the less explicit and fleeting nature of oral feedback. Given that oral feedback is less explicit and fleeting in nature, that is, learners may not know that they are being corrected and oral feedback is only provided to learners at the exact moment of the negotiation for meaning; therefore, learners do not have a chance to discuss the feedback later, unlike the explicit and permanence nature of WCF (i.e., learners know that they are being corrected and they can go back to look at the feedback later), to notice and process oral feedback might be more cognitively demanding than when processing WCF. Consequently, learners might need to rely more on their working memory resources when processing oral as opposed to written feedback. This may explain why WMC was found to moderate the effectiveness of oral feedback but was not found to have any effect on the processing of WCF in this study.

---

[66] Implicit feedback in this writing context referred to indirect WCF, the feedback that locates the errors for learners but does not provide them corrections.

Apart from the explicitness of WCF as elucidated, the second factor postulated to neutralize the moderating effect of WMC is learners' instructional context. As previously discussed, it is posited that meaning-based (i.e., implicit) instructional context which predisposed learners towards implicit learning reduced learners' reliance on conscious working memory operations. As most scholars suggest, working memory plays the less important role in implicit learning and processing but plays an integral part in explicit learning and processing (e.g., N. Ellis, 2005; R. Ellis, 2009c; Skehan, 2016). Given that the learners in this study seemed to orient themselves more towards implicit learning, which to the large part did not involve working memory operations, variations in WMC did not seem to have any influence here. The null WMC effect found in this study is also in line with the results of a number of previous studies which detected no WMC effect in implicit learning condition (e.g., Ando et al., 1992; Hamrick, 2015; McDonough & Trofimovich, 2016; Ruiz et al., 2019; Tagarelli, Mota, & Rebuschat, 2015; Walker et al., 2020). The result also corroborates the claim that learners' instructional context mediates the effect of WMC on L2 learning outcomes (e.g., Ando et al., 1992; Faretta-Stutenberg & Morgan-Short, 2018; Ruiz et al., 2019; Suzuki & DeKeyser, 2017; Tagarelli et al., 2015; Walker et al., 2020; Yang et al., 2017).

The last factor posited to mediate the effect of WMC is learners' proficiency level. As many SLA researchers have pointed out, lower proficient learners or learners at the beginning stage of L2 learning tend to consume more on their WMC than higher proficient learners (Harrington & Sawyer, 1992; Mitchell et al., 2015; Roberts, 2012; Sagarra, 2013, 2017; Serafini & Sanz, 2016).

This claim is also supported by skill acquisition theory tenets which suggest the more important role of explicit knowledge (i.e., declarative and procedural knowledge) during the beginning stage of language learning. Since for beginners or lower proficient L2 learners, their implicit knowledge is still very limited or undeveloped, and learners still need time to analyze new information in order to encode and/or consolidate new knowledge (involve conscious operations), these learners then can only largely rely on their explicit (i.e., conscious) knowledge. Given that explicit knowledge is processed through the explicit learning system (i.e., declarative memory) which to the large part implicates working memory operations, it can be said that learners at this stage greatly

rely on working memory resources. However, once learners have developed more L2 experiences, that is becoming more proficient in their L2, learners gradually develop and rely more on their automatized/implicit knowledge (Paradis, 2009; Suzuki & DeKeyser, 2015, 2017), which is processed and retrieved through the implicit learning system (i.e., procedural memory) which to the most part does not entail working memory operations (N. Ellis, 2005; R. Ellis, 2009b; Skehan, 2016, also see Buffington & Morgan-Short, 2019, for a comparison review of skill acquisition theory with declarative/procedural model). Consequently, providing that the learners in this study were at intermediate level and had already developed some established L2 knowledge, it is then probable that these learners may not largely resort to their explicit knowledge, which involves working memory operations, but turn to rely on other cognitive resources, i.e., implicit knowledge, when processing input or feedback. Accordingly, their WMC was not heavily taxed on, compared to that of the lower proficient learners, and in part, this may attribute to the lack of working memory effect within this study context.

Similar to this study's finding, a number of previous studies (e.g., Gilabert & Muñoz, 2010; Grey et al., 2015; Serafini & Sanz, 2016) also did not find any significant relationships between intermediate learners' WMC and their morphosyntactic performance (i.e., grammar learning). This, thus, further corroborates the claim that the impact of working memory is lessened as learners' proficiency level increases (Sagarra, 2013, 2017; Serafini & Sanz, 2016).

6.2. Implications

In the subsequent sections, theoretical and pedagogical implications drawing from the main findings are put forward.

6.2.1. Theoretical implications

The findings of the current investigation yielded several theoretical implications. First, the results provide evidence in support of the role of WCF, a form of negative evidence, in promoting learners' explicit and implicit knowledge of the non-salient linguistic features, especially the ones impacted by learners' L1 interference. Theoretically, the results suggest that negative evidence in the form of corrective feedback is truly beneficial for L2 learning as it could help learners further develop their L2 knowledge. The results also implicate that positive evidence alone may not be sufficient for L2 learners to improve certain aspects of L2 learning, especially the ones influenced or blocked by learners' L1 cues, because if positive evidence alone is adequate for L2 development, the learners in this study should be able to successfully acquire this plural structure after having learnt it for years. Thus, the current findings corroborate previous proposition contending that the provision of positive evidence alone is not adequate for L2 acquisition, and that negative evidence is deemed essential (e.g., Abbuhl & Gass, 2013; Anderson, 1981, 1982, 1993, 2007; Anderson & Fincham, 1994; Bitchener & Ferris, 2012; DeKeyser, 1994, 2007, 2015; R. Ellis, 2010; Gass, 1997, 2003, 2015; Gass & Mackey, 2015; Hulstijn & Schmidt, 1994; Long, 1996, 2007, 2014; Swain, 1995, 2005; Van Beuningen, 2010). The findings are also in accordance with the interactionist, skill acquisition, and usage-based framework, proposing that negative evidence in the form of corrective feedback is beneficial not only for the development of L2 explicit knowledge but also implicit knowledge (Abbuhl et al., 2018; Anderson, 1981, 1982, 1993, 2007; Anderson & Fincham, 1994; Biber et al., 2011; Bitchener & Storch, 2016; Brown, 2016; DeKeyser, 1994, 2007, 2015; N. Ellis, 2005; 2015b; R. Ellis, 2010; Gass, 1997, 2003, 2015; Gass & Mackey, 2015; Goldberg, 1995; Goo & Mackey, 2013; Han, 2002; Hulstijn & Schmidt, 1994; Kang & Han, 2015; Lee, 2019; Li, 2010; Li & Vuono,

2019; Long, 1996, 2007, 2014; Mackey et al., 2013; Mao & Lee, 2020; Robinson & N. Ellis, 2008; Russell & Spada, 2006; Van Beuningen, 2010). Hence, the findings refute Truscott's (1996, 1999, 2004, 2007, 2010) allegation regarding the ineffectiveness and harmfulness of WCF in L2 classrooms.

In addition to the beneficial effect of WCF, one form of negative evidence, on learners' L2 development, the study asserts that noticing is an essential construct for L2 development, especially for the development of the non-salient linguistic features influenced by learners' L1 interference or transfer, such as the English plurality examined in this study. The results of the qualitative data analysis substantiate Schmidt's Noticing Hypothesis (2001), which contends that "since many features of L2 input are likely to be infrequent, non-salient, and communicatively redundant, intentionally focused attention may be a practical (though not theoretical) necessity for successful language learning" (p.23). The study's proposition is also in line with N. Ellis' (2005) suggestion which indicates that in the situation where the target feature is low salient and influenced by learners' L1 cues and "we describe the learner as having *fossilized* or more correctly *stabilized*, the remedy is to bring the issue into the light of consciousness" (p. 324). In this case, feedback can be used as a consciousness-raising device enabling learners to efficiently notice the low salient cue and the mismatch between their current L2 state and the target-like production; thereby, prompting the destabilization and reconstruction of their interlanguage grammar, leading to successful L2 acquisition (Bitchener & Storch, 2016; Gass, 1997, 2003; Long, 1996, 2007, 2014; Polio, 2012; Van Beuningen, 2010, Williams, 2012).

Relatedly, the study proposes that learning without awareness[67] is possible considering that most learners in this study were not aware of their grammatical improvement (see previous sections). Nevertheless, as Rebuschat et al. (2015) comment, a lack of awareness evidence could also be a result of the fact that "awareness may happen more quickly than concurrent verbalization allows expression of" (p. 303), or

---

[67] In this study, learning without awareness means learners did not have intention to learn something (i.e., learning from feedback) to begin with, so they were not aware of what they had been learned, and, therefore, they were unaware of the knowledge they had acquired (Rebuschat, 2013, 2015, also see accidental learning). In short, learning occurred without learners being aware of what they had been learned.

awareness is already decayed from learners' memory when they have to report it. In other words, a lack of evidence of awareness does not necessarily mean a lack of awareness. Rebuschat et al. (2015) also argue that, in part, it could be the ineffectiveness of the measurement of awareness that attributes to a lack of awareness evidence. Consequently, a lack of awareness among learners detected in this study could also be a result of the aforementioned factor. However, note that a lack of awareness of what being learned described above does not equate to or implicate a lack of noticing. In other words, the study does not propose that learning can take place without noticing (at the attention level). Instead, in line with N. Ellis' (2005) argument, the study contends that even though learners can still successfully acquire L2 knowledge without awareness (through implicit or accidental learning as opposed to conscious or explicit learning)[68], noticing or attention is unquestionably essential construct for successful L2 development.

In addition to the aforementioned implications, the results of the study lend support to the widely accepted notion in SLA suggesting that explicit and implicit knowledge are distinctively developed (Andringa & Rebuschat, 2015; N. Ellis, 2005, 2012, 2015b; 2016a; Suzuki & DeKeyser, 2017; Paradis, 2009; Wulff & N. Ellis, 2018; Ullman, 2005, 2015; 2016; Ullman & Lovelett, 2018) and that they are fundamentally distinguished by awareness criterion (Andringa & Rebuschat, 2015; N. Ellis, 2005; R. Ellis et al., 2009; Suzuki & DeKeyser, 2017). As N. Ellis (2005) explains, the development of explicit knowledge pertains to conscious registration of linguistic features or patterns, which are later bound together to establish a coherent representation of the knowledge. Conscious awareness then plays a significant role in explicit knowledge development. However, the development of implicit knowledge begins with the integration of "new information with existing knowledge, using overlapping distributed representations to extract the general statistical structure of the environment" (N. Ellis, 2005, p.319). Conscious awareness is no longer necessary for implicit knowledge

---

[68] The researcher believes that most implicit learning processes, e.g., frequency stocking, exemplar tallying, etc., occur unconsciously (i.e., implicitly or without awareness) as we might not be able to consciously count the amount of input we tally into our learning systems, for example. However, the researcher also believes that "input to the associative network is gated by consciousness" (p.248, N. Ellis, 2016), that is, consciousness (i.e., attention) is a gateway to learning regardless.

development. Ullman (2016) also contends that implicit or meaning-focused instruction promotes unconscious (i.e., implicit) learning leading to the development of learners' implicit or unconscious knowledge. On the other hand, explicit or form-focused instruction facilitates conscious (i.e., explicit) learning, which aids the development of explicit or conscious knowledge.

Given that the learners in this study received meaning-focused instruction and performed significantly better on their implicit as opposed to explicit knowledge tests, and most of them were unaware of their grammatical improvement indicating learning without awareness, the findings substantiate the above proposition that the two types of knowledge have developmental processes that are independent to each other and that conscious awareness is a criterion distinguishing explicit from implicit knowledge. Moreover, as the two do not share the same developmental route (conscious vs. unconscious learning), the study corroborates the theoretical claim that explicit knowledge will never become implicit knowledge and vice versa. Yet, it is posited that there might be an interface between the two types of knowledge, given that the implicit learning condition provided in this study did not only lead to the development of implicit knowledge but explicit knowledge. In other words, the study proposes that the development of implicit knowledge might influence explicit knowledge advancement. That is, the increment in learners' implicit knowledge may, to some extent, facilitate the encoding and consolidation of explicit knowledge as learners can extract, analyze (hypothesis testing) or reanalyze, generalize the underlying rules from the accumulated exemplars already tallied in the implicit learning system and then contribute them to their explicit knowledge repertoire. Even though the study cannot provide direct evidence in support of the interface position, the increment of explicit or conscious knowledge as a result of implicit or unconscious learning indicates a high probability that the two types of knowledge might interact in a bidirectional direction. Nonetheless, more evidence is unquestionably needed to make a claim more affirmative.

Last, the study corroborates previous claims stating that instructional contexts and learners' proficiency levels are two key factors significantly mediating the extent to which learners can benefit from WMC (e.g., Adams & Mohammadtaghi, 2014; Ando et al., 1992; Faretta-Stutenberg & Morgan-Short, 2018; Harrington & Sawyer, 1992; Kang

& Han, 2015; Linck et al., 2014; Mitchell, Jarvis, O'Malley, & Konstantinova, 2015; Ruiz et al., 2019; Sagarra, 2013, 2017; Serafini & Sanz, 2016; Suzuki & DeKeyser, 2017; Tagarelli et al., 2015; Yang et al., 2017). Specifically, the study proposes that in the implicit learning condition, variations in WMC do not seem to affect learners' learning differently since the instructional context promotes implicit learning, which, for the most part, does not implicate working memory operations. In addition, the impact of working memory effect might be lessened as learners' proficiency level increases. That is, the higher the learners' proficiency levels, the higher probability the learners rely less on their explicit knowledge (i.e., conscious processing, involving control processing and working memory operations) but their implicit knowledge (i.e., automatic or unconscious processing). This is because high proficient learners could get access to their implicit knowledge faster than those at the lower levels who have lesser quality and quantity of implicit knowledge and, therefore, need to rely merely on their explicit knowledge.

6.2.2. Pedagogical implications

In addition to the theoretical implications, the current analysis also yielded a number of pivotal pedagogical implications that could be especially beneficial for L2 practitioners. First of all, the current findings warrant WCF, both direct *"focused"* and direct *"unfocused"* WCF as an effective pedagogical tool helping learners develop both their explicit and implicit L2 knowledge of the non-salient form influenced by learners' L1. As aforementioned, WCF can efficiently serve as an awareness-raising tool inducing noticing at both the attention and understanding levels facilitating learners' L2 development. Even though content feedback and the presumed high grammatical sensitivity aptitude had also shown to aid the development of the control group learners' plural knowledge similar to the provided WCF, it is postulated that if the treatment is prolonged, differences in the effects of the WCF and content feedback may have emerged.

Providing that the two types of WCF are both beneficial for L2 development, a question may arise as to which one between the two is the most effective in aiding L2 learning and should be adopted in L2 classrooms. In this respect, L2 practitioners need to

take learners' proficiency levels into consideration when determining which one of the two WCF is more corresponding to their learners' needs.

For learners at a high proficiency level (i.e., intermediate and advanced learners), who already have partial knowledge of the target linguistic structure(s) similar to the learners in this study, both types of WCF seem to be equally effective and practically either can be appropriately used in class. However, it is speculated that in the context where there are some persistent errors, direct *"focused"* WCF may be a better option, at least at the beginning, as learners only have to pay attention to the problematic forms and the practice on a few types of errors may better accelerate the acquisition of the target structures.

Nevertheless, for beginner learners, it is speculated that direct *"focused"* WCF may be more facilitative than direct *"unfocused"* WCF because the *"unfocused"* WCF can be overwhelming. Also, due to their limited L2 knowledge, beginner learners may not understand all the corrections given to them, and it is a waste of the instructor's time and effort to correct all types of grammatical errors for learners at this stage. It is suggested that L2 practitioners may provide beginner learners with direct *"focused"* WCF first (targeting one or few types of grammatical errors) and provide them with direct *"unfocused"* WCF when they have more L2 knowledge or show developmental readiness. In sum, for high proficient learners, both types of WCF are equally effective and either can be appropriately used in class. However, for beginner learners, direct *"focused"* WCF seems to be a more practical option at least at the beginning. Once learners acquire more L2 knowledge, direct *"unfocused"* WCF can be used in place of direct *"focused"* WCF.

Apart from the merits of WCF in L2 development, the study suggests that learners' educational contexts, i.e., ESL or EFL setting, should also be taken into account when designing classroom instruction. In other words, whether the instruction should extend more towards meaning or forms, this also depends on learners' educational context. For example, in an ESL setting where L2 learners expose to the target input both in and outside of class and most of them are already fluent communicators of the target language community, more form-focused instruction may be more beneficial. This is because while most of these learners can communicate fluently, they may overlook or

have problems with their grammatical usage since accuracy is less important than the ability to communicate fluently in their everyday life. Hence, in this case, explicit learning strategy is more helpful as it helps highlight the grammatical aspects they need to focus on or improve. The WCF role in the ESL classroom should also be explicitly emphasized as WCF can effectively draw learners' attention to problematic grammatical forms helping them understand and improve their accuracy better and faster.

On the other hand, in an EFL setting where learners have limited exposure to the target input[69], the use of explicit or implicit instruction has to be proportional. Both explicit and implicit instructions are equally important for EFL learners, as only through implicit instruction alone, the acquisition may take forever. In contrast, a more form-focused instruction may inhibit them from becoming fluent L2 users. Therefore, for EFL learners, the proportionate combination of form and meaning-focused instruction is preferable. In effect, it is suggested that for lower proficient EFL learners, a more meaning-focused instruction can be introduced first. WCF should still be provided but may not necessarily be emphasized. However, when learners become more fluent, the proportionate combination of form and meaning-focused instruction can be implemented, and the role of WCF can be more emphasized in class.

The last pedagogical suggestion of this study entails the potential role of WCF as an effective motivator. As the findings revealed, learners' motivation seemed to influence the current findings to some extent (especially during the administration of the delayed posttest), and this emphasizes the crucial role motivation plays in L2 learning. L2 practitioners, thus, need to find teaching strategies that not only suit their learners' needs but also motivate them to engage or invest more in their learning.

According to Bandura's (1986) self-efficacy belief theory, when a person has a positive belief that he/she has an ability to perform the task, this positive view motivates and enhances his/her performance leading to the successful accomplishment of his/her determined goal. The motivation self-system by Dörnyei (2005) also proposes that

---

[69] It is true that with the technological advancement, EFL learners have more channels to be exposed to the target input. Yet, it is unlikely that all of them are going to invest that much outside of class time except for the motivated ones. Even if they do, they still have less exposure to the target input compared to ESL learners who have to use the target language in their everyday life.

motivated behavior is directed by a positive view towards a goal a person aims to achieve or an activity a person is participating. That is, a person needs to first foresee the benefit of doing something or believe that the goal is achievable before they feel the need to pursue the course and invest in that action or activity. In short, motivation is primarily established based on a person's positive view and confidence towards his/her ability to achieve. Consequently, to motivate learners means to make them feel positive about themselves or the activities they are engaged in.

Based on the study's qualitative data, all learners in this study demonstrated to have positive views towards the provision of feedback. They indicated that WCF is an effective and useful tool that can help them improve the quality of their L2 writing and as a result, helping them become a confident L2 writer. These learners also expressed that it is necessary that they receive teachers' feedback. Without feedback, they are not confident if they have made any progress in their learning, and this could make them feel lost or even demotivated. Given that most learners in this study had positive views towards feedback provision, feedback can be perceived as an effective motivator that could help stimulate or increase learners' motivation to learn or invest in their learning, particularly in composition classrooms. Thus, the provision of feedback should not be viewed as a teaching option.

For composition classrooms, it is recommended that teachers provide WCF for every writing occasion. In addition, based on learners' comments, direct WCF seemed to be more favorable for those learners who are at a lower proficiency level or those who are not confident or fluent in their writing just yet, even though they are already at a higher proficiency level (similar to this study). For those at a more advanced level, it is postulated that indirect WCF, the feedback that locates the error for learners but correction is not provided (did not offer in this study), may be more appropriate considering that this group of learners already have ample knowledge necessary for their own hypothesis testing and correction.

For other L2 classes, teachers are also encouraged to provide learners with appropriate feedback or comments that could give learners clear guidance about learning aspects they need to improve and, from time to time, give them feedback that indicates their learning progress. It is undeniable that the provision of feedback is time-consuming.

However, considering that teachers' feedback could potentially serve as an effective motivator encouraging learners to invest more in their learning leading to their learning success, it is worthwhile for teachers to consistently provide feedback in their classrooms.

6.3. Limitations and future research

In this study, the effectiveness of direct *"focused"* and direct *"unfocused"* WCF was explored. The findings affirm the effectiveness of these two types of WCF in promoting the development of learners' explicit and implicit knowledge of English plurals, even though the differential effects between the two WCF and the content feedback was not detected. The results highlight the advantages of providing WCF to L2 learners and the role of negative input plays in L2 acquisition. In addition, the current findings corroborate previous claims contending that explicit and implicit knowledge develop independently of each other and that the two types of knowledge interact. The effects of learner differences in WMC on the efficacy of WCF were also investigated in this study. The results showed that variations in WMC did not affect the extent to which learners benefited from WCF. It is postulated that the explicitness of direct WCF adopted in this study, instructional context, and learners' proficiency moderated the facilitative effect of WMC. In particular, the findings suggest that WMC might play a less significant role in implicit learning condition, and high proficient L2 learners might not largely rely on WMC during input processing but other cognitive resources. Also, the explicitness or implicitness of the processed input (i.e., feedback) might mediate the effect of WMC to a certain extent. All in all, the findings generate more insights into the potential role of WCF in L2 development and the association between learners' WMC and the observed effectiveness of WCF.

Notwithstanding, despite the said contributions, some limitations of this study must be acknowledged. As delineated in previous chapters, one main limitation of this study concerns the small sample size resulting from high attrition, which led to insufficient statistical power to detect all the existing effects (see the discussion section of chapter 4). In part, this issue could have been prevented by conducting power analysis before the experiment. This may have impacted recruitment efforts (i.e., starting with a

larger sample) to help further offset the attrition, which is inevitable due to the nature of a longitudinal study. In addition, for a longitudinal study, it is critical to keep learners motivated throughout their participation, as demotivated behavior can bias the test results and lead to high attrition as shown in this study. Thus, future studies with a longitudinal nature similar to this study may need a larger sample size group for the results to be more robust. Researchers may also have to find a better way to keep learners incentivized and motivated during their participation.

A second limitation of this study pertains to the short treatment period (6 weeks), which assumingly led to the less discernable progress between the grammar and content feedback groups. As explained earlier, considering that the learners were at the intermediate level, which seemed to leave them smaller room for improvement, a longer treatment period may be more appropriate to help them to make more progress than the short treatment provided in this study. Especially when the recurrent errors are influenced by learners' L1, a longer treatment period seems to be a more appropriate approach. Thus, future research may need to incorporate a longer treatment phase, especially if the study aims for those at a higher proficiency level (i.e., intermediate and advanced levels) or when the recurrent errors are influenced by learners' L1. It is also advantageous that future replications include learners from other proficiency levels (i.e., beginner and advanced levels) and compare them to this group of intermediate learners since differences in proficiency might interact differently with different types of WCF or learners' WMC.

Third, because the study adopted the experimental design, which allowed for the efficacious control over potential confounding variables, the yielded results, to some extent, might not fully reflect or predict learners' real-world behaviors. As already discussed in chapter 5, in a real classroom setting, it is impossible to control all variables that may affect learners' learning outcomes. Learners' motivation and engagement levels within a controlled setting and a real classroom setting are also different. Accordingly, the results obtained from a real classroom setting may be different from the ones obtained in this study. It is, thus, worth investigating further whether within a less controlled environment, similar findings can be obtained so that the current findings can be affirmed.

Fourth, given that the learners in this study were directed to pay more attention to meaning rather than forms and the role of feedback was not emphasized during treatment, the instructional context may have influenced the yielded results. As mentioned earlier, the highly meaning-focused instructional context and the less focus on the given WCF may have pushed learners to rely more on their implicit rather than explicit learning systems; therefore, the learners in this study might not fully benefit from the provided WCF. In part, this may have led to the null differential effects between the WCF and content feedback found in this study. However, this premise is speculative. Future replications should adopt a similar instructional condition but the WCF role needs to be emphasized so that the results could shed more lights on the relationship between instructional context and the observed effectiveness of WCF.

The fifth limitation lies in the fact that there was a lack of a no-feedback exposure group in this study. This may raise a question as to whether learners who receive feedback could actually perform better than those who do not receive any. Indeed, it would be ideal to have a complete no-feedback exposure group so that the results could be more affirmativv. However, considering the longitudinal nature of the current study, if some learners did not receive any feedback during treatment at all, this might greatly impact learners' motivation and participation, potentially lead to a higher attrition rate, as most learners in this study expected to receive teachers' feedback (see Chen, Nassaji & Liu, 2016, and Mao & Lee, 2020, for similar arguments on the EFL learners' beliefs about feedback). Also, as the learners in this study were randomly assigned, meaning there would be learners from both the control and the two WCF groups practicing together in one intact class, if the control group learners noticed that they did not receive any feedback at all for six weeks straight, this might reveal the true purpose of the project adding more confounding variables to the analysis. For this reason, content feedback was provided for the control group learners.

In addition, a number of previous WCF studies, which adopted a no-feedback exposure group design (e.g., Bonilla et al., 2018; Frear & Chiu, 2015; Kim et al., 2020; Kurzer, 2018; Shintani & R. Ellis, 2013; Shintani et al., 2014; Van Beuningen et al., 2012; Wagner & Wulf, 2016), have already provided robust evidence in support of the proposition that it is more beneficial providing learners with WCF than not providing

them any. As shown in these studies, learners who received WCF performed significantly better than those from the non-feedback condition. Based on these studies' findings, it is quite affirmative in the WCF field that providing learners with WCF is more advantageous. Therefore, even though the current study did not have a complete no-feedback exposure group to be compared with, it can be hypothesized based on previous findings that learners who receive feedback might perform better than those who do not receive any. In essence, although the current study cannot provide direct evidence attesting the claim that providing feedback is better than not providing any, the observed effectiveness of WCF in promoting learners' L2 acquisition detected in this study warrants the positive role WCF plays in learners' L2 development regardless. However, it would be beneficial for future research with available resources to attempt a longitudinal study with a complete no-feedback exposure group.

Another limitation of this study involves the implicit knowledge measurements employed in this study. As mentioned earlier in chapter 4, some researchers may argue that the timed writing task used in this study is unlikely a valid measure of implicit knowledge. It is as Polio (2012) and Shintani and R. Ellis (2013) point out that even when the writing task is under time pressure and learners' attention is predominantly placed on meaning, learners are still able to monitor their grammatical usage, that is, the task still allows the learners to deploy their explicit knowledge. In addition, even though the timed OEIT used in this study is widely accepted as one of the most valid measures of implicit knowledge available to date (apart from the reaction time methods such as the word-monitoring task, the self-paced reading task, and the visual-world task[70]), the task appears to lack face validity when used in WCF studies, given that the main aim of all WCF studies is to improve learners' writing abilities (Shintani & R. Ellis, 2013). However, as already discussed (also see the discussion section of chapter 4), given that there is no pure measurement of implicit knowledge available to date, the best method that could be used to compensate for the lack of pure implicit knowledge measurement is to employ multiple measurements, instead of one, in a single study, so that data from various sources can be cross-checked yielding the more robust outcome. Nonetheless, it is worth noting that even

---

[70] These measures of implicit knowledge are also under development and still in need of further validation (Suzuki & DeKeyser, 2017).

though the current study adopted multiple measurements to compensate for a lack of pure implicit knowledge measurement as described, providing that the validity of the timed writing task is still subject to debate, any interpretation drawn from the current findings still has to be exercised with this caveat in mind.

Apart from the implicit knowledge measurement issue, another limitation of this study pertains to learners' self-report data derived from the exit questionnaire and qualitative interview. Even though the questionnaire and qualitative interview are the useful research tools that could help researchers obtain more insights about learners' thoughts and/or thinking processes, the data obtained from these methods may not always well reflect learners' perception as anticipated. There are many factors that could affect the reliability of the self-report data such as learners' memory decay as learners might already forget what they actually think or proceed back at the time so the data might not accurately represent what really is on their minds at that particular time, or learners may not want to tell the researcher the truth about how they feel because they are afraid to offend the researcher, etc. Besides, especially for the interview, it is possible that by participating in the interview (particularly the introspective one) like the one administered in the study, the interview learners might have learnt about the true purpose of the study and this might influence their post-treatment performance[71]. Also, sometimes the interview questions can be misleading creating fault memories among learners, and, as a result, learners might end up providing inaccurate data to the researcher. Consequently, while the self-report data from the interview and questionnaire are helpful for in-depth analyses and that explains why the current study employed these two qualitative methods in addition to the quantitative ones to begin with, the self-report data also has its own limitation and any conclusion drawn from it should be exercised with this caveat in mind.

The last limitation of this study has to do with the fact that the current study only investigated the effectiveness of WCF on a single linguistic feature, English plurals.

---

[71] Being aware of this possibility, the researcher rechecked the six interview learners' exit questionnaire responses. However, none of these learners seem to recognize that the study aimed to assess their learning from feedback. In part, this might be because of the short intervention period. It is possible that if the treatment is prolonged, the interview learners might be able to recognize the goal of the study especially after several more interview sessions.

Thus, the effectiveness of WCF on English plurals may not be generalizable to other linguistic categories. At best, the results could be generalizable to other simple rule-based features but not to those idiosyncratic features like prepositions for example. Previous feedback studies (e.g., Adams et al., 2011; Benson & DeKeyser, 2019; Bonilla et al., 2018; Frear et al., 2015; Shintani et al., 2014; Wagner & Wulf, 2016) suggest the high possibility that different types of WCF may interact differently with different categories of linguistic errors. Likewise, providing that this project focused only on EFL learners who had limited exposure to the target input, were familiar with the provision of WCF (i.e., grammar correction), and seemed to have high grammatical sensitivity aptitude due to their ample experiences with the form-focused instruction, the current findings might not be generalizable to those ESL learners who have different L2 learning experiences and reside in a different learning environment.

Notwithstanding, despite the above reservations, this study contributes to the fuller understanding of the role of WCF in L2 development as well as the relationship between WMC, WCF, and L2 learning. In what follows, some avenues for future research are proposed.

First, even though the WCF provided in this study demonstrated to have a positive effect in addressing learners' plural errors, to the best of my knowledge, this study is the only WCF study specifically examining the effect of WCF on this particular feature; therefore, more studies targeting English plurals are necessitated to make the finding conclusive. In particular, it will be more contributive if future replication is conducted on a group of learners whose L1 does not have plural inflectional morphemes similar to this study context. The study also suggests that future WCF studies focus more on a wider range of linguistic features, apart from articles, past tense, prepositions, subject-verb agreement, etc., which are usually targeted in most WCF studies, because different error categories might differently respond to different WCF types. Also, it will be more interesting to investigate the effectiveness of WCF on learners' recurrent errors resulting from learners' L1 transfer. The results would be worthwhile for the development of L2 theory and pedagogy if WCF could actually help learners overcome their L1 barrier.

In addition, the replication of the current study, exploring the differential effects of direct *"focused"* and direct *"unfocused"* WCF, is needed. As delineated, the

comparison of these two types of feedback in one single study is limited. To date, there are less than ten studies including the current project that examined the effects of these two feedback types in a single study, and the results are mixed. Moreover, none of these previous studies investigated the differential effects of direct *"focused"* and direct *"unfocused"* WCF in relation to learners' L2 development measured via explicit and implicit knowledge measurements. Hence, future replications are encouraged to examine the effectiveness of these two feedback types on learners' L2 development using explicit and implicit knowledge measurements so that the effectiveness of WCF could be fully examined.

Another avenue for future study entails the role of learner differences in grammatical sensitivity aptitude that needs to be further researched. As exhibited in this study, grammatical sensitivity is speculated to have some contributions to the progress of the control group learners. Based on the yielded results, it is postulated that grammatical sensitivity may play a more significant role in the EFL than ESL setting given that EFL learners might develop this aptitude factor due to their long-term exposure to form-focused instruction and, as a result, the learners might have a high tendency to rely more on this aptitude factor. Even though a number of prior studies have investigated the relationships between grammatical sensitivity and other types of instructional treatments, this line of research is still very limited within the WCF field. To the best of my knowledge, only a few studies (i.e., Benson & DeKeyser, 2019; Sheen, 2007; Shintani & R. Ellis, 2015; Stefanou & Révész, 2015) had investigated the impact of learners' grammatical sensitivity on the effectiveness of WCF to date. Thus, the role of grammatical sensitivity in relation to the effectiveness of WCF remains open for further investigation.

Finally, this study suggests that further replications concerning the relationship between WMC, WCF, and learners' gains should be pursued to confirm the current findings. In addition, as the current findings implicated that learners' proficiency level might moderate the extent to which learners benefit from WMC, future replications should attempt to include learners from different proficiency levels (i.e., beginner and advanced learners) in one single study so that the results could be compared across groups. Moreover, providing that the current study only employed complex working

memory measures, it might be more beneficial that future replications adopt both phonological working memory (PWM) and complex working memory measures so that the results could be compared, given that numerous WMC studies have also reported relationships between PWM and different aspects of L2 learning (e.g., N. Ellis & Sinclair, 1996; Mackey et al., 2002; Mackey & Sachs, 2012; Masoura & Gathercole, 1999; O'Brien, Segalowitz, Collentine, & Freed, 2006; Révész, 2012; Speciale, N. Ellis & Bywater, 2004; Trofimovich, Amnar & Gatbonton, 2007; Williams & Lovatt, 2003). Also, as N. Ellis (2005) has suggested, different components of working memory may differently interact with different learning aspects or conditions. That is, PWM may be more involved with learners' ability to retain feedback for future use (also see Alison et al., 2010), whereas the executive function of working memory (i.e., complex working memory) may be "more associated with explicit learning…and the production of output" (p. 339).

In conclusion, despite the study's limitations that may limit the generalizability of the findings, the current study confirms the efficacy of WCF in promoting learners' L2 acquisition (at least on English plurals) and proposes that the explicitness of feedback, instructional context as well as learners' proficiency level moderate the effect of WMC on learners' L2 learning. Future research should attempt to generate more insights into the role of WCF plays in instructed SLA and its relationship to learners' variations in WMC.

REFERENCES

Abbuhl, R., Ziegler, N., Mackey, A., & Amoroso, L. (2018). Interaction and learning grammar. In J. Liontas & M. DelliCarpini (Eds.), *The TESOL Encyclopedia of English Language Teaching* (pp. 1-7). Hoboken: Wiley.

Abrahamsson, N., & Hyltenstam, K. (2009). Age of onset and nativelikeness in a second language: Listener perception versus linguistic scrutiny. *Language Learning*, *59* (2), 249-306.

Ackerman, P., Beier, M., & Boyle, M. (2005). Working memory and intelligence. *Psychological Bulletin, 131* (1), 30–60.

Adams, R. (2003). L2 output, reformulation and noticing: Implications for IL development. *Language Teaching Research,* 7, 347–376.

Adams, R., & Mohammadtaghi, S. (2014). The relationship between working memory and L2 reading comprehension. *Applied Research on English Language, 3* (2), 19-34.

Adams, R., Nuevo, A., & Egi, T. (2011). Corrective feedback, learner responses, and SLA: Does feedback promote learning in NNS-NNS interactions? *Modern Language Journal*, *95* (S1), 42-63.

Afitska, O. (2015). Roles of focus-on-form instruction, corrective feedback and uptake in second language classrooms: some insights from recent second language acquisition research. *The Language Learning Journal, 43* (1), 57-73.

Alderson, J., Clapham, C., & Steel, D. (1997). Metalinguistic knowledge, language aptitude and language proficiency. *Language Teaching Research, 1*, 93-121.

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

Alptekin, C., & Erçetin, G. (2010). The role of L1 and L2 working memory in literal and inferential comprehension in L2 reading. *Journal of Research in Reading, 33,* 206–219.

Ammar, A. (2008). Prompts and recasts: Differential effects on second language morphosyntax. *Language Teaching Research, 12,* 183-210.

Amnuai, W. (2020). An error analysis of research project abstracts written by Thai undergraduate students. *Advances in Language and Literacy Studies, 11* (4), 13-20.

Anderson, J. (1981). *Cognitive skills and their acquisition*. Hilldale, NJ: Erlbaum.

Anderson, J. (1982). Acquisition of cognitive skill. *Psychological Review, 89,* 369-406.

Anderson, J. (1983). *The architecture of cognition. Cambridge*, MA: Harvard University Press.

Anderson, J. (1985). *Cognitive psychology and its implications*. New York: Freeman.

Anderson, J. (1993). *Rules of the mind*. Hillsdale. NJ: Lawrence Erlbaum.

Anderson, J. (2007). *How can the human mind occur in the physical universe?* New York: Oxford University Press.

Anderson, J., & Fincham, J. (1994). Acquisition of procedural skills from examples. *Journal of Experimental Psychology, 20,* 1322-1340.

Ando, J., Fukunaga, N., & Kurahachi, J. (1992). A comparative study on the two EFL teaching methods: The communicative and the grammatical approach. *Japanese Journal of Educational Psychology*, *40*, 247–256.

Andringa, S., & Rebuschat, P. (2015). New directions in the study of implicit and explicit learning: An introduction. *Studies in Second Language Acquisition, 37*, 185-196.

Ashwell, T. (2000). Patterns of teacher response to student writing in a multi-draft composition classroom: Is content feedback followed by form feedback the best method? *Journal of Second Language Writing, 9*, 227–57.

Atkins, P., & Baddeley, A. (1998). Working memory and distributed vocabulary learning. *Applied psycholinguistics*, *19* (4), 537-552.

Baddeley, A. (1986). *Working memory.* Oxford, UK: Oxford University Press.

Baddeley, A. (2000). The episodic buffer: A new component of working memory? *Trends in Cognitive Sciences*, *4*, 417–423.

Baddeley, A. (2003). Working memory and language: an overview. *Journal of Communication Disorders*, *36*, 189–208.

Baddeley, A. (2007). *Working memory, thought, and action*. Oxford, UK: Oxford University Press.

Baddeley, A. (2010). Working memory. *Current Biology*, *20*, 136–140.

Baddeley, A. (2017). Modularity, working memory and language acquisition. *Second Language Research*, *33*, 299–311.

Baddeley, A., & Hitch, G. (1974). Working memory. In G. Bower (Ed.), *Recent Advances in Learning and Motivation* (pp. 47-90). New York: Academic Press.

Bandura, A. (1986). *Social Foundations of Thought and Action: A Social Cognitive Theory*. Prentice-Hall.

Bao, M., Egi, T., & Han, Y. (2011). Classroom study on noticing and recast features: Capturing learner noticing with uptake and stimulated recall. *System, 39*, 215-228.

Baoshu, Y., & Chuanbi, N. (2015). Planning and working memory effects on L2 performance in Chinese EFL learners' argumentative writing. *Indonesian Journal of Applied Linguistics*, *5*, 44-53.

Baoshu, Y., & Luo, S. (2012). The effect of working memory capacity on written language production of second language learners. *Foreign Language Teaching and Research*, *44*, 536-546.

Bennui, P. (2008). A study of L1 interference in the writing of Thai EFL students. *Malaysian Journal of ELT Research*, *4*, 72-102.

Benson, S., & DeKeyser, R. (2019). Effects of written corrective feedback and language Aptitude on verb tense accuracy. *Language Teaching Research, 23* (6), 702-726.

Bergsleithner, J. (2010). Working memory capacity and L2 writing performance. *Ciências & Cognição*, *15* (2), 2-20.

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.

Biber, D., Nekrasova, T., & Horn, B. (2011). The effectiveness of feedback for L1 English and L2-writing development: A meta-analysis. *ETS research report series, 2011* (1), 1-110.

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

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

Bitchener, J. (2017). Why some L2 learners fail to benefit from written corrective feedback. In H. Nassaji & E. Kartchava (Eds.), *Corrective feedback in second language learning and teaching: Research, theory, applications, implications* (pp. 129-140). Milton: Taylor & Francis Group.

Bitchener, J., & Ferris, D. (2012). *Written corrective feedback in second language acquisition and writing*. New York; London: Routledge.

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

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

Bitchener, J., & Knoch, U. (2009b). The value of a focused approach to written corrective feedback. *ELT Journal, 63*, 204–211.

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

Bitchener, J., & Knoch, U. (2010b). The contribution of written corrective feedback to language development: A ten-month investigation. *Applied Linguistics*, *31* (2), 193–214.

Bitchener, J., & Knoch, U. (2015). Written corrective feedback studies: Approximate replication of Bitchener & Knoch (2010a) and Van Beuningen, De Jong & Kuiken (2012). *Language Teaching, 48* (3), 405-414.

Bitchener, J., & Storch, N. (2016). *Written corrective feedback for L2 development*. Bristol, UK; Buffalo, USA; Toronto, Canada: Multilingual Matters.

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

Bley-Vroman, R. (1989). What is the Logical Problem of Foreign Language Learning? In S. Gass & J. Schachter (Eds.), *Linguistic perspective on second language acquisition* (pp. 41-68). Cambridge, UK; New York, USA: Cambridge University Press.

Bley-Vroman, R. (1990). The logical problem of foreign language learning. *Linguistic Analysis, 20*, 3-49.

Bonilla López, M., Van Steendam, E., & Buyse, K. (2017). Comprehensive corrective feedback on low and high proficiency writers: Examining attitudes and preferences. *ITL-International Journal of Applied Linguistics*, *168*, 91–128.

Bonilla López, M., Van Steendam, E., Speelman, D., & Buyse, K. (2018). The differential effects of comparative feedback forms in the second language writing class. *Language Learning, 68* (3), 813-850.

Bowles, M. (2011). Measuring implicit and explicit linguistic knowledge. *Studies in Second Language Acquisition*, *33*, 247-271.

Brooks, L., & M. Swain (2009). Languaging in collaborative writing: creation of and response to expertise. In A. Mackey and C. Polio (Eds.), *Multiple perspectives on interaction: second language research in honor of Susan M. Gass* (pp. 58-89). New York: Routledge.

Brown, D. (2016). The type and linguistic foci of oral corrective feedback in the L2 classroom: A meta-analysis. *Language Teaching Research, 20* (4), 436-458.

Bruton, A. (2009). Improving accuracy is not the only reason for writing, and even if it were. *System, 37*, 600-613.

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

Buffington, J., & Morgan-Short, K. (2019). Declarative and procedural memory as individual differences in second language aptitude. In Z. Wen, P. Skehan, A. Biedroń, S. Li. & R. Sparks (Eds.), *Language aptitude: Advancing theory, testing, research and practice* (pp. 215–237). New York: Routledge.

Canovas Guirao, J., Roca de Larios, J., & Coyle, Y. (2015). The use of models as a written feedback technique with young EFL learners. *System, 52* (1), 63-77.

Carroll, J. (1962). The prediction of success in intensive foreign language training. In R. Glaser (Ed.), *Training research and education* (pp. 87-136). Pittsburgh: University of Pittsburgh Press.

Carroll, J. (1981). Twenty-five years of research in foreign language aptitude. In K. Diller (Ed.), *Individual differences and universals in language learning aptitude* (pp.

83–118). Rowley, MA: Newbury House.

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.

Chen, S., Nassaji, H., & Liu, Q. (2016). EFL learners' perceptions and preferences of written corrective feedback: A case study of university students from mainland China. *Asian-Pacific Journal of Second and Foreign Language Education, 1* (5), 1-17.

Chomsky, N. (1981). *Lectures on government and binding*. Dordrecht, Holland; Cinnaminson, NJ: Foris Publications.

Chomsky, N. (1987). Transformational Grammar: Past, Present, and Future. *Studies in English Language and Literature*, 33-80.

Christiansen, M., & Chater, N. (2001). *Connectionist psycholinguistics*. Westport, CO: Ablex.

Cintrón-Valentín, M., & Ellis, N. (2015). Exploring the interface: Explicit focus-on-form instruction and learned attentional biases in L2 Latin. *Studies in Second Language Acquisition*, *37*, 197–235.

Cintrón-Valentín, M., & Ellis, N. C. (2016). Salience in second language acquisition: Physical form, learner attention, and instructional focus. *Frontiers in Psychology, 7*, 1-21.

Clahsen, H. (1990). The comparative study of first and second language acquisition. *Studies in Second Language Acquisition, 12*, 135-153.

Clahsen, H., Felser, C., Neubauer, K., Sato, M., & Silva, R. (2010). Morphological structure in native and nonnative language processing. *Language Learning*, *60*, 21-43.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale, NJ: Erlbaum.

Cohen, J. (1992). A power primer. *Psychological Bulletin, 112* (1), 155-159.

Conway, R., Kane, M., & Engle, R. (2003). Working memory capacity and its relation to general intelligence. *Trends in Cognitive Science, 7*, 547-552.

Conway, A., Cowan, N., Bunting, M., Therriault, D., & Minkoff, S. (2002). A latent variable analysis of working memory capacity, short-term memory capacity, processing speed, and general fluid intelligence. *Intelligence, 30*, 163-184.

Conway, A., Jarrold, C., Kane, M., Miyake, A., & Towse, J. (2007). *Variation in working memory*. Oxford: New York.

Conway, A., Kane, M., Bunting, M., Hambrick, Z., Wilhelm, O., & Engle, W. (2005). Working memory span tasks: A methodological review and user's guide. *Psychonomic Bulletin & Review*, *12*, 769-786.

Cook, V. (1988). *Chomsky's Universal Grammar: An introduction.* Oxford: Blackwell.

Cowan, N. (2005). *Working memory capacity*. Hove, UK: Psychology Press.

Coyle, Y., & Roca de Larios, J. (2014). Exploring the role played by error correction and models on children's reported noticing and output production in a L2 writing task. *Studies in Second Language Acquisition, 36* (3), 451-485.

Coyle, Y., & Roca de Larios, J. (2020). Exploring young learners' engagement with models as a written corrective technique in EFL and CLIL settings. *System, 95*, 1-14.

Culbertson, G., Andersen, E., & Christiansen, M. (2020). Using utterance recall to assess second language proficiency. *Language Learning, 70* (S1)*, 104-132.

Csizér, K., & Dörnyei, Z. (2005). The internal structure of language learning motivation and its relationship with language choice and learning effort. *Modern Language Journal, 89*, 19-36.

Daneman, M., & Merikle, P. (1996). Working memory and language comprehension: A meta-analysis. *Psychonomic Bulletin & Review*, *3*, 422-433.

Davies, W., & Kaplan, T. (1998). Native speaker vs. L2 learner grammaticality judgments. *Applied Linguistics*, *19*, 183-203.

DeKeyser, R. (1994). Implicit and explicit learning of second language grammar: A pilot study. *TESOL Quarterly, 28* (1), 188-194.

DeKeyser, R. (1995). Learning second language grammar rules: An experiment with a miniature linguistic system. *Studies in Second Language Acquisition, 17*, 379 – 410.

DeKeyser, R. (1997). Beyond explicit rule learning: Automatizing second language morphosyntax. *Studies in Second Language Acquisition, 19,* 195-222.

DeKeyser, R. (1998). Beyond Focus on Form: Cognitive perspectives on learning and practicing second language grammar. In C. Doughty & J. Williams (Eds.), *Focus on form in classroom second language acquisition* (pp. 42-63). Cambridge: Cambridge University Press.

DeKeyser, R. (2000). The robustness of critical period in second language acquisition. *Studies in Second Language Acquisition, 22,* 499–534.

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

DeKeyser, R. (2007). Skill Acquisition Theory. In B. VanPatten & J. Williams (Eds.), *Theories in second language acquisition: An introduction* (pp. 97-113). Mahwan, NJ: Erlbaum.

DeKeyser, R. (2015). Skill Acquisition Theory. In B. VanPatten & J. Williams (Eds.), *Theories in second language acquisition: An introduction* (pp. 95-112). New York, USA; Abingdon, Oxon, UK: Routledge.

Denhovska, N., Serratrice, L., & Payne, J. (2016). Acquisition of second language grammar under incidental learning conditions: The role of frequency and working memory. *Language Learning, 66* (1), 159-190.

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

Dörnyei, Z. (1990). Conceptualizing motivation in foreign-language learning. *Language Learning, 40,* 45-78.

Dörnyei, Z. (2003). Attitudes, orientations, and motivations in language learning: Advances in theory, research, and applications. *Language Learning, 53*, 3-32.

Dörnyei, Z. (2005). The psychology of the language learner: Individual differences in second language acquisition*.* Mahwah, NJ: Lawrence Erlbaum.

Dörnyei, Z. (2010). The relationship between language aptitude and language learning motivation. In E. Macaro (Ed.), *Continuum companion to second language acquisition* (pp. 247-267). London: Continuum.

Dörnyei. Z., & Ryan, S. (2015). *The psychology of the language learner revisited*. New York: Routledge.

Dörnyei, Z., & Ushioda, E. (2009). *Motivation, language identity and the L2 self*. Bristol, UK: Multilingual Matters.

Dörnyei, Z., & Ushioda, E. (2011). *Teaching and researching motivation*. Harlow, UK: Longman.

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. Long (Eds.), *Handbook of second language acquisition* (pp. 256-310). Oxford: Blackwell.

Doughty, C., & Williams, J. (1998). Pedagogical choices in focus on form. In C. Doughty & J. Williams (Eds.), *Focus on form in classroom second language acquisition* (pp. 197-261). Cambridge: Cambridge University Press.

Dulay, H., & Burt, M. (1974). Natural sequences in child second language acquisition. *Language Learning, 24*, 37-53.

Ellis, N. (2005). At the interface: How explicit knowledge affects implicit language learning. *Studies in Second Language Acquisition, 27*, 305-352.

Ellis, N. (2008). The psycholinguistics of the interaction hypothesis. In A. Mackey & C. Polio (Eds.), *Multiple perspectives on interaction in SLA: Second language research in honor of Susan M. Gass* (pp. 11-40). New York: Routledge.

Ellis, N. (2011). Implicit and explicit SLA and their interface. In C. Sanz & R. Leow (Eds.), *Implicit and explicit language learning* (pp. 35-47). Washington, DC: Georgetown University Press.

Ellis, N. (2012). Frequency-based accounts of second language acquisition. In S. Gass & A. Mackey (Eds.), *The routledge handbook of second language acquisition* (pp. 193-210). London: Routledge.

Ellis, N. (2015a). Cognitive and social aspects of learning from usage. In T. Cadierno & S. Eskildsen (Eds.), *Usage-based perspectives on second language learning* (pp. 49-73). Berlin: De Gruyter Mouton.

Ellis, N. (2015b). Implicit and explicit language learning: Their dynamic interface and complexity. In P. Rebuschat (Ed.), *Implicit and explicit learning of languages* (pp. 3-23). Philadelphia: John Benjamins Publishing Company.

Ellis, N. (2016a). Frequency in language learning and language change. In H. Behrens & S. Pfander (Eds.), *Experience counts: Frequency effects in language acquisition, language change and language processing* (pp. 239-254). Berlin; Boston: De Gruyter.

Ellis, N. (2016b). Salience, cognition, language complexity, and complex adaptive systems. *Studies in Second Language Acquisition, 38*, 341-351.

Ellis, N., & Sagarra, N. (2010). The bounds of adult language acquisition: Blocking and learned attention. *Studies in Second Language Acquisition*, *32* (4), 553-580.

Ellis, N., & Sagarra, N. (2011). Learned attention in adult language acquisition: A replication and generalization study and meta-analysis. *Studies in Second Language Acquisition*, *33*, 589-624.

Ellis, N., & Sinclair, S. (1996). Working memory in the acquisition of syntax: Putting language in good order. *Quarterly Journal of Experimental Psychology*, *49A*, 234-250.

Ellis, N., & Wulff, S. (2015). Usage-based approaches to SLA. In B. VanPattern & J. Williams (Eds.), *Theories in second language acquisition. An introduction* (pp. 75-93). Mahwah, NJ: Erlbaum.

Ellis, N., & Wulff, S. (2018). Usage-based approaches to second language acquisition. In D. Miller, F. Bayram, J. Rothman & L. Serratrice (Eds.), *Bilingual cognition and language: The state of the science across its subfields* (pp. 37-56). Amsterdam; Philadelphia: John Benjamins.

Ellis, N., Hafeez, K., Martin, K., Chen, L., Boland, J., & Sagarra, N. (2014). An eye-tracking study of learned attention in second language acquisition. *Applied Psycholinguistics*, *35*, 547-579.

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

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

Ellis, R. (1997). *SLA research and language teaching.* Oxford: Oxford University Press.

Ellis, R. (2001). Investigating form-focused instruction. *Language Learning, 51* (1), 1-46.

Ellis, R. (2002). Does form-focused instruction affect the acquisition of implicit knowledge?: A review of the research. *Studies in Second Language Acquisition, 24* (2), 223-236.

Ellis, R. (2004). The definition and measurement of L2 explicit knowledge. *Language Learning*, *54*, 227-275.

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

Ellis, R. (2005b). Principles of instructed language learning. *System, 33* (2), 209-224.

Ellis, R. (2006). Researching the effects of form-focused instruction on L2 acquisition. In K. Bardovi-Harlig & Z. Dörnyei (Eds.), *Themes in SLA research* (pp. 18-41). Amsterdam: Benjamins.

Ellis, R. (2008a). A typology of corrective feedback types. *English Language Teaching Journal*, 97–107.

Ellis, R. (2008b). *The study of second language acquisition*. Oxford: Oxford University Press.

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

Ellis, R. (2009b). Implicit and explicit learning, knowledge and instruction. In R. Ellis, S. Loewen, C. Elder, R. Erlam, J. Philp, & H. Reinders (Eds.), Implicit and explicit knowledge in second language learning, testing and teaching (pp. 3-25). Bristol, UK: Multilingual Matters.

Ellis, R. (2009c). Measuring implicit and explicit knowledge of a second language. In R. Ellis, S. Loewen, C. Elder, R. Erlam, J. Philp, & H. Reinders (Eds.), *Implicit and explicit knowledge in second language learning, testing and teaching* (pp. 31-64). Bristol, UK: Multilingual Matters.

Ellis, R. (2010). A framework for investigating oral and written corrective feedback. *Studies in Second Language Acquisition*, *32*, 335-349.

Ellis, R. (2012a). *Form-focused instruction and second language learning*. Oxford: Wiley Blackwell.

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

Ellis, R., & Barkhuizen, G. (2005). *Analysing Learner Language*. Oxford: Oxford University Press.

Ellis, R., & Sheen, Y. (2006). Re-examining the role of recasts in second language acquisition. *Studies in Second Language Acquisition, 28* (4), 575-600.

Ellis, R., Basturkmen, H., & Loewen, S. (2001). Learner uptake in communicative ESL lessons. *Language Learning*, *51*, 281-318.

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., Sheen, Y., Murakami, M., & Takashima, H., (2008). The effects of focused and unfocused written corrective feedback in an English as a foreign language context. *System, 36* (3), 353-371.

Ellis, R., Loewen, S., Elder, C., Erlam, R., Philp, J., & Reinders, H. (2009). *Implicit and explicit knowledge in second language learning, testing and teaching*. Bristol, UK; Buffalo, USA; Toronto, Canada: Multilingual Matters.

Engeström, Y. (1987). *Learning by expanding: An activity-theoretical approach to developmental research*. Helsinki, Finland: Orienta-Konsultit.

Engle, R., & Kane, M. (2004). Executive attention, working memory capacity, and a two-factor theory of cognitive control. In B. Ross (Ed.), *The psychology of learning and motivation* (pp. 145-199). New York, NY: Academic Press.

Engle, R., Carullo, J., & Collins, K. (1991). Individual differences in working memory for comprehension and following directions. *Journal of Educational Research, 84*, 253-262.

Erçetin, G., & Alptekin, C. (2013). The explicit/implicit knowledge distinction and working memory: Implications for second-language reading comprehension. *Applied Psycholinguistics*, *34*, 727-753.

Erlam, R. (2005). Language aptitude and its relationship to instructional effectiveness in second language acquisition. *Language Teaching Research*, *9*, 147-171.

Erlam, R. (2006). Elicited imitation as a measure of L2 implicit knowledge: An empirical validation study. *Applied Linguistics*, *27*, 464-491.

Erlam, R., Ellis, R., & Batstone, R. (2013). Oral corrective feedback on L2 writing: Two approaches compared. *System, 41* (2), 257-268.

Evans, N., Hartshorn, K., & Strong-Krause, D. (2011). The efficacy of dynamic written corrective feedback or university matriculated ESL learners. *System, 39*, 229-239.

Evans, N., Hartshorn, J., McCollum, M., & Wolfersberger, M., (2010). Contextualizing corrective feedback in L2 writing pedagogy. *Language Teaching Research, 14* (4). 445-463.

Faretta-Stutenberg, M., & Morgan-Short, K. (2018). The interplay of individual differences and context of learning in behavioral and neurocognitive second language development. *Second Language Research*, *34*, 67-101.

Farrokhi, F., & Sattarpour, S. (2012). The effects of direct written corrective feedback on improvement of grammatical accuracy of high-proficient L2 learners. *World Journal of Education, 2* (2), 49-57.

Fathman, A., & Whalley, E. (1990). Teacher response to student writing: Focus on form versus content. In B. Kroll (Ed.), *Second language writing: Research insights for the classroom* (pp. 178-190). Cambridge: Cambridge University Press.

Ferris, D. (1995). Teaching ESL composition students to become independent self-editors. *TESOL Journal, 4* (4), 18-22.

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

Ferris, D. (1999). The case for grammar correction in L2 writing classes: A response to Truscott (1996). *Journal of Second Language Writing, 8*, 1-10.

Ferris, D. (2003). *Response to student writing: Implications for second language students*. Mahwah, NJ: Lawrence Erlbaum.

Ferris, D. (2004). The "grammar correction" debate in L2 writing: Where are we, and where do we go from here? (and what do we do in the meantime?). *Journal of Second Language Writing, 13*, 49-62.

Ferris, D. (2006). Does error feedback help student writers? New evidence on the short- and long-term effects of written error correction. In K. Hyland & F. Hyland

(Eds.), *Feedback in second language writing: Contexts and issues* (pp. 81-104). Cambridge: Cambridge University Press.

Ferris, D. (2010). Second language writing research and written corrective feedback in SLA: Intersections and practical applications. *Studies in Second Language Acquisition, 32*, 181-201.

Ferris, D. (2012). Written corrective feedback in second language acquisition and writing studies (research timeline). *Language Teaching, 45* (4), 446-459.

Ferris, D., & Hedgcock, J. (2014). *Teaching L2 composition: Purpose, process, and practice*. New York: Routledge.

Ferris, D., & 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, Vancouver, British Columbia, Canada.

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

Ferris, D., Liu, H., Sinha, A., & Senna, M. (2013). Written corrective feedback for individual L2 writers. *Journal of Second Language Writing*, *22*, 307-329.

Field, A. (2018). *Discovering statistics using IBM SPSS statistics.* London: SAGE publications Ltd.

Flynn, S. (1996). A parameter-setting approach to second language acquisition. In W. Ritchie & T. Bhatia (Eds.), *Handbook of Second Language Acquisition* (pp. 121-158). San Diego: Academic Press.

Foote, R. (2011). Integrated knowledge of agreement in early and late English-Spanish bilinguals. *Applied Psycholinguistics, 32*, 187-220.

Frear, D. (2012). The effect of written feedback and revision on intermediate Chinese learners' acquisition of English. Unpublished doctoral dissertation, The University of Auckland, New Zealand.

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.

Gardner, R. (1985). *Social psychology and second language learning: The role of attitudes and motivation*. London: Edward Arnold.

Gardner, R. (2000). Correlation, causation, motivation, and second language acquisition. *Canadian Psychology, 41* (1), 10-24.

Gardner, R. (2001). Integrative motivation and second language acquisition. In Z. Dörnyei & R. Schmidt (Eds.), *Motivation and second language acquisition* (pp. 1-20). Honolulu, HI: University of Hawai'i Press.

Gardner, R., & Lambert, W. (1959). Motivational variables in second language acquisition. *Canadian Journal of Psychology, 13*, 266-272.

Gardner, R., & MacIntyre, P. (1991). An instrumental motivation in language study: Who says it isn't effective? *Studies in Second Language Acquisition, 13*, 57-72.

Gass, S. (1997). *Input, interaction and the development of second languages.* Mahwah, NJ: Lawrence Erlbaum Associates.

Gass, S. (2003). Input and interaction. In C. Doughty & M. Long (Eds.), *The Handbook of second language acquisition* (pp. 224-255). Oxford, England: Blackwell,

Gass, S. (2015). Comprehensible input and output in classroom interaction. In N. Markee (Ed.), *The Handbook of classroom discourse and interaction* (pp. 182-197)*. Hoboken: Wiley.

Gass, S., & Mackey, A. (2015). Input, interaction, and output in second language acquisition. In B. VanPatten & J. Williams (Eds.), *Theories in second language acquisition: An introduction* (pp. 180-206). London: Routledge.

Gass, S., & Mackey, A. (2016). *Stimulated recall methodology in applied linguistics and L2 research.* New York and London: Routledge.

Gilabert, R., & Muñoz, C. (2010) Differences in attainment and performance in a foreign language: The role of working memory capacity. *International Journal of English Studies, 10*, 19-41.

Godfroid, A. (2016). The effects of implicit instruction on implicit and explicit knowledge development. *Studies in Second Language Acquisition*, *38* (2), 177-215.

Godfroid, A., Loewen, S., Jung, S., Park, J., Gass, S., & Ellis, R. (2015). Timed and untimed grammaticality judgements measure distinct types of knowledge. *Studies in Second Language Acquisition*, *37*, 269-297.

Goldberg, A. (1995). *Constructions: A construction grammar approach to argument structure.* Chicago: University of Chicago Press.

Goldberg, A. (2003). Constructions: A new theoretical approach to language. *Trends in Cognitive Science, 7, 219*-224.

Goo, J. (2012). Corrective feedback and working memory capacity in interaction-driven L2 learning. *Studies in Second Language Acquisition, 34*, 445-474.

Goo, J., & Mackey, A. (2013). The case against the case against recasts. *Studies in Second Language Acquisition*, *35*, 127-165.

Goss, N., Ying-Hua, Z., & Lantolf, J. (1994). Two heads may be better than one: Mental activity in second-language grammaticality judgments. In E. Tarone, S. M. Gass, & A. Cohen (Eds.), *Research methodology in second-language acquisition* (pp. 263-286). Mahwah, NJ: Erlbaum.

Granena, G., & Long, M. (2013). Age of onset, length of residence, aptitude and ultimate L2 attainment in three linguistic domains. *Second Language Research, 29* (3), 311-343.

Grey, S., Serafini, E., Cox, J., & Sanz, C. (2015). The role of individual differences in the study abroad context: Cognitive capacity and language development during short-term intensive language exposure. *The Modern Language Journal, 99* (1), 137-157.

Gries, S., & Ellis, N. (2015). Statistical measures for usage-based linguistics. *Language Learning, 65* (1), 228-255.

Grüter, T., Lew-Williams, C., & Fernald, A. (2012). Grammatical gender in L2: A production or a real-time processing problem? *Second Language Research, 28*, 191-215.

Guénette, D. (2007). Is feedback pedagogically correct? Research design issues in studies of feedback on writing. *Journal of Second Language Writing*, *16*, 40-53.

Guo, Q. (2015) The effectiveness of written CF for L2 development: A mixed-method study of written CF types, error categories and proficiency levels. Unpublished doctoral thesis. AUT University, Auckland, New Zealand.

Guo, X., & Yang, Y. (2018). Effects of corrective feedback on EFL learners' acquisition of third-person singular form and the mediating role of the cognitive style. *Journal of Psycholinguist Research, 47*, 841-858.

Gutiérrez, X. (2013). The construct validity of grammaticality judgment tests as measures of implicit and explicit knowledge. *Studies in Second Language Acquisition*, *35*, 423-449.

Hale, S., Myerson, J., Rhee, S., Weiss, C., & Abrams, R. (1996). Selective interference with the maintenance of location information in working memory. *Neuropsychology, 10*, 228-240.

Hama, M., & Leow, R. (2010). Learning without awareness revisited: Extending Williams (2005). *Studies in Second Language Acquisition*, *32*, 465-491.

Hambrick, D., Oswald, F., Darowski, E., Rench, T., & Brou, R. (2010). Predictors of multitasking performance in a synthetic work paradigm. *Applied Cognitive Psychology*, *24* (8), 1149-1167.

Hamrick, P. (2015). Declarative and procedural memory abilities as individual differences in incidental language learning. *Learning and Individual Differences*, *44*, 9-15.

Han, Z. (2002). A study of the impact of recasts on tense consistency in L2 output. *TESOL Quarterly, 36,* 543-572.

Hanaoka, O., & Izumi, S. (2012). Noticing and uptake: Addressing pre-articulated covert problems in L2 writing. *Journal of Second Language Writing, 21* (4), 332-347.

Harrington, M., & Sawyer, M. (1992). L2 working memory capacity and L2 reading skill. *Studies in Second Language Acquisition, 14*, 25-38.

Hartshorn, K., & Evans, N. (2012). The differential effects of comprehensive corrective feedback on L2 writing accuracy. *Journal of Linguistics ad Language Teaching, 3* (2), 217-247.

Hartshorn, K., & Evans, N. (2015). The effects of dynamic written corrective feedback: A 30-week study. *Journal of Response to Writing*, *1* (2), 6-34.

Hartshorn, K., Evans, N., Merrill, P., Sudweeks, R., Strong-Krause, D., & Anderson, N. (2010). Effects of dynamic corrective feedback on ESL writing accuracy. TESOL Quarterly, *44*, 84-109.

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

Hedgcock, J., & Lefkowitz, N. (1996). Some input on input: Two analyses of student responses to expert feedback on L2 writing. *Modern Language Journal, 80*, 287-308.

Hopp, H. (2013). Grammatical gender in adult L2 acquisition: Relations between lexical and syntactic variability. *Second Language Research, 29*, 33-56.

Housen, A. & Pierrard, M. (2006). Investigating instructed second language acquisition. In A. Housen, A. & M. Pierrard, M. (Eds.), *Investigations in instructed second language acquisition* (pp. 1-27). Berlin: Mouton de Gruyter.

Hulstijn, J., & Schmidt, R. (1994). Guest editors' introduction (to special issue on consciousness in second language learning. *AILA Review, 11*, 5-10.

Hyland, F. (2010). Future directions in feedback on second language writing: Overview and research agenda. *International Journal of English Studies,10* (2), 171-182.

Hyland, K., & Hyland, F. (2006). Context and issues in feedback on L2 writing: An introduction. In K. Hyland & F. Hyland (Eds.), *Feedback in second language writing* (pp. 1-20). Cambridge, England: Cambridge University Press.

Ionin, T., & Montrul, S. (2010). The role of L1 transfer in the interpretation of articles with definite plurals in L2 English. *Language Learning*, *60* (4), 877-925.

Juffs, A. (2004). Representation, processing and working memory in a second language. *Transactions of the Philological Society*, *102* (2), 199-225.

Juffs, A. (2005). The influence of first language on the processing of wh-movement in English as a second language. *Second Language Research*, *21* (2), 121-151.

Juffs, A., & Harrington, M. (2011). Aspects of working memory in L2 learning, *Language Teaching, 44*, 137-166.

Kane, M., Hambrick, D., & Conway, A. (2005) Working memory capacity and fluid intelligence are strongly related constructs: Comment on Ackerman, Beier, and Boyle (2005). *Psychological Bulletin, 131*, 66-71.

Kane, M., Bleckley, M., Conway, A., & Engle, R. (2001). A controlled-attention view of working-memory capacity. *Journal of Experimental Psychology: General*, *130*,

169-183.

Kang, E., & Han, Z. (2015). The efficacy of written corrective feedback in improving L2 written accuracy: A meta-analysis. *The Modern Language Journal, 99* (1), 1-18.

Kartchava, E., & Amnar, A. (2014). Learners' beliefs as mediators of what is noticed and learned in the language classroom. *TESOL Quarterly, 48* (1), 86-109.

Kellogg, R. (1996). A model of working memory in writing. In C. Levy, & S. Ransdell (Eds.). *The science of writing: Theories, methods, individual differences, and applications* (pp. 57-71). Mahwah, NJ: Lawrence Erlbaum Associates.

Kellogg, R., Whiteford, A., Turner, C., Cahill, M., & Mertens, A. (2013). Working memory in written composition: An evaluation of the 1996 model. *Journal of Writing Research, 5* (2), 159-190.

Kepner, C. (1991). An experiment in the relationship of types of written feedback to the development of second language writing skills. *The Modern Language Journal, 75*(3), 305-313.

Khumpee, S., & Yodkumlue, B. (2017). Grammatical errors in English essays written by Thai EFL undergraduate students. *Journal of Education, 11* (4), 139-153.

Kim, H, & Bowles, M. (2019). How deeply do second language learners process written corrective feedback? Insights gained from think aloud. *TESOL Quarterly, 53* (4), 913-938.

Kim, J., & Nam, H. (2017). Measures of implicit knowledge revisited: Processing modes, time pressure, and modality. *Studies in Second Language Acquisition*, *39*, 431-457.

Kim, Y., & Schatschneider, C. (2017). Expanding the developmental models of writing: A Direct and Indirect Effects Model of Developmental Writing (DIEW). *Journal of Educational Psychology, 109* (1), 35-50.

Kim, Y., Payant, C., & Pearson, P. (2015). The intersection of task-based interaction, task complexity and working memory: L2 question development through recasts in a laboratory setting. *Studies in Second Language Acquisition, 37* (3), 549-581.

Kim, Y., Choi, B., Kang, S., Kim, B., & Yun, H. (2020). Comparing the effects of direct and indirect synchronous written corrective feedback: Learning outcomes and students' perceptions. *Foreign Language Annals, 53*, 176-199.

Koo, T., & Li, M. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of Chiropractic Medicine, 15,* 155-163.

Kormos, J., & Safar, A. (2008) Phonological short-term memory, working memory and foreign language performance in intensive language learning. *Bilingualism: Language and Cognition, 11*, 261-271.

Krashen, S. (1981). *Second Language Acquisition and Second Language Learning.* Oxford: Pergamon Press.

Krashen, S. (1982). *Principles and practice in second language acquisition*. Englewood Cliffs: Prentice-Hall.

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

Kuiken, F., & Vedder, I. (2011). Task complexity and linguistic performance in L2 writing and speaking: The effect of mode. In P. Robinson (Ed.), *Second language task complexity. Researching the cognition hypothesis of language learning and performance* (pp. 91-104). Amsterdam: John Benjamins.

Kurzer, K. (2018). Dynamic written corrective feedback in developmental multilingual writing classes. *TESOL Quarterly, 52* (1), 5-33.

Lalande, J. (1982). Reducing composition errors: An experiment. *The Modern Language Journal, 66* (2), 140-149.

Lantolf, J., & Poehner, M. (2014). *Sociocultural theory and the pedagogical imperatives in L2 education. Vygotskian praxis and the research/practice divide*. New York: Routledge.

Lantolf, J., Thorne, S., & Poehner, M. (2015). Sociocultural theory and second language Development. In B. VanPatten & J. Williams (Eds.), *Theories in second language acquisition* (pp. 207-226). New York: Routledge.

Lapkin, S., Swain, M., & Smith, M. (2002). Reformulation and the learning of French pronominal verbs in a Canadian French immersion context. *Modern Language Journal, 86*, 485-507.

Lee, I. (2019). Teachers' frequently asked questions about focused written corrective feedback. *TESOL Journal, 10*, 1-15.

Leeser, M. (2007). Learner-based factors in L2 reading comprehension and processing grammatical form: Topic familiarity and working memory. *Language Learning, 57*, 229-270.

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

Li, S. (2013). The interactions between the effects of implicit and explicit feedback and individual differences in language analytic ability and working memory. *The Modern Language Journal, 97* (3), 634-654.

Li, S. (2017). Cognitive differences and ISLA. In S. Loewen & M. Sato (Eds.), *The Routledge handbook of instructed second language acquisition* (pp. 396-417). New York: Routledge.

Li, S., & Roshan, S. (2019). The associations between working memory and the effects of four different types of written corrective feedback. *Journal of Second Language Writing, 45*, 1-15.

Li, S., & Vuono, A. (2019). Twenty-five years of research on oral and written corrective feedback in System. *System, 84*, 93-109.

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

Lightbown, P. (2007). Transfer appropriate processing as a model for classroom second language acquisition. In Z. Han (Ed.), *Understanding second language process* (pp. 27-44) Clevedon, UK: Multilingual Matters.

Lightbown, P., & Spada, N. (1990). Focus-on-form and corrective feedback in communicative language teaching: Effects on second language learning. *Studies in Second Language Acquisition, 12*, 429-448.

Linck, J., & Weiss, D. (2011). Working memory predicts the acquisition of explicit L2 knowledge. In C. Sanz & R. Leow (Eds). *Implicit and explicit language learning* (pp. 101-114). Washington, DC: Georgetown University Press.

Linck, J., & Weiss, D. (2015). Can working memory and inhibitory control predict second language learning in the classroom? *SAGE Open*, 1-11.

Linck, J., Osthus, P., Koeth, J., & Bunting, M. (2014). Working memory and second language comprehension and production: A meta-analysis. *Psychon Bull Rev, 21,* 861-883.

Liu, Q., & Brown, D. (2015). Methodological synthesis of research on the effectiveness of corrective feedback in L2 writing. *Journal of Second Language Writing, 30*, 66-81.

Loewen, S. (2004). Uptake in incidental focus on form in meaning-focused ESL lessons. *Language Learning, 54* (1), 153-188.

Loewen, S. (2005). Incidental focus on form and second language learning. *Studies in Second Language Acquisition*, *27* (3), 361-386.

Loewen, S. (2009). Grammaticality judgment tests and the measurement of implicit and explicit L2 knowledge. In R. Ellis, S. Loewen, C. Elder, R. Erlam, J. Philp, & H. Reinders (Eds.), *Implicit and explicit knowledge in second language learning, testing and teaching* (pp. 94-112). Bristol, UK: Multilingual Matters.

Loewen, S. (2012). The role of feedback. In S. Gass & A. Mackey (Eds.), *The Routledge handbook of second language acquisition* (pp. 24-40). New York: Routledge.

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

Loewen, S., Li, S., Fei, F., Thompson, A., Nakatsukasa, K., Ahn, S., & Chen, X. (2009). L2 learners' beliefs about grammar instruction and error correction. *The Modern Language Journal, 93*, 91-104.

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

Long, M. (1998). Focus on form in task-based language teaching. *University of Hawaii Working Papers in ESL, 16* (2), 35-49.

Long, M. (2000). Focus on form in task-based language teaching. In R. Rambert & E. Shohamy (Eds.), *Language policy and pedagogy* (pp. 179-192). Amsterdam: John Benjamins.

Long, M. (2007). *Problems in SLA*. Mahwah, NJ: Lawrence Erlbaum Associates.

Long, M. (2014). Psycholinguistic underpinnings: A Cognitive-interactionist theory of Instructed Second Language Acquisition (ISLA). In M. Long (Ed.), *Second language acquisition and task-based language teaching* (pp. 30-62). Hoboken: Wiley.

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

Luk, Z., & Shirai, Y. (2009). Is the acquisition order of grammatical morphemes impervious to L1 knowledge? Evidence from the acquisition of plural -*s*, articles, and possessive'*s*. *Language Learning*, *59* (4), 721-754.

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

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

Lyster, R. (2007). *Learning and teaching languages through content: A counterbalanced approach*. Amsterdam: Benjamins.

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

Lyster, R., & Ranta, L. (2013). Counterpoint piece: The case for variety in corrective feedback research. *Studies in Second Language Acquisition*, *35*, 1-18.

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

Lyster, R., Saito, K., & Sato, M. (2013). State-of-the-art article: Oral corrective feedback in second language classrooms. *Language Teaching, 46*, 1-40.

MacArthur, C., Graham, S., & Fitzgerald, J. (2006). *Handbook of writing research*. New York: Guilford Press.

MacIntyre, P., Baker, S., Clement, R., & Conrod, S. (2001). Willingness to communicate, social support, and language-learning orientations of immersion students. *Studies in Second Language Acquisition, 23*, 369-388.

Mackey, A. (1999). Input, interaction and second language development: An empirical study of question formation in ESL. *Studies in Second Language Acquisition, 21,* 557-587.

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

Mackey, A. (2012). *Input, interaction, and corrective feedback in L2 learning*. Oxford, UK: Oxford University Press.

Mackey, A., & Goo, J. (2007). Interaction research in SLA: A meta-analysis and research synthesis. In: A Mackey (Ed.), *Conversational interaction in SLA: A collection of empirical studies* (pp. 408-452). New York: Oxford University Press.

Mackey, A., & Philp, J. (1998). Conversational interaction and second language development: Recasts, response and red herrings? *Modern Language Journal, 82,* 338-356.

Mackey, A., & Polio, C. (2009). Introduction. In A. Mackey & C. Polio (Eds.), *Multiple perspectives on interaction: Second language research in honor of Susan M. Gass* (pp. 1-10). New York: Routledge.

Mackey, A., & Sachs, R. (2012). Older learners in SLA research: A first look at working memory, feedback, and L2 development. *Language Learning, 62*, 704-740.

Mackey, A., Abbuhl, R., & Gass, S. (2013). Interactionist Approach. In S. Gass & A. Mackey (Eds.), *The Routledge handbook of second language acquisition* (pp.7-23). London: Routledge.

Mackey, A., Adams, R., Stafford, C., & Winke, P. (2010). Exploring the relationship between modified output and working memory capacity. *Language Learning*, *60*, 501-533.

Mackey, A., Philp, J., Egi, T., Fujii, A., & Tatsumi, T. (2002). Individual differences in working memory, noticing of interactional feedback, and L2 development. In P. Robinson (Ed.), *Individual differences and instructed language learning* (pp. 181-209). Philadelphia, PA: John Benjamins.

MacWhinney, B. (1999). *The emergence of language*. Hillsdale, NJ: Erlbaum.

Manchón, R. (2009). Broadening the perspective of L2 writing scholarship: The contribution of research on foreign language writing. In R. Manchón (Ed.),

Writing in foreign language contexts: Learning, teaching, and research (pp.1-19). Bristol: Multilingual Matters.

Manchón, R. (2011). Writing to learn the language: Issues in theory and research. In R. Manchón (Ed.), *Learning to write and writing to learn in an additional language* (pp. 61-82). Amsterdam: John Benjamins.

Manchón, R., & Williams, J. (2016). L2 writing and SLA studies. In R. Manchón & P. Matsuda (Eds.), *Handbook of second and foreign language writing* (pp. 567-586). Berlin; Boston: De Gruyter Mouton.

Mao, Z., & Lee, I. (2020). Feedback scope in written corrective feedback: Analysis of empirical research in L2 contexts. *Assessing Writing, 45*, 1-14.

Martin, K., & Ellis, N. (2012). The roles of phonological short-term memory and working memory in L2 grammar and vocabulary learning. *Studies in Second Language Acquisition, 34,* 379-413.

Masoura, E., & Gathercole, S. (1999). Phonological short-term memory and foreign vocabulary learning. *International Journal of Psychology*, *34*, 383-388.

Masoura, E., & Gathercole, S. (2005). Contrasting contributions of phonological short-term memory and long-term knowledge to vocabulary learning in a foreign language. *Memory, 13*, 422-429.

McDonough, K. (2005). Identifying the impact of negative feedback and learners' responses on ESL questions development. *Studies in Second Language Acquisition, 27* (1), 79-103.

McDonough, K. (2007). Interactional feedback and the emergence of simple past activity verbs in L2 English. In A. Mackey (Ed.), *Conversational interaction in second language acquisition: A series of empirical studies* (pp. 323-338). Oxford: Oxford University Press.

McDonough, K., & Mackey, A. (2006). Responses to recasts: Repetitions, primed production and linguistic development. *Language Learning*, *56*, 693-720.

McDonough, K., & Trofimovich, P. (2016). The role of statistical learning and working memory in L2 speakers' pattern learning. *The Modern Language Journal*, *100*, 428-445.

McLaughlin, B. (1987). *Theories of second language acquisition*. London: Edward Arnold.

McLaughlin, B. (1990). "Conscious" versus "unconscious" learning. *TESOL Quarterly, 24*, 617-631.

Meisel, J. (2011). *First and Second Language Acquisition*. Cambridge: CUP.

Michel, M., Kormos, J., Brunfaut, T., & Ratajczak, M. (2019). The role of working memory in young second language learners' written performances. *Journal of Second Language Writing, 45*, 31-45.

Mifka-Profozic, N. (2015). Effects of corrective feedback on L2 acquisition of tense-aspect verbal morphology. *Language, Interaction, and Acquisition, 6* (1), 149-180.

Mitchell, A., Jarvis, S., O'Malley, M., & Konstantinova, I. (2015) Working memory measures and L2 proficiency. In W. Zheng, M. Mota & A. McNeill (Eds.) *Working memory in second language acquisition and processing* (pp. 270-284). Bristol: Multilingual Matters.

Miyake, A., & Friedman, N. (1998). Individual differences in second language proficiency: Working memory as language aptitude. In A. Healey & L. Bourne (Eds.), *Foreign language learning: Psycholinguistic studies on training and retention* (pp. 339-364). Mahwah, NJ: Lawrence Erlbaum Associates.

Mohammadreza, (2020). The effect of comprehensive written corrective feedback on EFL learners' written syntactic accuracy. *Advances in Language and Literary Studies, 11* (1), 17-26.

Morgan-Short, K., & Ullman, M. (2012). The neurocognition of second language. In S. Gass & A. Mackey (Eds.), *The Routledge handbook of second language acquisition* (pp. 282-300). London: Routledge.

Murakami, A., & Alexopoulou, T. (2016). L1 influence on the acquisition order of English grammatical morphemes: A learner corpus study. *Studies in Second Language Acquisition, 38*, 365-401.

Nassaji, H. (2015). *The interactional feedback dimension in instructed second language learning: Linking theory, research, and practice*. London: Bloomsbury.

Nassaji, H. (2016). Anniversary article: Interactional feedback in second language teaching and learning: A synthesis and analysis of current research. *Language Teaching Research*, *20*, 535-562.

Nassaji, H., & Fotos, S. (2007). Current issues in form-focused instruction. In S. Fotos & H. Nassaji (Eds.), *Form-focused instruction and teacher education: Studies in honour of Rod Ellis* (pp. 7-15). Oxford: Oxford University Press.

Nassaji, H., & Fotos, S. (2011). *Teaching grammar in second language classrooms: Integrating form-focused instruction in communicative context.* New York: Routledge.

Nassaji, H., & Kartchava, E. (2017). *Corrective feedback in second language teaching and learning*. New York: Routledge.

Nassaji, H., & Swain, M. (2000). A Vygotskian perspective on corrective feedback in L2: The effect of random versus negotiated help on the learning of English articles. *Language Awareness* 9.1, 34-51.

Nemati, M., Alavi, S., & Mohebbi, H. (2019). Assessing the effect of focused direct and focused indirect written corrective feedback on explicit and implicit knowledge of language learners. *Language Testing in Asia, 9* (1), 1-18.

Newport, E. (1990). Maturational constraints on language learning. *Cognitive Science, 14* (1), 11-28.

Noels, K., Clément, R., & Pelletier, L. (1999). Perceptions of teachers' communicative style and students' intrinsic and extrinsic motivation. *The Modern Language Journal, 83*, 23-34.

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

Norris, J., & Ortega, L. (2001). Does type of instruction make a difference? Substantive findings from a meta-analytic review. *Language Learning, 51* (1), 157-213.

Norris, J., & Ortega, L. (2003). Defining and measuring SLA. In C. Doughty & M. Long (Eds.), *The handbook of second language acquisition* (pp. 716-761). Malden, Mass: Blackwell.

Oberauer, K., Martin Süβ, H., Wilhelm, O., & Sander, N. (2007). Individual differences in working memory capacity and reasoning ability. In A. Conway, C. Jarrold, M.

Kane, A. Miyake & J. Towse (Eds.), *Variation in working memory* (pp. 49-75). Oxford: Oxford University Press.

O′Brien, I., Segalowitz, N., Collentine, J., & Freed, B. (2006). Phonological memory and lexical, narrative and grammatical skills in second-language oral production by adult learners. *Applied Psycholinguistics*, *27*, 377-402.

O′Grady, W. (1996). Language acquisition without universal grammar: A general nativist proposal for learning. *Second Language Research, 4,* 374-397.

Olive, T. (2004). Working memory in writing: Empirical evidence from the dual-task technique. *European Psychologist, 9*, 32-42.

Osaka, M., & Osaka, N. (1992). Language independent working memory as measured by Japanese and English Reading Span Test. *Bulletin of the Psychonomic Society, 30,* 287-289.

Ortega, L. (2009). *Understanding second language acquisition*. London: Hodder Education.

Paradis, M. (2009). *Declarative and procedural determinants of second languages.* Amsterdam: John Benjamins.

Phoocharoensil, S., Moore, B., Gampper, C., Geerson, E., Chaturongakul, P., Sutharoj, S., & Carlon, W. (2016). Grammatical and lexical errors in low-proficiency Thai graduate students' writing. *Language Education and Acquisition Research Network Journal, 9* (1), 11-24.

Pica, T. (1991). Foreign language classrooms: Making them research-ready and researchable. In B. Freed (Ed.), *Foreign language acquisition research and the classroom* (pp. 393-412). Lexington, MA: D.C. Heath.

Pica, T. (1994). Research on negotiation: What does it reveal about second-language learning conditions, processes, and outcomes? *Language Learning, 44*, 493-527.

Pienemann, M. (1989). Is language teachable? *Applied Linguistics, 10* (1), 52-79.

Pienemann, M. (1998). *Language processing and second language development: Processability theory*. Amsterdam: Benjamins.

Pienemann, M. (2007). Processability theory. In B. VanPatten & J. Williams (Eds.), *Theories in second language acquisition: An introduction* (pp. 137-154). Mahwah, NJ: Erlbaum.

Pienemann, M., & Lenzing, A. (2015). Processability theory. In B. VanPatten & J. Williams (Eds.), *Theories in second language acquisition: An introduction* (pp. 159-179). New York; London: Routledge.

Plonsky, L., & Oswald, F. (2014). How big is "big"? Interpreting effect sizes in L2 research. *Language Learning*, *64*, 878-912.

Plonsky, L., Marsden, E., Crowther, D., Gass, S., & Spinner, P. (2020). A methodological synthesis and meta-analysis of judgment tasks in second language research. *Second Language Research, 36* (4), 583-621.

Polio. C. (2012). The relevance of second language acquisition theory to the written error correction debate. *Journal of Second Language Writing, 21*, 375-389.

Pongsiriwet, C. (2001). Relationships among grammatical accuracy, discourse features, and the quality of second language writing: The case of Thai EFL learners. Doctoral thesis. West Virginia University, USA.

Qi, D., & Lapkin, S. (2001). Exploring the role of noticing in a three-stage second language writing task. *Journal of Second Language Writing, 10*, 277-303.

Reber, A. (1967). Implicit learning of artificial grammars. *Journal of Verbal Learning and Verbal Behavior, 6*, 317-327.

Rebuschat, P. (2008) Implicit learning of natural language syntax. PhD thesis, Cambridge University.

Rebuschat, P. (2013). Measuring implicit and explicit knowledge in second language research. *Language Learning, 63* (3), 595-626.

Rebuschat, P. (2015). Introduction: Implicit and explicit learning of languages. P. Rebuschat (Ed.) *Implicit and explicit learning of languages* (pp. 13-22). Amsterdam: John Benjamins.

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

Rebuschat, P., Hamrick, P., Riestenberg, K., Sachs, R., & Ziegler, N. (2015). Triangulating measures of awareness. *Studies in Second Language Acquisition, 37*, 299-334.

Révész, A. (2011). Task complexity, focus on L2 constructions and individual differences: A classroom-based study. *The Modern Language Journal, 95,* 162-

181.

Révész, A. (2012). Working memory and the observed effectiveness of recasts on different L2 outcome measures. *Language Learning, 62*, 93-132.

Révész, A., Michel, M., & Lee, M. (2017) Investigating IELTS academic writing task 2: relationships between cognitive writing processes, text quality, and working memory. IELTS Research Report. London: The British Council.

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

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

Roberts, L. (2012) Individual differences in second language sentence processing. *Language Learning, 62*, 172-188.

Roberts, L., & Liszka, S. (2013). Processing tense/aspect-agreement violations on-line in the second language: A self-paced reading study with French and German L2 learners of English. *Second Language Research*, *29*, 413-439.

Robinson, P. (1997). Individual differences and the fundamental similarity of implicit and explicit adult second language learning, *Language Learning, 47* (1), 45-99.

Robinson, P. (2002). Learning conditions, aptitude complexes and SLA: A framework for research and pedagogy. In P. Robinson (Ed.), *Individual differences and instructed language learning* (pp. 113-135). Philadelphia/Amsterdam: John Benjamins.

Robinson, P. (2005). Cognitive abilities, chunk-strength, and frequency effects in implicit artificial grammar and incidental L2 learning: Replications of Reber, Walkenfeld, and Hernstadt (1991) and Knowlton and Squire (1996) and their relevance for SLA. *Studies in Second Language Acquisition*, *27*, 235-268.

Robinson, P. (2012). Individual differences, aptitude complexes, SLA processes, and aptitude test development. In M. Pawlak (Ed.), *New perspectives on individual differences in language learning and teaching* (pp. 57-75). Berlin Heidelberg: Springer-Verlag.

Robinson, P., & Ellis, N. (2008). Conclusion: Cognitive linguistics, second language acquisition and L2 instruction-Issues for research. In P. Robinson & N. Ellis (Eds.), *Handbook of cognitive linguistics and second language acquisition* (pp. 489-545). London: Routledge.

Rogers, J., Révész, A., & Rebuschat, P. (2016). Implicit and explicit knowledge of inflectional morphology. *Applied Psycholinguistics, 37*, 781-812.

Rosa, E., & Leow, R. (2004). Awareness, different learning conditions, and second language development. *Applied psycholinguistics, 24*, 269-292.

Ruiz, S., Rebuschat, P., & Meurers, D. (2019). The effects of working memory and declarative memory on instructed second language vocabulary learning: Insights from intelligent CALL. *Language Teaching Research, 3* (1), 1-30.

Ruiz, S., Tagarelli, K., & Rebuschat, P. (2018). Simultaneous acquisition of words and syntax: Effects of exposure condition and declarative memory. *Frontiers in Psychology*, *9* (1168), 1-11.

Rummel, S. (2014). Student and teacher beliefs about written CF and the effect these beliefs have on uptake: A multiple case study of Laos and Kuwait. Unpublished doctoral thesis, AUT University, Auckland, New Zealand.

Russell, 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* (pp. 131-164). Amsterdam: Benjamins.

Sachs, R., & Polio, C. (2007). Learners' use of two types of written feedback on a L2 writing task. *Studies in Second Language Acquisition*, *29*, 67-100.

Saeb, F. (2014). The effects of focused and unfocused written corrective feedback on the grammaticality accuracy of beginner EFL learners. *International Journal of Applied Linguistics & English Literature, 3* (2), 22-26.

Sagarra, N. (2007). From CALL to face-to-face interaction: The effect of computer-delivered recasts and working memory on L2 development. In A. Mackey (Ed.), *Conversational interaction in second language acquisition: A collection of empirical studies* (pp. 229-248). Oxford: Oxford University Press.

Sagarra, N. (2008). Working memory and L2 processing of redundant grammatical forms. In Z. Han (Ed.) *Understanding second language process* (pp. 133-147). Clevedon: Multilingual Matters.

Sagarra, N. (2012). Working memory in second language acquisition. In C. Chapelle (Ed.), *The Encyclopedia of Applied Linguistics* (pp.1-8). Oxford: Wiley Blackwell.

Sagarra, N. (2013). Optimizing the noticing of recasts via computer-delivered feedback: Evidence that oral input enhancement and working memory help second language learning. *The Modern Language Journal, 97* (1), 196-216.

Sagarra, N. (2017). Longitudinal effects of working memory on L2 grammar and reading abilities. *Second Language Research, 33* (3), 341-363.

Sagarra, N., & Abbuhl, R. (2013). Optimizing the noticing of recasts via computer-delivered feedback: Evidence that oral input enhancement and working memory help second language learning. *The Modern Language Journal, 97* (1), 196-216.

Saito, K., & Lyster, R. (2012). Effects of form-focused instruction and corrective feedback on L2 pronunciation development of /ɹ/ by Japanese learners of English. *Language Learning, 62*, 595-633.

Santos, M., López-Serrano, S., & Manchón, R. (2010). The differential effect of two types of direct written corrective feedback on noticing and uptake: Reformulation vs. error correction. *IJES - International Journal of English Studies*, *10*, 131-154.

Sanz, C., Lin, H., Lado, B., Stafford, C., & Bowden, H. (2016). One size fits all? Learning conditions and working memory in ab initio language development. *Applied Linguistics, 37* (5), 669-692.

Sarandi, H. (2015). Reexamining elicited imitation as a measure of implicit grammatical knowledge and beyond ...? *Language Testing, 32* (4), 485-501.

Sato, M., & Loewen, S. (2018). Metalinguistic instruction enhances the effectiveness of corrective feedback: Variable effects of feedback types and linguistic targets. *Language Learning, 68* (2), 507-545.

Schachter, J. (1988). Second language acquisition and its relationship to universal grammar. *Applied Linguistics, 9,* 219-235.

Schachter, J. (1991). Issues in the Accessibility Debate: A reply to Felix. In L. Eubank (Ed.), *Point counterpoint: Universal grammar in the second language* (pp.105-116). Amsterdam: John Benjamins.

Schmidt, R. (1990). The role of consciousness in second language learning. *Applied Linguistics, 11*(2), 129-158.

Schmidt, R. (1992). Psychological mechanisms underlying second language fluency, *SSLA, 14,* 357-385.

Schmidt, R. (1994). Deconstructing consciousness in search of useful definitions for applied linguistics. *AILA Review, 11*, 11-26.

Schmidt, R. (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: Hawaii: University of Hawaii, Second Language Teaching and Curriculum Center.

Schmidt, R. (2001). Attention. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 3-32). New York: Cambridge University Press.

Schmidt, R., & Watanabe, Y. (2001). Motivation, strategy use, and pedagogical preferences in foreign language learning. In Z. Dörnyei & R. Schmidt (Eds.), *Motivation and second language acquisition* (pp. 313-359). Honolulu: University of Hawai'i Second Language Teaching and Curriculum Center.

Schwartz, B. (1993). On explicit and negative data effecting and affecting competence and linguistic behavior. *Studies in Second Language Acquisition, 15*, 147-163.

Schwartz, B., & Sprouse, R. (1996). L2 cognitive states and the full transfer/full access model. *Second Language Research*, *12*, 40-72.

Selinker, L. (1972). Interlanguage. *IRAL: International Review of Applied Linguistics in Language Teaching, 10*, 209-241.

Selinker, L., & Lakshmanan, U. (1992). Language transfer and fossilization: The multiple effects principle. In S. Gass & L. Selinker (Eds.), *Language Transfer in Language Learning* (pp. 47-70). Amsterdam: John Benjamins.

Semke, H. (1984). The effects of the red pen. *Foreign Language Annuals, 17* (3), 195-202.

Service, E. (1992). Phonology, working memory, and foreign-language learning. *The Quarterly Journal of Experimental Psychology, 45*(1), 21-50.

Serafini, E., & Sanz, C. (2016) Evidence for the decreasing impact of cognitive ability on second language development as proficiency increases. *Studies in Second Language Acquisition, 38*, 607-646.

Sheen, Y. (2004). Corrective feedback and learner uptake in communicative classrooms across instructional settings. *Language Teaching Research*, *8* (3), 263-300.

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

Sheen, Y. (2010). The role of oral and written corrective feedback in SLA. *Studies in Second Language Acquisition, 32*, 169-179.

Sheen, Y., & Ellis, R. (2011). Corrective feedback in language teaching. In E. Hinkel (Ed.), *Handbook of research in second language teaching and learning* (pp. 593-610). New York; London: Routledge.

Sheen, Y., Wright, D., & Moldawa, A. (2009). Differential effects of focused and unfocused written correction on the accurate use of grammatical forms by adult ESL learners. *System, 37* (4), 556-569.

Sheppard, K. (1992). Two feedback types: Do they make a difference? *RELC Journal, 23* (1), 103-110.

Shintani, N. (2016). The effects of computer-mediated synchronous and asynchronous direct corrective feedback on writing: A case study. *Computer Assisted Language Learning, 29*, 517-538.

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

Shintani, N., Ellis, R., & Suzuki, W. (2014). Effects of Written Feedback and Revision on Learners' Accuracy in Using Two English Grammatical Structures. *Language Learning, 64* (1), 103-131.

Shooshtari, Z., Vahdat, S., & Negahi, M. (2019). The effect of unfocused direct and indirect written corrective feedback on the implicit and explicit grammatical knowledge and writing accuracy of EFL learners. *Applied Linguistics Research Journal, 3* (5), 32-48.

Simard, D., Guénette, D., & Bergeron, A. (2015). L2 learners' interpretation and understanding of written corrective feedback: Insights from their metalinguistic reflections. Language Awareness, 24, 233-254.

Skehan, P. (1998). *A cognitive approach to language learning.* Oxford: Oxford University Press.

Skehan, P. (2002). Theorizing and updating aptitude. In P. Robinson (Ed.), *Individual Differences and Instructed Language Learning* (pp. 69-94). John Benjamins.

Skehan, P. (2015). Foreign language aptitude and its relationship with grammar: A critical overview. *Applied Linguistics, 36* (3), 367-384.

Skehan, P. (2016). Foreign language aptitude, acquisitional sequences, and psycholinguistic processes. In G. Granena, D. Jackson & Y. Yilmaz (Eds.), *Cognitive individual differences in L2 processing and acquisition* (pp.17-40). Amsterdam: John Benjamins.

Skehan, P., & Foster, P. (2001). Cognition and tasks. In P. Robinson (Ed.), *Cognition and second language instruction* (pp.183-205). Cambridge, UK: Cambridge University Press.

Slabakova, R. (2014). The bottleneck of second language acquisition. *Foreign Language Teaching and Research*, *46* (4), 543-559.

Spada, N. (1997). Form-focused instruction and second language acquisition: a review of classroom and laboratory research. *Language Teaching, 30*, 73-87.

Spada, N. (2011). Beyond form-focused instruction: Reflections on past, present, and future research. *Language Teaching*, *44*, 225-236.

Spada, N., Shiu, J., & Tomita, Y. (2015). Validating an elicited imitation task as a measure of implicit knowledge: Comparisons with other validation studies. *Language Learning, 65* (3), 723-751.

Spada, N., Jessop, L., Tomita, Y., Suzuki, W., & Valeo, A. (2014). Isolated and integrated form-focused instruction: Effects on different types of L2 knowledge.

Language Teaching Research, *18*, 453-473.

Speciale, G., Ellis, N., & Bywater, T. (2004). Phonological sequence learning and short-term store capacity determine second language vocabulary acquisition. *Applied Psycholinguistics*, 25 (2), 293-321.

Stavans, A., Seroussi, B., & Ehrlich, S. (2019). Literary-related abilities' effects on argumentative text quality structure. *Journal of Literacy Research, 5* (3), 315-335.

Stefanau, C. (2014). L2 general use for generic and specific plural reference: The role of written CF, learner factors and awareness. Unpublished doctoral thesis, Lancaster University, UK.

Stefanou, C., & Révész, A. (2015). Direct written corrective feedback, learner differences, and the acquisition of second language article use for generic and specific plural reference. *Modern Language Journal, 99*, 263-282.

Storch, N. (2010). Critical feedback on written corrective feedback research. *International Journal of English Studies, 10,* 29-46.

Storch, N. (2018). Written corrective feedback from sociocultural theoretical perspectives: A research agenda, *Language Teaching, 51* (2), 262-277.

Storch, N., & Wigglesworth, G. (2010). Learners' processing, uptake, and retention of corrective feedback on writing: Case studies. *Studies in Second Language Acquisition, 32* (2), 303-334.

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

Suzuki, Y. (2017). Validity of new measures of implicit knowledge: Distinguishing implicit knowledge from automatized explicit knowledge. *Applied Psycholinguistics, 38* (5), 1229-1261.

Suzuki, Y., & DeKeyser, R. (2015). Comparing elicited imitation and word monitoring as measures of implicit knowledge. *Language Learning*, *65*, 860-895.

Suzuki, Y., & DeKeyser, R. (2017). The interface of explicit and implicit knowledge in a second language: Insights from differences in cognitive aptitudes. *Language Learning, 67* (4), 747-790.

Suzuki, Y., Nakata, T., & DeKeyser, R. (2019). Optimizing second language practice in the classroom: Perspectives from cognitive psychology. *The Modern Language Journal, 103* (3), 551-561.

Swain, M. (1984). A review of immersion education in Canada Research and evaluation studies, *ELT Documents, 119*, 35-51.

Swain, M. (1985). Communicative Competence: Some roles of comprehensible input and comprehensible output in its development. In S. Gass & C. Madden (Eds.), *Input in second language acquisition* (pp. 235-253). Rowley, MA: Newbury House.

Swain, M. (1993). The Output Hypothesis: Just speaking and writing are not enough. *The Canadian Modern Language Review, 50* (1), 158-164.

Swain, M. (1995). Three functions of output in second language learning. In G. Cook & B. Seidlhofer (Eds.), *Principle and practice in applied linguistics: Studies in honor of H. G. Widdowson* (pp.125-144). Oxford: Oxford University Press.

Swain, M. (1998). Focus on Form through Conscious Reflection. In C. Doughty & J. Williams (Eds.), *Focus on form in classroom second language acquisition* (pp. 64-81). Cambridge: Cambridge University Press.

Swain, M. (2005). The output hypothesis: Theory and research. In E. Hinkel (Ed.), *Handbook on research in second language teaching and learning* (pp. 471-83). Mahwah, NJ: Lawrence Erlbaum Associates.

Swain, M., & Lapkin, S. (1982). *Evaluating bilingual education: A Canadian case study.* Clevedon, UK: Multilingual Matters.

Swain, M., & Lapkin, S. (1986). Immersion French at the secondary level: "The goods" and "the bads." *Contact*, 2-9.

Swain, M., & Lapkin, S. (1998). Interaction and second language learning: Two adolescent French immersion students working together. *Modern Language Journal, 82*, 320-337.

Swain, M., & Lapkin, S. (2002). Talking it through: Two French immersion learners' responses to reformulation. *International Journal of Educational Research, 37*, 285-304.

Taber, K. (2018). The use of Cronbach's Alpha when developing and reporting research instruments in science education. *Research in Science Education, 48*, 1273-1296.

Tagarelli, K., Mota, M., & Rebuschat, P. (2015). Working memory, learning conditions, and the acquisition of L2 syntax. In Z. Wen, M. Borges Mota & A. McNeill (Eds.), *Working memory in second language acquisition and processing: Theory, research and commentary* (pp. 224-247). Buffalo, NY: Multilingual Matters.

Tan, K., & Manochpinyo, A. (2017). Improving grammaticality accuracy in Thai learners' writing: Comparing direct and indirect written corrective feedback. *The Journal of Asia TEFL, 14* (3), 430-442.

Tarone, E. (2014). Enduring questions from the Interlanguage Hypothesis. In Z. Han, & E. Tarone (Eds.), *Interlanguage: Forty years later* (pp.7-26). Amsterdam: John Benjamins.

Tavakol, M., & Dennick, R. (2011). Making sense of Cronbach's Alpha. *International Journal of Medical Education, 2*, 53-55.

Thep-Ackrapong, T. (2005). Teaching English in Thailand: An uphill battle. *Manutsart Paritat: Journal of Humanities (Srinakarinwirote University), 27* (1), 51-62.

Tomita, Y., Suzuki, W., & Jessop, L. (2009). Elicited imitation: Toward valid procedures to measure implicit second language grammatical knowledge. *TESOL Quarterly, 43* (2), 345-350.

Trahey, M., & White, L. (1993). Positive evidence and preemption in the second language classroom. *Studies in Second Language Acquisition, 15*, 181-204.

Trofimovich, P., Ammar, A., & Gatbonton, E. (2007). How effective are recasts? The role of attention, memory and analytical ability. In A. Mackey (Ed.), *Conversational interaction in second language acquisition* (pp.171-195). New York: Oxford University Press.

Truscott, J. (1996). The case against grammar correction in L2 writing classes. *Language Learning, 46* (2), 327-369.

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

Truscott, J. (2001). Selecting errors for selective error correction. *Studies in English Literature and Linguistics, 27* (2), 93-108.

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

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

Truscott, J. (2010). Further thoughts on Anthony Bruton's critique of the correction debate. *System, 38* (4), 626-633.

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

Turner, M., & Engle, R. (1989). Is working memory capacity task dependent? *Journal of Memory & Language*, *28,* 127-154.

Ullman, M. (2001). The neural basis of lexicon and grammar in first and second language, the declarative/procedural model. *Bilingualism: Language and Cognition, 4,* 105-122.

Ullman, M. (2004). Contributions of memory circuits to language: The declarative/procedural model. *Cognition, 92*, 231-270.

Ullman, M. T. (2005). A cognitive neuroscience perspective on second language acquisition: The declarative/procedural model. In C. Sanz (Ed.), *Mind and context in adult second language acquisition: Methods, theory and practice* (pp. 141-178). Washington, DC: Georgetown University Press.

Ullman, M. (2015). The declarative/procedural model: A neurobiologically motivated theory of first and second language. In B. VanPatten & J. Williams (Eds.), *Theories in second language acquisition: An introduction* (pp.135-160). New York; London: Routledge.

Ullman, M. (2016). The declarative/procedural model: A neurobiological model of language learning, knowledge and use. In G. Hickok & S. Small (Eds.), *Neurobiology of language* (pp. 953-968). New York: Elsevier Science and Technology.

Ullman, M., & Lovelett, J. (2018). Implications of the declarative/procedural model for improving second language learning: The role of memory enhancement techniques. *Second Language Research, 34* (1), 39-65.

Unsworth, N., Schrock, J. C., & Engle, R. W. (2004). Working memory capacity and the antisaccade task: Individual differences in voluntary saccade control. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 30*, 1302-1321.

Unsworth, N., Heitz, R., Schrock, J., & Engle, R. (2005). An automated version of the operation span task. *Behavior Research Methods, 37,* 498-505.

Vafaee, P., Suzuki, Y., & Kachisnke, I. (2017). Validating grammaticality judgment tests: Evidence from two new psycholinguistic measures. *Studies in Second Language Acquisition, 39*, 59-95.

Van Beuningen, C. (2010). Corrective feedback in L2 writing: Theoretical perspectives, empirical insights, and future directions. *International Journal of English Studies, 10* (2), 1-27.

Van Beuningen, C., De Jong, N., & Kuiken, F. (2008). The effect of direct and indirect corrective feedback on L2 learners' written accuracy. *ITL-Review of Applied Linguistics, 156*, 279-296.

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

Verhagen, J., & Leseman, P. (2016). How do verbal short-term memory and working memory relate to the acquisition of vocabulary and grammar? A comparison between first and second language learners. *Journal of Experimental Child Psychology, 141*, 65-82.

Vyatkina, N. (2010). The effectiveness of written corrective feedback in teaching beginning German. *Foreign Language Annals*, *43*, 671-689.

Vygotsky, L. (1978). *Mind in Society: The Development of Higher Psychological Processes*. Cambridge, MA: Harvard University Press.

Vygotsky, L.S. (1981). The genesis of higher mental functions. In J. Wertsch (Ed.), *The Concept of Activity in Soviet Psychology* (pp.144-188). Armonk, NY: M.E. Sharpe.

Wagner, J., & Wulf, D. (2016). Understanding written corrective feedback in second language grammar acquisition. *Journal of Education and Learning, 5,* 259-277.

Walker, N., Monaghan, P., Schoetensack, C., & Rebuschat, P. (2020). Distinctions in the acquisition of vocabulary and grammar: An individual differences approach. *Language Learning, 70* (2), 221-254.

Walter, C. (2004). Transfer of reading comprehension skills to L2 is linked to mental representations of text and to L2 working memory. *Applied Linguistics, 25*, 315-339.

Wang, W., & Li, S. (2021). Corrective feedback and learner uptake in American ESL and Chinese EFL classrooms: A comparative study. *Language, Culture and Curriculum, 34* (1), 35-50.

Watcharapunyawong, S., & Usaha, S. (2013). Thai EFL students' writing errors in different text types: The interference of first language. *English Language Teaching, 6* (1), 67-78.

Waters, G., & Caplan, D. (2003). The reliability and stability of verbal working memory measures. *Behavior Research Methods, Instruments, & Computers, 35*, 550-564.

Wechsler, D. (1997) *The Wechsler Adult Intelligence Scale: Revised.* New York: The Psychological Corporation.

Wen, Z. (2016). Phonological and executive working memory in L2 task-based speech planning and performance. *The Language Learning Journal, 44* (4), 418-435.

Wen, Z., Mota, M., & McNeill, A. (2015). *Working memory in second language acquisition and processing*. Bristol: Multilingual Matters.

White, L. (2003). *Second language acquisition and universal grammar*. New York: Cambridge University Press.

White, L. (2015). Linguistic theory, universal grammar, and second language acquisition. In B. VanPatten & J. Williams (Eds.), *Theories in second language acquisition: An introduction* (pp. 34-53). New York; London: Routledge.

White, L. (2018). Universal grammar. In J. Liontas & M. DelliCarpini (Eds.), *The TESOL encyclopedia of English language teaching* (pp.1-6). Hoboken, NJ, USA: John Wiley & Sons, Inc.

Williams, J. (2012). The potential role (s) of writing in second language development. *Journal of Second Language Writing, 21* (4), 321-331.

Williams, J. N. (2009). Implicit learning in second language acquisition. In W. Ritchie & T. Bhatia (Eds.), *The new handbook of second language acquisition* (pp. 319-353). Bingley, UK: Emerald Press.

Williams, J. N. (2012). Working memory and SLA. In P. Robinson, A. Mackey, S. Gass. & R. Schmidt. (Eds.), *The Routledge handbook of second language acquisition* (pp. 427-441). New York: Routledge.

Williams, J. N., & Lovatt, P. (2003). Phonological memory and rule learning. *Language Learning*, *53*, 67-121.

Wills, A. (2005). *New directions in human associative learning*. Mahwah, NJ: Erlbaum.

Wu, S., & Ortega, L. (2013). Measuring global oral proficiency in SLA research: A new elicited imitation test of L2 Chinese. *Foreign Language Annals*, *46*, 680-704.

Wulff, S., & Ellis, N. (2018). Usage-based approaches to second language acquisition. In D. Miller, F. Bayram, J. Rothman & L. Serratrice (Eds.), *Bilingual cognition and language: The state of the science across its subfields* (pp.37-56). Amsterdam; Philadelphia: John Benjamins Publishing Company.

Yan, X., Maeda, Y., Lv, J., & Ginther, A. (2016). Elicited imitation as a measure of second language proficiency: A narrative review and meta-analysis. *Language Testing, 33* (4), 497-528.

Yang, L., & Zhang, L. (2010). Exploring the role of reformulations and a model text in EFL students' writing performance. *Language Teaching Research, 14* (4), 464-484.

Yang, Y., Shintani, N., Li, S., & Zhang, Y. (2017). The effectiveness of post-reading word-focused activities and their associations with working memory. *System*, *70*, 38-49.

Yilmaz, Y. (2013). Relative effects of explicit and implicit feedback: The role of working memory capacity and language analytic ability. *Applied Linguistics, 34* (3), 344-368.

Yilmaz, Y., & Granena, G. (2019). Cognitive individual differences as predictors of improvement and awareness under implicit and explicit feedback conditions. *The Modern Language Journal, 103* (3), 686-702.

Zabihi, R. (2018). The role of cognitive and affective factors in measures of L2 writing. *Written Communication, 35* (1), 32-57.

Zhang, R. (2015). Measuring university-level L2 learners' implicit and explicit linguistic knowledge. *Studies in Second Language Acquisition, 37* (3), 457-486.

APPENDICES
Appendix A:
Informed Consent Documentation
<u>PARTICIPANT INFORMATION SHEET</u>

My name is Miss Ornuma Chingchit and I am a current PhD student from the department of Linguistics and English Language, Lancaster University, UK. I would like to invite you to take part in my research study which is about – (1) how writing affects second language development, and (2) how learner differences in their working memory capacity affect their second language learning outcomes.

Please take time to read the following information carefully before you decide whether or not you wish to take part in this study.

This study aims to (1) investigate the effect of writing on second language development, and (2) examine the extent to which learners' working memory capacity mediates learners' learning outcomes.

I have approached you because I am interested in understanding how writing could help Thai learners improve their second language learning and how Thai learners' working memory capacity mediates their learning outcomes. I would be very grateful if you would agree to take part in this study. However, it is completely up to you to decide whether or not you will take part in this study. Your participation is completely voluntary. **If you decide not to take part in this study, this will not affect your study, the way you are assessed on this course or the relationship with the teacher in anyway.**

I**f you agree to participate and want to withdraw later, I will extract any information (i.e., data) you contribute to the study and destroy it immediately. Nevertheless, it is difficult and often impossible to take out data from one specific participant when the data has already been anonymized or pooled together with other people's data. Therefore, you can only withdraw up to 5 weeks after taking part in the study.**

It should be acknowledged here that only I, the researcher conducting this study, my supervisor and another research assistant would have access to the information you share with me. I will keep all your personal information (e.g., your name and other

information about you that can identify you) confidential that is I will not share it with others apart from the ones mentioned.

In addition, I will use the information you have shared with me only in the following ways: my PhD thesis and other relevant publications, for example journal articles. I may also present the results of this study at academic conferences.

If you have any queries or you are unhappy with anything regarding your participation, please contact me at the following email address: o.chingchit@lancaster.ac.uk**. Also,** if you are interested in obtaining the study's results, please feel free to contact me once the study is completed.

Thank you for considering your participation in this study.

CONSENT FORM

Project Title: Written Corrective Feedback, Working Memory and the Development of Explicit and Implicit Knowledge of English Plurals
Name of Researcher: Ornuma Chingchit
Email: o.chingchit@lancaster.ac.uk

Please tick each box

1. I confirm that I have read and understand the information sheet for the above study. I have had the opportunity to consider the information, ask questions and have had these answered satisfactorily.	☐
2. I understand that my participation is voluntary and that I am free to withdraw at any time during my participation in this study and within 5 weeks after taking part in the study, without giving any reason. If I withdraw within [5 weeks] of taking part in the study, my data will be removed.	☐
3. I understand that any information disclosed within my group remains confidential to the group, and I will not discuss any information about my group with anyone who is not involved.	☐
4. I understand that any information given by me may be used in the researcher's PhD thesis, future reports, academic articles, publications or presentations, but my personal information will not be included, and I will not be identifiable.	☐
5. I understand that my name/my organisation's name will not appear in any thesis, report, article or presentation without my consent.	☐
6. I understand that any audio-recorded information will be transcribed, and the data will be protected on encrypted devices and kept secure.	☐
7. I understand that data will be kept according to University guidelines for a minimum of 10 years after the end of the study.	☐
8. I agree to take part in the above study.	☐

_____ _____ _____
Name of Participant   Date   Signature

I confirm that the participant was given an opportunity to ask questions about the study, and all the questions asked by the participant have been answered correctly and to the best of my ability. I confirm that the individual has not been coerced into giving consent, and the consent has been given freely and voluntarily.

Signature of Researcher /person taking the consent_____ **Date** _____ Day/month/year

EXPERIMENT DEBRIEF INFORMATION

First of all, thank you for participating in this study. This debriefing form is given to inform you some information about this experimental study. Please feel free to ask any questions or to comment on any aspect of the study.

This study is designed to (1) examine how writing practice affects learners' second language development, and (2) investigate the extent to which learners' working memory capacity mediates learners' learning outcomes. This study is experimental in nature. Participants will be randomly assigned to different experimental groups. At the beginning of the experiment, participants will be asked to complete a battery of English and working memory tests. After that, for 6 weeks, participants in each group will receive different writing practice. After the 6-week treatment, participants will complete the battery of English tests once again. The scores of the participants' pre, post and delayed posttests will then be compared and further analyzed.

As you know, your participation in this study is voluntary. If you so wish, you may withdraw after reading this debriefing form or up to 5 weeks after the participation, at which point all records of your participation will be destroyed. You will not be penalized if you withdraw. However, if you decide to participate, you can rest assured that all your personal data as well as responses will be absolutely confidential. Your name will be converted to a code number, and only people who are associated with this research will see your name or your responses. In return, it will be greatly appreciated if you could maintain the research's confidentiality by not sharing any detail of this study to other students. That is because if other students know about the study before they participate, their data will be biased and thus cannot be included.

Once again, your participation in this study is greatly appreciated. If you'd be interested in obtaining a copy of the results once the study is complete, you may contact the researcher at o.chingchit@lancaster.ac.uk. If you are interested in this area of research, you may wish to consult the following introductory sources of references:

Ellis, R., Loewen, S., Elder, C., Erlam, R., Philp, J., & Reinders, H. (2009). Implicit and
     explicit knowledge in second language learning, testing and teaching. Bristol, UK;

Buffalo, USA; Toronto, Canada: Multilingual Matters

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.

Rebuschat, P. (2013). Measuring implicit and explicit knowledge in second language
research. *Language Learning*, 63 (3), 595-626.

Wen, Z., Mota, M., & McNeill, A. (2015). Working memory in second language
acquisition and processing. Bristol, UK; New York, USA, Ontario, Canada:
Multilingual Matters.

This study has been granted clearance according to the recommended principles of
Lancaster University's ethics guidelines and policies. If you have any complaints,
concerns, or questions about this research, please feel free to directly contact the
researcher's supervisor Dr. Jenefer Philp at j.philp@lancaster.ac.uk.

Thank you very much for your participation!!

Week 1: Argumentative Essay Writing Task

Instruction: Write an argumentative essay responding to the question below. You have **30 minutes** to plan, write, and revise your essay. Use specific reasons and/or examples to support your answer. An effective essay contains a minimum of 300 words.

Should richer people pay more taxes?

Week 2: Argumentative Essay Writing Task

Instruction: Write an argumentative essay responding to the question below. You have **30 minutes** to plan, write, and revise your essay. Use specific reasons and/or examples to support your answer. An effective essay contains a minimum of 300 words.

Professional athletes and celebrities earn more money than other professions. Is this justified?

Week 3: Argumentative Essay Writing Task

Instruction: Write an argumentative essay responding to the question below. You have **30 minutes** to plan, write, and revise your essay. Use specific reasons and/or examples to support your answer. An effective essay contains a minimum of 300 words.

Should Thailand make English another official language?

Week 4: Argumentative Essay Writing Task

Instruction: Write an argumentative essay responding to the question below. You have **30 minutes** to plan, write, and revise your essay. Use specific reasons and/or examples to support your answer. An effective essay contains a minimum of 300 words.

Should smoking lounges or smoking areas be eliminated from airports?

Week 5: Argumentative Essay Writing Task

Instruction: Write an argumentative essay responding to the question below. You have **30 minutes** to plan, write, and revise your essay. Use specific reasons and/or examples to support your answer. An effective essay contains a minimum of 300 words.

Should Thailand give asylum to refugees/asylum seekers?

Week 6: Argumentative Essay Writing Task

Instruction: Write an argumentative essay responding to the question below. You have **30 minutes** to plan, write, and revise your essay. Use specific reasons and/or examples to support your answer. An effective essay contains a minimum of 300 words.

Has social media changed us for the worse?

Appendix C:

Background and Exit Questionnaire

<u>BACKGROUND QUESTIONNAIRE</u>

Name-Surname: _____

Gender: _____   Native Language: _____

Date of Birth: _____ Age: _____

Major: _____   Year: _____

English National Entrance Exam Scores (ONET scores)_____

Other English Test Scores [e.g., TOEIC, TOEFL, IELTS, etc. (if there is any)]:

_____

Other Foreign Language Background [e.g., French, Chinese, Japanese, etc. (if there is any)]:

_____

Number of Years of Formal English Language Education (from kindergarten or primary school to

university): _____

Abroad Education Experience (if there is any, please state the country and duration of your stay, for

example, America -3 months-):

_____

Phone: _____   E-mail: _____

EXIT QUESTIONNAIRE

1) What aspect (s) of writing usually concern you the most when writing? (e.g., content, style, organization, vocabulary, grammar, spelling, punctuation etc.) Why?

_____

_____

2) How do you feel about writing in general? Is it beneficial for your English language learning/knowledge development?

_____

_____

_____

3) Has your writing improved during the past months? If it has improved, how? Please explain.

_____

_____

_____

4) What aspect (s) of writing do you want to improve and want your teacher to help with the most in class? (e.g., content, style, organization, vocabulary, grammar, spelling, punctuation etc.)

_____

_____

5) What do you think is the main purpose/goal of this research study?

_____

_____

TIMED ARGUMENTATIVE ESSAY WRITING TEST

PRETEST

Instruction: Write an argumentative essay responding to the question below. You have **30 minutes** to plan, write, and revise your essay. Use specific reasons and/or examples to support your answer. An effective essay contains a minimum of 300 words.

University degree: is it necessary for success?

POSTTEST

Instruction: Write an argumentative essay responding to the question below. You have **30 minutes** to plan, write, and revise your essay. Use specific reasons and/or examples to support your answer. An effective essay contains a minimum of 300 words.

Should students blame their teachers if they have low test scores?

Instruction: Write an argumentative essay responding to the question below. You have **30 minutes** to plan, write, and revise your essay. Use specific reasons and/or examples to support your answer. An effective essay contains a minimum of 300 words.

Should students wear school uniforms?

OEIT Instruction sheet

Instruction: In this task, you will hear some statements in English. You will hear each statement only once. Listen to the statement, then decide whether the content of the statement you hear is a real fact or not.

If the statement represents the real fact, say "True" in English and circle "True" on your answer sheet. If you think the statement is not true, say "False" and circle "False" on your answer sheet.

For example, after hearing the sentence, "America and England are neighboring countries", you should say "False", as in reality America and England are not neighboring countries.

After you answer "True" or "False" and circle the answer on your answer sheet, please say the statement again in correct English.

OEIT Answer Sheet

1.   True   False

2.   True   False

3.   True   False

4.   True   False

5.   True   False

6.   True   False

7.   True   False

8.   True   False

9.   True   False

10.   True   False

11.   True   False

12.   True   False

13.   True   False

14.   True   False

15.   True   False

16.   True   False

17.   True   False

18.   True   False

19.   True   False

20.   True   False

21.   True   False

22.   True   False

UNTIMED GRAMMATICALITY JUDGMENT TEST (UGJT)

UGJT: FORM A

Instruction:
(1) Judge whether the sentences are grammatically correct or incorrect. Tick the "Correct" box if the sentence is grammatically correct and tick the "Incorrect" box if it is incorrect.
(2) If the sentence is grammatically incorrect, underline the part that makes the sentence incorrect and substitute it with the correct one.

Example: I haven't seen him for a long time = correct
He has been living in Australia since three years = incorrect
I have been studying English since a long time = incorrect
Are you going to London? = correct
Is Fai teach yesterday? = incorrect

Item
1. Mai completed her assignment and print it out. ☐ Correct   ☐ Incorrect
2. They would like to have Dr. Grey as their biology teacher. ☐ Correct   ☐ Incorrect
3. The teacher explained the problem to the students. ☐ Correct   ☐ Incorrect
4. Ploy wanted to know who Mandy went out with earlier. ☐ Correct   ☐ Incorrect
5. She often watch comedy movies before going to bed. ☐ Correct   ☐ Incorrect

6. I would love to work every day from noon to midnight.

☐ Correct   ☐ Incorrect

7. Peter bought his dinner and go straight back home.

☐ Correct   ☐ Incorrect

8. George flew to LA last night to meet up with his old friends.

☐ Correct   ☐ Incorrect

9. I think he has intelligent because he studies really hard.

☐ Correct   ☐ Incorrect

10. Dan's parents and friends usually visit him during long holidays.

☐ Correct   ☐ Incorrect

11. Mira always wanted to study medicine when she was young.

☐ Correct   ☐ Incorrect

12. Laura is currently living in Australia with her cousins.

☐ Correct   ☐ Incorrect

13. Clearly, Matteo must to speak Italian very well.

☐ Correct   ☐ Incorrect

14. Somebody in this class needs to volunteer for this assignment.

☐ Correct   ☐ Incorrect

15. When you miss someone badly, you can just video call them.

☐ Correct   ☐ Incorrect

16. There are many beautiful river and mountains in Northern Scotland.

☐ Correct   ☐ Incorrect

17. Each of them play soccer really well.

☐ Correct   ☐ Incorrect

18. She has always a diligent and devoting student.

☐ Correct   ☐ Incorrect

19. Mike might have to go to the library this afternoon.

☐ Correct   ☐ Incorrect

20. Technology has become one of the necessity in human life.

☐ Correct   ☐ Incorrect

21. John worked really hard but earn very little in the past.

☐ Correct   ☐ Incorrect

22. Jamie is tall but not quite smart.

☐ Correct   ☐ Incorrect

23. Having access to information technology are very important these days.

☐ Correct   ☐ Incorrect

24. Bill wanted to know where had I been for the last two nights.

☐ Correct   ☐ Incorrect

25. Everyone prefers their own way of living.

☐ Correct   ☐ Incorrect

26. Chompoo loves eating a lot of sushi.

☐ Correct   ☐ Incorrect

27. Mary got really sick after coming back from France.

☐ Correct   ☐ Incorrect

28. Sarah decided moving out of her old house.

☐ Correct   ☐ Incorrect

29. Fiona wants to further her study and needs advices on it.

☐ Correct   ☐ Incorrect

30. Arya often gets up late on weekend.

☐ Correct   ☐ Incorrect

31. George just came back from his exciting adventurous safari trip.

☐ Correct   ☐ Incorrect

32. Most people absolutely want to send us kids to best schools.

☐ Correct   ☐ Incorrect

33. Tim asked me why did I not go to work today.

☐ Correct   ☐ Incorrect

34. The toffee pudding at this shop always tastes amazingly delicious.

☐ Correct   ☐ Incorrect

35. Kelly wants to go to Portugal this weekend.

☐ Correct   ☐ Incorrect

36. The elephant is the kind of mammal.

☐ Correct   ☐ Incorrect

37. Joe's favorite drinks are red French wine and German beer.

☐ Correct   ☐ Incorrect

38. He expected sending a gift to his parents' anniversary.

☐ Correct   ☐ Incorrect

39. The student is working on his homework diligently.

☐ Correct   ☐ Incorrect

40. Sandy thought it good to move to Edinburgh.

☐ Correct   ☐ Incorrect

UGJT: FORM B

Instruction:

(1) Judge whether the sentences are grammatically correct or incorrect. Tick the "Correct" box if the sentence is grammatically correct and tick the "Incorrect" box if it is incorrect.

(2) If the sentence is grammatically incorrect, underline the part that makes the sentence incorrect and substitute it with the correct one.

Example: I haven't seen him for a long time = correct
He has been living in Australia since three years = incorrect
I have been studying English since a long time = incorrect
Are you going to London? = correct
Is Fai teach yesterday? = incorrect

Item
1. Facebook is one of many powerful tool available nowadays. ☐ Correct ☐ Incorrect
2. Ken's new job is teaching English to hill tribe children. ☐ Correct ☐ Incorrect
3. Everyone is exciting to see the new season of Game of Thrones. ☐ Correct ☐ Incorrect
4. I usually cooked spaghetti for dinner in the past. ☐ Correct ☐ Incorrect
5. He can fly an airplane. ☐ Correct ☐ Incorrect

6. I agree doing this cleaning job for him so that he can take a rest.

   ☐ Correct   ☐ Incorrect

7. Sally and I enjoyed our trip to Maldives.

   ☐ Correct   ☐ Incorrect

8. Dan will never visit those crowded tourist destination like Rome or Paris.

   ☐ Correct   ☐ Incorrect

9. Studying can be tiresome and boring from time to time.

   ☐ Correct   ☐ Incorrect

10. She explained why did he study German.

   ☐ Correct   ☐ Incorrect

11. Running make me feel happy.

   ☐ Correct   ☐ Incorrect

12. Laura loved eating tomatoes when she was young.

   ☐ Correct   ☐ Incorrect

13. Mathematics is a difficult subject for most female students.

   ☐ Correct   ☐ Incorrect

14. John's father really wants him to be successful.

   ☐ Correct   ☐ Incorrect

15. Some people like to work really hard while some enjoy doing nothing.

☐ Correct   ☐ Incorrect

16. A technology has both advantages and disadvantages.

☐ Correct   ☐ Incorrect

17. When I back from work, I take a shower and then watch a movie.

☐ Correct   ☐ Incorrect

18. Technology makes our life easier and helps save us times.

☐ Correct   ☐ Incorrect

19. The teacher explains the problem to the students.

☐ Correct   ☐ Incorrect

20. I must have to clean my room tonight.

☐ Correct   ☐ Incorrect

21. That cute little child is playing with its puppy.

☐ Correct   ☐ Incorrect

22. She asked him sending her a text message.

☐ Correct   ☐ Incorrect

23. Some people believes that exercising is a waste of time.

☐ Correct   ☐ Incorrect

24. Vegetables contain many vitamins and minerals necessary for human life.

☐ Correct   ☐ Incorrect

25. George wants to buy a new car.

☐ Correct   ☐ Incorrect

26. I hope being successful one day.

☐ Correct   ☐ Incorrect

27. It system is needed to be improved.

☐ Correct   ☐ Incorrect

28. Their great grandparents move to America in the 1980s.

☐ Correct   ☐ Incorrect

29. My first car is a white Honda Jazz.

☐ Correct   ☐ Incorrect

30. Katherine plays badminton very well.

☐ Correct   ☐ Incorrect

31. He told me he wanted to work harder.

   ☐ Correct   ☐ Incorrect

32. Martha and Jordan already booked their flight to Tokyo.

   ☐ Correct   ☐ Incorrect

33. Anna can be described as a sweetly and delightful girl at school.

   ☐ Correct   ☐ Incorrect

34. Rosemary reports the crime to the police.

   ☐ Correct   ☐ Incorrect

35. The boys bought the show tickets from an online website.

   ☐ Correct   ☐ Incorrect

36. Daniel might joining our trip to Phuket.

   ☐ Correct   ☐ Incorrect

37. Derek went to their teacher's retirement party with his friends last night.

   ☐ Correct   ☐ Incorrect

38. Tom asked whether was I going.

   ☐ Correct   ☐ Incorrect

39. Martin is living in a big white wooden house near the beach.

☐ Correct   ☐ Incorrect

40. Jonathan and I was best friends since fourth grade.

☐ Correct   ☐ Incorrect

MKT: FORM A

Instruction: In this part, there are 10 sentences. **These 10 sentences are all ungrammatical.** The part of the sentence that contains the error is underlined. For each sentence, if you know a rule that explains why the sentence is ungrammatical, write it in English in the space provided. If you do not know a rule, leave it blank and go on to the next sentence.

Here are some examples.

Example 1: I have lost <u>mine ring</u>.

_____

_____

Example 2: He saw <u>a elephant</u>.

_____

_____

Now start the test.

1. Sarah made <u>we</u> cry yesterday.

_____

_____

_____

2. Ken grew <u>many rose</u> in his garden.

_____

_____

_____

3. Because I was late, I called <u>taxi.</u>

_____

_____

_____

4. You <u>will to</u> wash your hands before meals.

_____

_____

_____

5. <u>Game online</u> is popular among teenagers and young adults.

_____

_____

_____

6. Technology <u>make</u> our lives easier.

_____

_____

_____

7. Ton likes <u>very much his current job</u>.

_____

_____

_____

8. Nowadays, I usually <u>watched</u> news on TV in the morning.

_____

_____

_____

9. I <u>like visit</u> my grandparents from time to time.

_____

_____

_____

10. They were interested in what <u>was I doing</u>.

_____

_____

_____

The End
Thank you ☺

MKT: FORM B

Instruction: In this part, there are 10 sentences. **<u>These 10 sentences are all ungrammatical.</u>** The part of the sentence that contains the error is underlined. For each sentence, if you know a rule that explains why the sentence is ungrammatical, write it in English in the space provided. If you do not know a rule, leave it blank and go on to the next sentence.

Here are some examples.

Example 1: I have lost <u>mine ring</u>.

_____

_____

Example 2: He saw <u>a elephant</u>.

_____

_____

Now start the test.

1. Currently, everyone prefers to communicate with <u>their loved ones</u> through social media.

_____

_____

_____

2. <u>Many of my friend</u> enjoy surfing during summer.

_____

_____

_____

3. I need <u>new IPhone</u> as my old one is already broken.

_____

_____

_____

4. Nowadays, you <u>can getting</u> free access to the internet almost everywhere.

_____

_____

_____

5. <u>That old house stand-alone</u> looks creepy from the distance.

_____

_____

_____

6. Som <u>always ask</u> people to help her with little things.

_____

_____

_____

7. John likes <u>sometimes</u> climbing the wall.

_____

_____

_____

8. Nowadays, I <u>preferred</u> online shopping to going out to the malls.

_____

_____

_____

9. I <u>would love eating</u> grill meat every other week.

_____

_____

_____

10. They always ask <u>what am I doing</u>.

_____

_____

_____

The End
Thank you ☺

Appendix E:

ICC Values of Timed Argumentative Writing Task

PRETEST

Intraclass Correlation Coefficient

	Intraclass Correlation	95% Confidence Interval		F Test with True Value 0			
		Lower Bound	Upper Bound	Value	df1	df2	Sig
Single Measures	.988	.970	.995	170.776	18	18	.000
Average Measures	.994	.985	.998	170.776	18	18	.000

POSTTEST

Intraclass Correlation Coefficient

	Intraclass Correlation	95% Confidence Interval		F Test with True Value 0			
		Lower Bound	Upper Bound	Value	df1	df2	Sig
Single Measures	.993	.983	.997	278.944	18	18	.000
Average Measures	.997	.991	.999	278.944	18	18	.000

DELAYED-POSTTEST

Intraclass Correlation Coefficient

	Intraclass Correlation	95% Confidence Interval		F Test with True Value 0			
		Lower Bound	Upper Bound	Value	df1	df2	Sig
Single Measures	.977	.858	.993	141.385	18	18	.000
Average Measures	.988	.924	.997	141.385	18	18	.000

ICC Values of Metalinguistic Knowledge Test (MKT)

PRETEST

Intraclass Correlation Coefficient

| | Intraclass Correlation | 95% Confidence Interval | | F Test with True Value 0 | | | |
		Lower Bound	Upper Bound	Value	df1	df2	Sig
Single Measures	.949	.875	.980	37.237	18	18	.000
Average Measures	.974	.933	.990	37.237	18	18	.000

POSTTEST

Intraclass Correlation Coefficient

| | Intraclass Correlation | 95% Confidence Interval | | F Test with True Value 0 | | | |
		Lower Bound	Upper Bound	Value	df1	df2	Sig
Single Measures	.965	.913	.986	55.889	18	18	.000
Average Measures	.982	.954	.993	55.889	18	18	.000

DELAYED-POSTTEST

Intraclass Correlation Coefficient

| | Intraclass Correlation | 95% Confidence Interval | | F Test with True Value 0 | | | |
		Lower Bound	Upper Bound	Value	df1	df2	Sig
Single Measures	.812	.363	.936	14.855	18	18	.000
Average Measures	.896	.532	.967	14.855	18	18	.000

Appendix F:

Normality Test Results

Tests of Normality of the UGJT Scores						
	Kolmogorov-Smirnov			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Pretest	.259	72	.000	.858	72	.000
Posttest	.217	72	.000	.890	72	.000
Delayed	.205	72	.000	.912	72	.000

Tests of Normality of the MKT Scores						
	Kolmogorov-Smirnov			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Pretest	.332	72	.000	.717	72	.000
Posttest	.505	72	.000	.450	72	.000
Delayed	.416	72	.000	.640	72	.000

Tests of Normality of the Combined Explicit Knowledge Test Scores						
	Kolmogorov-Smirnov			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Pretest	.195	72	.000	.926	72	.000
Posttest	.178	72	.000	.936	72	.001
Delayed	.162	72	.000	.957	72	.015

Tests of Normality of the Timed Writing Scores

	Kolmogorov-Smirnov			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Pretest	.077	72	.200	.981	72	.333
Posttest	.044	72	.200	.984	72	.502
Delayed	.084	72	.200	.981	72	.358

Tests of Normality of the OEIT Scores

	Kolmogorov-Smirnov			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Pretest	.353	72	.000	.679	72	.000
Posttest	.290	72	.000	.751	72	.000
Delayed	.278	72	.000	.793	72	.000

Tests of Normality of the Combined Implicit Knowledge Test Scores

	Kolmogorov-Smirnov			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Pretest	.073	72	.200	.973	72	.117
Posttest	.069	72	.200	.987	72	.683
Delayed	.085	72	.200	.962	72	.030

Tests of Normality of the Working Memory Tests

	Kolmogorov-Smirnov			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Backword	.179	67	.000	.868	67	.000
OSpan	.090	67	.200[*]	.975	67	.197
Composite	.074	67	.200[*]	.967	67	.075

Appendix G:

Between-Group Equivalent Test Results

Tests		Sum of Squares	df	Mean Square	F	Sig.
PreW	Between Groups	1415.889	2	707.944	2.258	.112
PreOEIT	Between Groups	2.146	2	1.073	1.492	.232
PreUG	Between Groups	.582	2	.291	.396	.675
PreMKT	Between Groups	.180	2	.090	1.068	.349

Note: Between groups = between the three participating groups: Control, Unfocused and Focused groups.

WM Tests		Sum of Squares	df	Mean Square	F	Sig.
Backward	Between Groups	.348	1	.348	.112	.739
OSpan	Between Groups	1.229	1	1.229	.029	.866

Note: Between groups = between the WCF (i.e., Unfocused and Focused groups) and non-feedback (i.e., Control group) groups.