Explicit versus implicit corrective feedback during videoconferencing: effects on the accuracy and fluency of L2 speech

by

Reza Shirani

M.A, University of Victoria, 2020

B.A, Azad University, 2016

A Thesis Submitted in Partial Fulfillment

of the Requirements for the Degree of

MASTER OF ARTS

in the Department of Linguistics

University of Victoria

Supervisory Committee

Explicit versus Implicit Corrective Feedback During Videoconferencing: Effects on the Accuracy and Fluency of L2 Speech

by

Reza Shirani

MA University of Victoria, 2020

BA Azad University, Isfahan Branch, 2016

Supervisory Committee

Dr. Hossein Nassaji (Department of Linguistics)
Supervisor

Dr. Li-Shih Huang (Department of Linguistics)
Departmental Member

Abstract

A growing body of research has compared the effects of explicit and implicit corrective feedback (CF) on L2 accuracy. However, L2 performance is not limited to accuracy. Fluency is another important aspect of L2 performance, but less is understood about its relationship with CF and CF explicitness/implicitness. This experimental study examined the effects of explicit correction versus implicit recasts on not only the accuracy but also the fluency of L2 speech during videoconferencing. Forty-eight lower-intermediate learners of English as a foreign language (EFL) were assigned to an explicit correction group, an implicit recast group, and a no-feedback group. Each engaged in eight picture description tasks with the researcher and received feedback according to the group they came from. Pre and posttests (immediate and delayed) of accuracy and fluency were conducted using additional picture tasks. Accuracy was measured by calculating the percentage of learners' (a) error-free clauses and (b) error-free T-units. Fluency was measured by calculating the number of (a) syllables per minute and (b) meaningful syllables per minute. Statistical analyses included (a) two-way repeated measures ANOVAs with feedback type as the between-subject factor and time as the within subject factor, (b) Planned comparisons, which treated the two experimental groups as one group and compared their mean with the mean of the control group, (c) Bonferroni post hoc tests, which examined the pairwise differences, and where needed, (d) paired sample *t*-tests, which examined each group's pretest-posttest differences. As for accuracy, planned comparisons showed that videoconferencing CF, irrespective of its explicitness/implicitness, improved accuracy. Further analyses showed that whereas the explicit correction group outperformed the control group on both the immediate and

delayed posttests, the recast group did not. However, the explicit feedback group produced a significantly less fluent speech compared to the recast group and the control group. But this was true on the immediate posttest and not on the delayed posttest. Pretest-posttest comparisons further indicated a negative effect for explicit correction but a positive effect for recasts on L2 fluency. The results suggest that (a) while explicit correction assisted accuracy, it negatively influenced fluency, and (b) while implicit correction seemed to assist fluency, it was not as effective as the effect of explicit correction on L2 accuracy. Further analyses indicated that the explicit correction group exhibited a large amount of monitoring behaviour on the immediate posttest, whereas the other two groups did not. The results are explained using an information-processing perspective of language performance and a knowledge proceduralization model of language development. The theoretical, empirical, and pedagogical implications are also discussed.

Table of Contents

Supervisory Committee ii
Abstract iii
Table of Contents v
List of Tables vii
List of Figures viii
Acknowledgments ix
Dedication x
Chapter 1 1
1.1 Background 1
1.2 Purpose of the Study 3
1.3 Significance of the Study 4
Chapter 2 6
2.1 Introduction 6
2.2 Theoretical Foundations of Corrective Feedback 6
2.2.1 Long's Interaction Hypothesis 7
2.2.2 Swain's Output Hypothesis 9
2.2.3. Schmidt's Noticing Hypothesis 10
2.3 Explicit and Implicit Corrective Feedback 10
2.4 Accuracy and Fluency 13
2.4.1 Levelt's Model of Language Production and Monitoring Theory 15
2.4.2 The Development of Accuracy and Fluency 18
2.5 Explicit and Implicit Corrective Feedback and L2 Accuracy and Fluency 19
2.6 Empirical Background 22
2.6.1 Face-to-Face Studies 23
2.6.2 Corrective Feedback in Online Communication 25
2.6.3 Explaining the Need for the Present Study 27
2.7 Research Questions 30
Chapter 3 32
3.1 Participants 32

3.2 Operationalization of Implicit and Explicit Corrective Feedback ……32
3.3 Design ……..34
3.3.1 Videoconferencing and Audio-Recordings…… .35
3.4 Tasks ……36
3.5 Measurement of Accuracy and Fluency ……..36
3.6 Learners' Monitoring Behavior……40
3.7 Procedures ……...40
3.8 Inter-Rater Reliability……...44
3.9 Statistical Analysis of the Data……45
Chapter 4……46
4.1 Treatment Data……46
4.2 Self-Repair (Monitoring) Behavior ……47
4.3 Pretests…… 47
4.4 Tests of Accuracy ……49
4.4.1 Error-Free Clauses……51
4.4.2 Error-Free T-units……...53
4.5 Tests of Fluency……55
4.5.1 The Number of Syllables per Minute ……58
4.5.2 The Number of Meaningful Syllables per Minute ……60
4.6 Summary of the Results ……..62
Chapter 5……65
5.1 Discussion …….. 65
5.1.1 The First Research Question ……65
5.1.2 The Second Research Question ……67
5.1.3 The Third and Fourth Research Questions……71
5.2 Implications ……74
5.2.1 Theoretical Implications ……...74
5.2.2 Pedagogical Implications ……..74
5.2.3 Implications for Videoconferencing CF ……75
5.3 Conclusions, Limitations, and Directions for Future Research ……76
References ……81

List of Tables

Table 1 Research design ……..…..35
Table 2 Identification of T-units, clauses, errors, and syllables……38
Table 3 Reliability scores ……44
Table 4 Feedback and uptake frequencies…….....46
Table 5 The frequency of learners' self-repair behavior……...47
Table 6 Descriptive and one-way ANOVA results for accuracy pretests…….48
Table 7 Descriptive and one-way ANOVA results for fluency pretests……...48
Table 8 Descriptive statistics for accuracy measures ……...49
Table 9 Tests of normality for error-free clauses and error-free T-units…….……..50
Table 10 Results of Mauchly's test of Sphericity for the measures of accuracy……...50
Table 11 Results of Levene's test of equality of error variances for accuracy……..51
Table 12 Tests of within- and between- subject effects for error-free clauses ……..52
Table 13 Tests of within- and between- subject effects for error-free T-units ……..…54
Table 14 Descriptive statistics for the No. of learners' syllables and meaningful syllables per min…….....56
Table 15 Tests of normality for the measures of fluency……56
Table 16 Results of Mauchly's test of Sphericity for the measures of fluency…….57
Table 17 Results of Levene's test of equality of error variance for fluency measures…..57
Table 18 Tests of within- and between- subject effects on syllables per min……59
Table 19 Tests of within- and between- subject effects on meaningful syllables per min…….……61
Table 20 Summary of the findings…….……...64

List of Figures

Figure 1 Levelt's (1989) model of language production……..…….16
Figure 2 Data collection procedures……..……....43
Figure 3 Group means for tests of error-free clauses……..……..52
Figure 4 Group means for tests of error-free T-units……..……..54
Figure 5 Group means for the number of syllables per min……..……58
Figure 6 Group means for the number of meaningful syllables per min…….…..61

Acknowledgments

I would like to thank my supervisor, Dr. Hossein Nassaji for all he has done for me during my adventure at the University of Victoria. Dr. Nassaji has constantly supported me during my M.A program and has answered all of my questions swiftly and patiently. It has been a great pleasure to be his student. I also appreciate everything Dr. Li-Shih Huang has done for me during my MA program. I am thankful for her useful comments on this thesis and on my other research projects. My sincere thanks also go to Dr. Hua Lin for her constructive comments on my proposal. I would also like to thank Dr. Catherine Caws of the Department of French, University of Victoria, for accepting to serve as the external examiner and for reading and evaluating this thesis. Thanks also go to Jenny Jessa and Maureen Kirby for their administrative support. Finally, I would like to thank my wife, Shadi Rahimipour Anaraki for her constant emotional and motivational support during our stay in Canada.

Dedication

To my family

Chapter 1 – Introduction

1.1 Background

Since Long's (1991, 1996) interaction hypothesis and focus on form proposal, corrective feedback (CF) has become a major topic in second language (L2) research. From Long's perspective, CF is defined as a pedagogical technique used to briefly call language learners' attention to the linguistic problems that arise during meaningful interaction. CF is beneficial in different ways: For example, by integrating form with meaning and helping learners develop the form-meaning mapping needed for L2 learning (Gass, 2003; Long, 1996; Pica, 1994); by helping learners notice the hole in their interlanguage while pushing them to produce modified output (Swain, 1985,1993, 1995); and by providing opportunities for noticing the gap, a process in which learners compare their erroneous utterances with the correct input and notice the differences between their interlanguage and the target language (Doughty, 2001; Schmidt & Frota, 1986). The role of such feedback becomes more highlighted in foreign language (FL) contexts, where exposure to L2 and L2 learning is usually limited to formal classrooms and where teachers' feedback can be a great source of learning.

A variety of feedback types have been identified in different studies (e.g., Lyster & Ranta, 1997; Fu & Nassaji, 2016; Nassaji, 2007; Panova & Lyster, 2002; Shirani, 2019), but generally, CF takes the form of (a) pushing or prompting a learner to self-repair an error (i.e., prompts), (b) reformulating the whole or a part of an erroneous utterance (i.e., recasts), or (c) correcting an error directly through explicit explanation

(Ellis et al., 2006). As such, CF varies in explicitness depending on how it is provided and how directly or indirectly it draws learners' attention to erroneous utterances. For example, recasts are considered relatively implicit as they keep the interlocutors' primary focus on meaning rather than form, and as they rarely interrupt the flow of communication. On the other hand, explicit correction directly points out to learners that their utterance is erroneous. Furthermore, CF can be provided during face-to-face interaction or delivered with the help of technology via computers and smartphones. In either case, empirical evidence suggests that CF, especially explicit CF, improves linguistic accuracy (see Nassaji, 2016a for an updated and chronological review).

However, L2 performance is not limited to accuracy. Fluency is another component of L2 performance that has received increasing attention, especially from a psycholinguistic perspective (e.g., Derwing, 2017; Segalowitz, 2000, 2010). Yet, increasing accuracy and fluency at the same time is not easy. Part of the difficulty, according to an information processing perspective, comes from the selectiveness of our attentional resources and the limitedness of our working memory (Anderson, 2000; Schmidt, 2001; Tomlin & Villa, 1994). Accordingly, attending to one aspect of L2 performance (e.g., form, accuracy) may negatively influence the other performance aspects (e.g., meaning, fluency) (i.e., Skehan, 1996, 1998, 2009 trade-off hypothesis). Moreover, the development of accuracy and fluency may depend on how procedural (i.e., implicit) versus declarative (i.e., explicit) L2 learning or L2 knowledge is. Accuracy, for example, is thought to be associated with learners' ongoing interlanguage knowledge, which may be in part declarative and in part procedural (Housen & Kuiken, 2009; Housen et al., 2012). Fluency, on the other hand, is believed to depend on proceduralized

knowledge to which access is fast and automatic (Anderson, 1993, 1996, 2005, 2007; Anderson & Lebiere, 1998; Anderson et al., 2004; DeKeyser, 2007; Segalowitz, 2010; Segalowitz et al., 1998).

Particularly, Nassaji (2020) has argued that it is important for CF research to also examine the effect of feedback on L2 fluency for two main reasons. First, because CF is accuracy-oriented, it may interrupt the flow of communication and negatively influence fluency during meaning-based interaction. This is important in light of the finding that many teachers are suspicious of error correction because they agree with this claim (Basturkmen, et al., 2004). Second, because our attention is selective and our working memory is limited, by focusing attention to form or accuracy, CF, especially explicit CF, may draw attention away from meaning or fluency. The relationship between CF and L2 fluency is also important regarding the role of error correction in the proceduralization and automatization of L2 knowledge to which fluency is linked (Sato & Lyster, 2012) (see Section 2.5 for a detailed discussion on the relationship between CF, feedback explicitness, and fluency).

1.2 Purpose of the Study

To date, many studies have compared the effectiveness of explicit and implicit CF on L2 learning (e.g., Carroll & Swain, 1993; Ellis, et al., 2006; Erlam & Loewen, 2010; Kim & Mathes, 2001; Leeman, 2003; Lyster, 2004; Monteiro, 2014; Yilmaz, 2012, 2013; Zhao & Ellis, 2020, see Nassaji, 2015, 2016a for a review). However, most of these studies have focused mainly on accuracy and not fluency (Nassaji, 2020). They have also focused on the accuracy of one or two particular target structures, rather than overall

accuracy. This study is the first that examined the possible effects of CF on not only the accuracy but also the fluency of FL learners' speech performance. Furthermore, most of the previous studies have delivered CF during face-to-face interactions, and although some have investigated the effects of computer- and mobile-delivered explicit and implicit CF on L2 writing (e.g., AbuSeileek & Abualsha'r, 2014; Sarré et al., 2019; Sauro, 2009), as for L2 speech, research is scant (cf. Monteiro, 2014). Accordingly, this experimental study was designed to examine (a) whether videoconferencing CF, irrespective of its explicitness/implicitness, had any effects on the overall accuracy and fluency of L2 speech performance, and (b) whether such effects were different for explicit and implicit types of such feedback (i.e., explicit correction and recasts). The study was conducted with participants from an English as an FL context.

1.3 Significance of the Study

The study reported in this thesis is important for empirical, theoretical, and pedagogical reasons. Empirically, the study builds on research into CF and addresses some of the gaps in the previous studies. For example, this study was the first to examine the effects of explicit versus implicit CF on L2 fluency. Also, as for accuracy, in contrast to the previous studies that have focused on one or two particular target structures, this study targeted learners' overall accuracy. I believe overall accuracy helps us obtain a bigger picture of CF effectiveness because it concerns all errors and is "more sensitive to detecting differences between experimental conditions" (Skehan & Foster, 1999, p. 106). Additionally, in contrast to most of the previous studies that have provided CF during face-to-face interactions, this study used WhatsApp videoconferencing to do so. Research

into computer- and mobile-mediated CF is still in its infancy (especially when it comes to L2 speech), and a comparison of the results of this study with those of the face-to-face ones can broaden our understanding of the role of technology in language learning.

Theoretically, this study uses two new theoretical frameworks in the investigation of CF: an information processing perspective of language performance (e.g., Levelt, 1989, 1999; Skehan, 1996, 1998, 2009) and a knowledge automatization model of language development (e.g., Anderson, 1993, 1996, 2005, 2007; Anderson & Lebiere, 1998; Anderson et al., 2004; DeKeyser, 2007), which have seldom been discussed in CF research. I believe these two perspectives provide an interesting discussion of the effects of CF, particularly explicit versus implicit CF, on the accuracy and fluency of L2 speech.

Pedagogically, the study has implications for language teachers who make frequent use of CF in their classes. For example, an important concern that language teachers may have is what types of feedback are most effective for their students. Should they provide their students with more explicit and direct types of CF or should they rely on more implicit and indirect types? This concern becomes even more highlighted when it comes to FL contexts, where most of the learning usually takes place in the classroom and where the teacher is one of the few important sources of learning. Therefore, using the most appropriate pedagogical techniques becomes indispensable in such a context. The results of this study may help teachers understand how differently the explicit and implicit types of CF influence the performance of their students.

Chapter 2 – Literature Review

2.1 Introduction

This chapter reviews the literature related to this thesis. The chapter begins with Section 2.2 where the theoretical underpinnings of CF are discussed regarding the cognitive theories of second language acquisition (SLA). Section 2.3 then sheds further light on the independent variable of the study - feedback explicitness and implicitness. Because explicit correction and recasts are considered the two ends of the feedback explicitness, Section 2.3 is limited to these two feedback types. Section 2.4 presents a discussion of the dependent variables of the study - accuracy and fluency. Section 2.5 then explains the possible relationship between the independent and dependent variables of the study: CF, CF explicitness, and the accuracy and fluency of L2 speech. Section 2.6 is devoted to the empirical background to the study, where the previous face-to-face and computer-mediated studies of explicit versus implicit CF are reviewed. Section 2.6 ends with an explanation of the limitations of the previous research, which provide the rationale for the study reported in this thesis. Finally, Section 2.7. introduces the research questions.

2.2 Theoretical Foundations of Corrective Feedback

The effectiveness of CF has been gauged form different perspectives, among which Longs' interaction hypothesis (1991, 1996), Swain's output hypothesis (1985, 1993, 1995), and Schmidt's noticing hypothesis (1995, 2001), have been widely cited and

discussed. These theories, known as the cognitive theories of (SLA), are discussed in the following sections.

2.2.1 Long's Interaction Hypothesis

The major source of the theoretical support for CF comes from Long's (1991, 1996) interaction hypothesis, whose genesis was Krashen's (1981, 1982, 1985) comprehensible input hypothesis. Krashen claimed that to acquire an additional language, it is only necessary and sufficient to receive input which is understandable, and which contains linguistic structures that are only slightly above the learners' current interlanguage knowledge. Long acknowledged the importance of input, but his argument was different from Krashen's in terms of what enables learners to comprehend input. Krashen claimed that the learner's current L2 knowledge, together with available extralinguistic knowledge, makes input comprehensible. Long, however, argued that what does so is *negotiation*, defined as modifications of interaction made "when learners and their interlocutors anticipate, perceive, or experience difficulties in message comprehensibility" (Pica, 1994, p. 494). The rationale for Long's claims was provided by the findings of his research (Long, 1983) into dyadic interactions of native and non-native speakers. Long found that during such interactions, the interlocutors used particular interactional moves to avoid message incomprehensibility. For example, when communication breakdown was likely, the interlocutors used confirmation, clarification requests, and comprehension checks to negotiate for meaning. Accordingly, Long concluded that interaction facilitates language learning because it provides learners with

comprehensible input; that is, input which has been made comprehensible through negation.

Earlier works of Long mainly concerned negotiation for meaning (e.g., Long, 1983), but he later argued for negation for form (e.g., Long, 1991; 1996; see also Long & Robinson, 1998). Long (1991, 1996) stated that brief attention to linguistic forms, for example by CF, facilitates L2 learning provided that the primary focus of the interaction is kept on meaning. Long's claims provided the basis for *focus on form* (FonF), a term used to refer to pedagogical techniques which aim to attract learners' attention to linguistic forms as engaging in meaning-based tasks (Doughty, 2001; Doughty & Williams, 1998; Ellis, 2001, 2016; Fotos & Nassaji, 2007; Loewen, 2011; Long, 1991, 2000; Long & Robinson, 1998; Nassaji, 2010; Nassaji & Fotos, 2004, 2007, 2011). Long (1991) distinguished FonF from *focus on forms* (FonFs) and *focus on meaning* (FonM) approaches to L2 pedagogy. FonFs refers to traditional approaches where L2 forms are taught explicitly in isolation from meaning. In contrast, FonM concerns meaning-focused approaches such as immersion and content-based programs where implicit learning is emphasized over explicit learning, and where learners are exposed to abundant communicative language with minimal attention to form. FonF, however, strikes a balance between form and meaning and is believed to compensate for the inadequacies of traditional structure-based instruction on the one hand and of the purely meaning-based approaches on the other hand.

A useful technique to implement FonF is CF (e.g., Ellis et al., 2001; Lightbown & Spada, 1990; Loewen, 2012; Lyster & Ranta, 1997; Nassaji, 2016b; Nassaji & Fotos, 2004, 2011; Nassaji & Kartchava, 2017). CF makes it possible to integrate FonFs with

FonM by drawing the learner's attention to linguistic forms within the context of meaning-based interaction. In other words, when receiving CF, learners attend to the grammaticality of certain structures in their output at the same time as they are conveying their message. Consequently, learners are enabled to process both meaning and form and are provided with opportunities to strengthen the form-meaning mapping necessary for language learning (Gass, 2003; Long, 1996; Nassaji, 2016b; Pica, 1994).

2.2.2 Swain's Output Hypothesis

A second theoretical source of support for CF has been provided by Swain's output hypothesis (1985, 1993, 1995), which states that receiving input is necessary but not sufficient and that language learners must generate output as well. According to Swain (1995), producing output is highly effective as it requires learners to "stretch their interlanguage in order to meet communicative goals" (p.127). Swain (1993) puts particular emphasis on the role of pushed and *modified* output in language learning and argues that learners "need to reflect on their output and consider ways of modifying it to enhance comprehensibility, appropriateness, and accuracy" (p.160). According to Swain (1993, 2005), pushed and modified output facilitates L2 learning through *noticing the hole*, a process in which learners figure out that they are not completely able to produce the output they wish to.

CF provides great opportunities for noticing the hole. For example, when learners receive CF on their erroneous utterances, especially in the form of prompts (i.e., output providing feedback), they are pushed to produce target-like output and are allowed to notice the hole in their interlanguage. The output that is modified in response to CF also

contributes to the internalization of L2 knowledge. For example, modified output in response to recasts leads to learning new structures, and modified output in response to prompts (i.e., self-repair) facilitates and automatizes the use of already known structures (McDonough, 2005; Nassaji, 2016b; Panova & Lyster, 2002; Swain, 1995).

2.2.3 Schmidt's Noticing Hypothesis

Another theoretical source of support for CF comes from Schmidt (1995, 2001) who has argued that "people learn about the things that they attend to and do not learn about the things they do not attend to" and that noticing is "the first step in language building" (Schmidt, 2001, pp. 30-31). According to Schmidt (1995, 2001), it is conscious noticing that makes input intake. In other words, the linguistic features of the input are not acquired unless they are noticed (see Schmidt, 2010 for an updated discussion). CF provides great opportunities for learners to notice the linguistic structures of the input they receive and the output they produce. For example, when learners make an error and receive input providing CF (i.e., recasts), their attention is focused on the reformulated structures that they are provided with. When learners receive output providing CF (i.e., prompts), their attention is focused on the self-reformulated structures in their output.

2.3 Explicit and Implicit Corrective Feedback

CF falls on a continuum of explicitness depending on how it is provided and how directly or indirectly it draws the learner's attention to errors. For instance, when feedback contains certain prompts such as rising intonation or added stress, it becomes more noticeable to the learner and is, therefore, more explicit (Nassaji, 2007). Also, when

CF contains a direct provision of the correct form, it overtly and explicitly shifts the learner's attention from meaning to form. In contrast, implicit CF is more meaning-focused and usually does not interrupt the communication flow (Long, 1996, 2007). *Recasts* and *explicit correction* are arguably on the two ends of the feedback explicitness/implicitness continuum.

Recasts are defined as "a reformulation of all or part a leaners' immediately preceding utterance in which one or more nontargetlike (lexical, grammatical, etc.) items are/is replaced by the corresponding target language form(s), and where, throughout the exchange, the focus of the interlocutors is on meaning, but not language as object" (Long, 2007, p. 77). In the interaction below, for example, the native speaker reformulates the learners' erroneous utterance without providing any explicit explanation. Thus, recasts are considered relatively implicit:

> Example 1
>
> Learner: The table camera have three drawers.
>
> Native Speaker: has three drawers. (Nassaji, 2017, p. 357)

Recasts are probably the most controversial type of CF. Some have argued for them (e.g., Ayoun, 2001; Doughty & Varela, 1998; Doughty & William, 1998; Goo, 2012, 2020; Goo & Mackey, 2013; Leeman, 2003; Loewen & Nabei, 2007; McDonough, 2007; McDonough & Mackey, 2006; Nassaji, 2009, 2017) while some have argued against them (e.g., Carroll, 2001; Lyster, 1998a, 1998b; Lyster & Ranta, 1997; Panova & Lyster, 2002). Long (1996, 2007), for example, supported recasts because they promote

implicit learning by indirectly and implicitly connecting form with meaning without breaking the flow of communication. Similarly, Doughty and Williams (1998) argued that recasts are more likely to enhance implicit learning because, unlike explicit feedback that interferes with communication, recasts minimize communication interruptions and maximize focus on meaning (see Goo & Mackey, 2013 for more information on the effectiveness of recasts). In contrast, Lyster (e.g., 1998a, 1998b) has claimed that because recasts are implicit, they are vague. Lyster argued that recasts are not noticed adequately by learners, and therefore, their corrective nature is not salient to learners. Lyster's claim has been based on his uptake studies (e.g., Lyster & Ranta, 1997, Linares & Lyster, 2014; Panova & Lyster, 2002) that have shown that learners are less responsive to recasts compared to more explicit types of CF. Carroll (2001) and Yilmaz (2012, 2013), too, have stated that more explicit types of CF are more effective than recasts because they provide the learner with explicit information about the location of their error and about how the error can be corrected best.

One type of CF that is much more explicit than recasts is explicit correction, defined as "the explicit provision of the correct form. As the teacher provides the correct form, he or she clearly indicates that what the student had said was incorrect" (Lyster & Ranta, 1997, p. 46). In the interaction below, for example, the teacher directly points out to the learner that what he said was wrong. In such situations, the attention of the learner is directly focused on their error:

Example 2

Learner: There is boy.

Native Speaker: *There is boy* is wrong, you should say there is *a* boy.

2.4 Accuracy and Fluency

The dichotomy between accuracy and fluency dates back to the early 1980s, where researchers began to distinguish fluent from accurate language performance as two key constructs of L2 competence (e.g., Brumfit, 1984; Hammerly, 1991). Brumfit (1984), for example, was among the first to differentiate accuracy-focused activities, which aimed to improve the grammaticality of language production, from fluency-focused activities, which aimed to enhance the smoothness and saponaceousness of L2 speech. Later, Skehan (1996, 1998) devised a multicomponent model of language performance, which included accuracy and fluency (and also complexity, which is not within the focus of this thesis) at the same time. Skehan has argued that language performance and language performance development are multidimensional, and that accuracy and fluency are two independent but related aspects of such phenomena.

Accuracy is probably the most traditional dimension of L2 performance, which is defined as the extent to which a learner conforms to the norms and the rule system of the target language (Bui & Skehan, 2018; Foster & Wigglesworth, 2016; Housen & Kuiken, 2009; Housen et al., 2012; Michel, 2017; Pallotti, 2009). There are two types of accuracy: specific and overall. Specific accuracy concerns the correct use of particular linguistic forms such as articles, auxiliaries, verbs, tenses, and so on, while overall accuracy concerns all structures and represents all types of errors. According to Skehan and Foster (1999), overall accuracy provides a more thorough measure for L2 proficiency and thus is "more sensitive to detecting differences between experimental conditions" (p. 106). Overall accuracy has been measured in different ways, but its most common measure is

the percentage of error-free clauses to all clauses (see Ellis, 2005, 2008; Ellis & Barkhuizen, 2005, for an inventory of accuracy measures).

Fluency, on the other hand, is more complicated to define. Traditionally, fluency is viewed as a learner's overall language proficiency and is equated with the smoothness and ease with which one speaks or writes (Freed, 2000; Riggenbach, 2000). Research, however, has suggested that fluency is a multifaceted construct (e.g., Kormos & Dénes, 2004; Lennon, 1990; Towell et al., 1996; Tavakoli & Skehan, 2005). Segalowitz (2010), for example, made a distinction between utterance fluency, perceived fluency, and cognitive fluency (see also Segalowitz, 2000). Utterance fluency is related to the learner's actual performance, which is arguably measurable, and which is further divided into repair fluency (false starts, repetitions, reformulations, and replacements), breakdown fluency (the length and number of pauses), and speed (the number of words or syllables per minute) (Skehan, 2003; Tavakoli & Skehan, 2005). Perceived fluency is associated with how fluent a learner sounds to a listener or rater. Cognitive fluency pertains to the learner's ability to plan their speech and to "the efficiency of the operation of the cognitive mechanisms underlying performance" (Segalowitz, 2000, p. 202). Utterance fluency is probably what most of the language learners try to achieve, but "it is highly dependent on the knowledge and skills of the speaker, which are the bases of cognitive fluency" (de Jong & Perfetti, 2011, p. 534).

Since cognitive psychology began to gain popularity in SLA (e.g., Anderson, 1993, 1996; Levelt, 1989, 1999), accuracy and fluency have been discussed regarding the psycholinguistic processes and mechanisms underlying language production and development. It is believed that accuracy is related to controlled processing, which may

require conscious attention, while fluency is linked with automatic processing, which requires little or no conscious attention (Schmidt, 1992, Segalowitz, 2010). Accuracy is thought to be related to the learners' current interlanguage knowledge, which may be in part explicit declarative and in part implicit procedural, and to which access may be slow and flexible (Ellis, 2008; Housen & Kuiken, 2009). Fluency, in contrast, is thought to depend on proceduralized knowledge and automatized language use, to which access is fast and rigid (de Jong & Perfetti, 2011; Segalowitz, 2010; Towell et al., 1996). Accuracy and fluency have been discussed from different theoretical frameworks (see Mitchel & Myles, 2004 for different theories), among which Levelt's (1989, 1992, 1993, 1999) model of language production and knowledge automatization model of language development (e.g., skill acquisition theory) (Anderson, 1993, 2007; Anderson & Lebiere, 1998; Anderson et al., 2004; DeKeyser, 2007) have been widely cited. These two models are discussed in the following sections.

2.4.1 Levelt's Model of Language Production and Monitoring Theory

Levelt (1989) introduced a model that accounted for the language production of a mature native speaker. This model has been widely applied to L2 as well (e.g., de Bot, 1992; Doughty, 2001; Kormos, 1999, 2011, 2014; Skehan, 2009, 2018; Towell et al., 1996; Towell & Dewaele, 2005). Levelt's model is based on two types of knowledge, namely declarative and procedural. The former concerns explicit, conscious knowledge *about* the world while the latter refers to the implicit, automatic knowledge of *how* to do something. Levelt has argued that the knowledge underlying L1 fluent language production is procedural (due to speed requirements) to which access is fast and

automatic, but different kinds of declarative knowledge can also be accessed in different parts of language production.

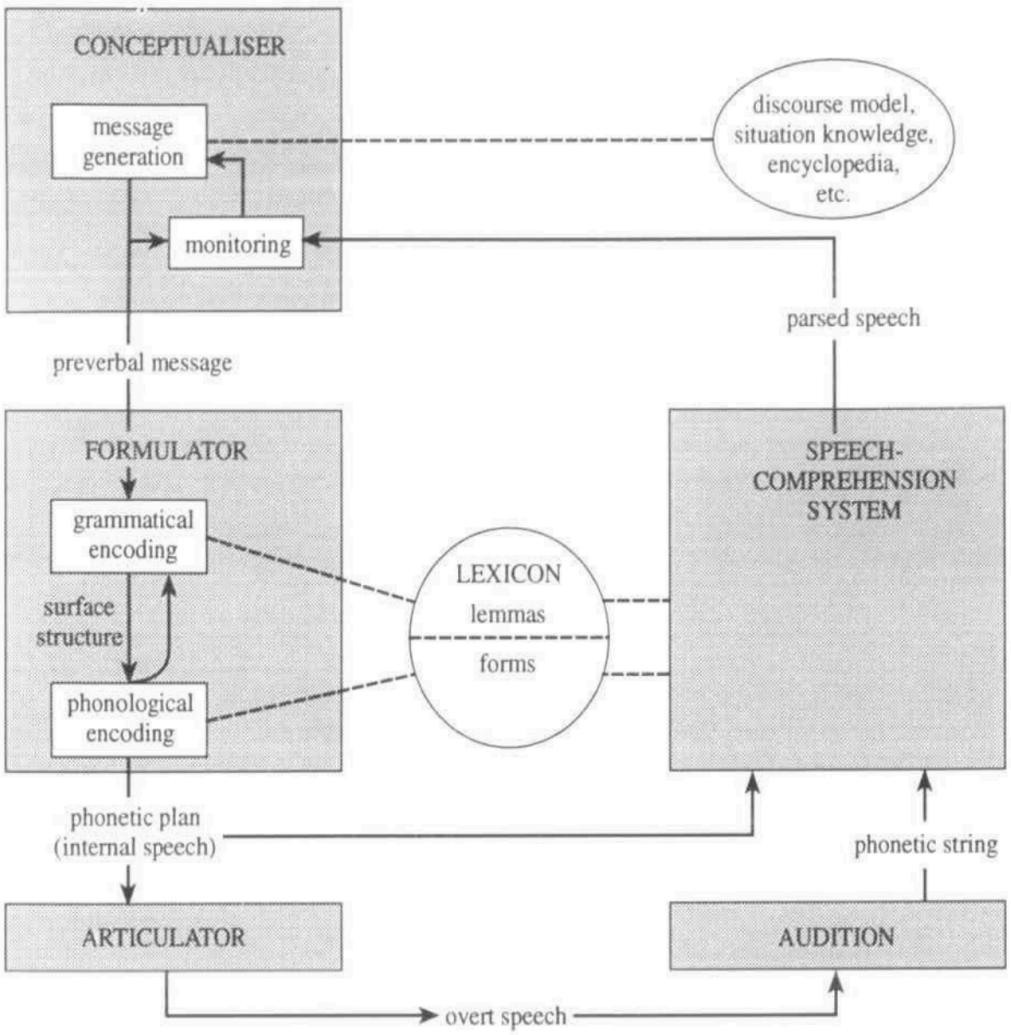

Figure 1. Levelt's (1989, p. 9) model of language production.

As can be seen on the left side of Figure 1, the model contains three main components: the conceptualizer, the formulator, and the articulator. In the conceptualizer, concepts (ideas) are encoded into propositions and the propositional content of a speech

is generated. This is done by accessing declarative knowledge about discourse, situation, and encyclopedia. The propositions then take the form of a preverbal message that acts as the output to the formulator. In the formulator, an acceptable grammatical and phonological form is assigned to the preverbal message in the following way: First, the formulator catches the semantic and pragmatic meanings of the message and then explores the lexicon to find the suitable means of expressing them. The lexicon has two parts. The higher half of it contains form/meaning pairs called lemmas, and the lower half includes morpho-phonological information. The formulator extracts the appropriate lemmas from the higher half of the lexicon and puts them together according to their syntactic and semantic obligations. This results in the creation of surface syntactic structures which are then phonologically encoded based on the morpho-phonological information available in the lower half of the lexicon. The internally formulated utterances are then moved to the articulator via phonetic encoding, and the overt speech is produced.

In the meantime, according to Levelt (1989), speakers attend to their language production and *monitor* and inspect their output. Levelt argued that one can inspect their speech the same way they inspect another persons' speech. Thus, Levelt's comprehension system model (on the right side of Figure 1), which was first used to explain how listeners perceived and inspected a speaker's speech, has also been used to explain *self-monitoring*. Levelt's (1989) model of monitoring has been referred to as the *perceptual loop theory*, which states that there are three loops for monitoring the outcome of the processes involved in language production (see also Kormos, 1999). In the first loop, the preverbal message, which has not been sent to the formulator yet, is checked against the

actual message intent of the speaker. In the second loop, pre-articulatory monitoring occurs in which the message is inspected before production. In the third loop, post-articulatory monitoring takes place in which the formulated utterance is monitored after production.

2.4.2 The Development of Accuracy and Fluency

Levelt's (1989) model explains how language is produced, but it does not account for how language proficiency develops. The development of accuracy and fluency have been largely discussed form a knowledge automatization perspective or what has come to be known as skill acquisition theory (e.g., Anderson, 1993, 1996, 2005, 2007; Anderson & Lebiere, 1998; Anderson et al., 2004; DeKeyser, 2007; Segalowitz, 2010; Segalowitz et al., 1998). According to this position, the knowledge underlying L2 language production in Levelt's model is gained from the frequent use and practice of the language. This perspective holds that fluent L2 speech depends on completely proceduralized knowledge to which access can become fast, rigid, and to some extent automatic (DeKeyser, 2007; Segalowitz, 2010; Towell et al., 1996). According to this perspective, declarative knowledge *about* lexicon and syntax, which is gained from explicit learning, first becomes proceduralized, and then the proceduralized knowledge becomes automatized. DeKeyser (2007) argues that if the relevant declarative knowledge is relevant, available, and accessible during language performance, proceduralization does not take long and "can be complete after just a few trials/instances" (p. 95). What takes longer and occurs more gradually is language automatization, which requires a substantial amount of repeated *practice* (Anderson, 1996, 2005, 2007; Anderson et al.,

2004; DeKeyser, 2007), defined as “specific activities in the second language, engaged in systematically, deliberately, to develop knowledge of and skills in the second language” (DeKeyser, 2007, p. 1). According to DeKeyser (2007), during the proceduralization stage, the learner relies on both declarative and procedural knowledge, but in the automatization stage, knowledge becomes completely procedural, and not only the accuracy but also the fluency of L2 speech develops.

2.5 Explicit and Implicit Corrective Feedback and L2 Accuracy and Fluency

Different instructional and contextual variables may affect the different parts of Levelt’s model of language production, and as a result, lead to changes in the accuracy and fluency of L2 performance. For example, when the context of L2 speech demands high levels of accuracy, learners engage in more monitoring and focus more on the syntactic stage of formulation to avoid errors (Kormos, 1999; Skehan, 2009). Relevant to the monitoring and formulation stages of Levelt’ model is CF. As for the formulation stage, Skehan (2009) argues that the presence of an interlocutor by itself increases learners’ tendency to avoid errors, and accordingly, learners may attend more to the syntactic stage of formulation in such contexts. Attention to accuracy may increase even further when the interlocutor becomes corrective and provides feedback on errors. Furthermore, CF can help learners monitor their language production and make cognitive comparisons of their current interlanguage with the target language (Doughty, 2001; Ellis, 2016; Nassaji, 2016b; Schmidt & Frota, 1986; Swain, 1995, 2005; Swain & Lapkin, 1995). For example, when learners receive recasts or explicit correction, they can compare their erroneous output with the correct input provided by their interlocutors and

learn from the differences between the two. Learners can also compare their erroneous utterance with their subsequent self-corrected output, for example in response to more output provided feedback types. The degree to which CF draws attention to accuracy, however, varies across more explicit and more implicit types of CF. Implicit CF is more meaning-oriented and draws learners' attention indirectly to errors, but explicit CF results in a conscious and explicit shift of attention from meaning to form or accuracy. Based on empirical evidence, many have suggested that CF, especially when provided explicitly, promotes L2 accuracy (e.g., Carroll, 2001; Carroll & Swain, 1993; Ellis et al. 2006; Havranek & Cesnik, 2003; Yilmaz, 2012, 2013). An increase in accuracy, however, could be at the expense of fluency. The reason for this is that due to the limitations of working memory, our attention is selective. Selective attention is defined as "the act of purposively focusing conscious attention on some particular object or goal while ignoring extraneous information that may be present in the situational context" (Ellis, 2016, p. 411). The selectiveness of our attentional resources may make it difficult for learners to focus simultaneously on their accuracy and fluency. This has been explained best by Skehan's (1998, 2009) trade-off hypothesis, which states that "committing attention to one area [of language performance], other things being equal, might cause lower performance in others" (p. 112). Thus, on the one hand, it can be argued that CF, especially explicit CF, negatively influences fluency because it interrupts the communication flow and makes a shift of focus from meaning or fluency to accuracy or form (Nassaji, 2020). Evidence for this argument may come from the language teaching era, influenced by Krashen's (1981, 1982) ideas of Comprehensible Input and Zero Grammar, where learners developed a high level of fluency at the expense of accuracy.

An example is early versions of French immersion programs in Canada where teachers' feedback was focused on content rather than language (Lyster, 2004; Swain & Carroll, 1987; Swain & Lapkin, 1989). Another example is study abroad learners who may not receive much formal instruction or explicit CF during their stay. Some (e.g., Freed, 1995; Freed et al., 2004; Segalowitz & Freed, 2004) have found that immersion and study abroad learners achieve significantly higher levels of fluency than formal classroom learners.

From a knowledge automatization or skill acquisition perspective (Anderson, 1993, 1996, 2005, 2007; Anderson & Lebiere, 1998; Anderson et al., 2004; DeKeyser, 2007; Segalowitz et al., 1998; Segalowitz, 2010), however, it can be argued that CF, particularly explicit CF, not only does not interfere with fluency, but it may also facilitate the proceduralization and automatization to which fluency is linked (e.g., Sato & Lyster, 2012). The reason for this argument is twofold. First, CF, especially when provided explicitly, provides learners with the relevant declarative knowledge required to trigger the process of proceduralization. For example, when learners receive explicit correction, they are provided with explicit knowledge *about* the location of their error and *about* how the error should be corrected. Another example is metalinguistic feedback which delivers declarative knowledge through explicit grammatical comments on learners' errors. The way knowledge is transmitted to learners through such feedback types can be a good example of DeKeyser's (2007) claim that declarative knowledge is mostly "transmitted in verbal form from one who knows to one who does not" (p. 95). Such knowledge can later become proceduralized if it is "drawn on in the execution of the target behavior" (DeKeyser, 2007, p. 95). Second, CF, if given repeatedly, can provide opportunities

for *repeated practice*. Sato and Lyster (2012), for example, have claimed that *uptake*, which is generated in response to CF, is a good example of using the language for communication and of the repeated practice required for automatization. To date, no study has examined whether explicit and implicit types of CF have any effects on the fluency of L2 performance and its development.

2.6 Empirical Background

The effectiveness of implicit versus explicit CF has been investigated by both descriptive (e.g., Nassaji, 2007; Sheen, 2006; Shirani, 2019) and experimental studies (e.g., Carroll, 2001; Carroll & Swain, 1993; Ellis et al. 2006; Goo, 2012; Havranek & Cesnik, 2003; Kim & Mathes, 2001; Loewen & Erlam, 2006; Yilmaz, 2012, 2013; Zhao & Ellis, 2020). Previous descriptive studies have compared implicit with explicit CF on uptake measures (i.e., a learner's immediate response to feedback). In general, it has been shown that implicit feedback might not be as effective as explicit feedback in provoking learner uptake and that explicit feedback may be more successful in drawing learners' attention to the corrective force of feedback. Nassaji (2007), for example, showed that there is a higher possibility of uptake when a feedback move is accompanied by certain signals (e.g., intonational and verbal cues) that make it more salient. Sheen (2006) and Shirani (2019) also found that feedback moves with more explicit characteristics are noticed and uptaken more frequently by learners. However, examining the effectiveness of CF on uptake measures is already so controversial an issue. Some have suggested that there may be no direct relationship between uptake and the acquisition of target forms (e.g., Loewen, 2004; Loewen & Philp, 2006; Nassaji, 2011). Thus, studies with more

controlled experimental designs have been conducted to address the issue. Such studies have either touched on the superiority of explicit feedback over implicit feedback (e.g., Carroll & Swain, 1993; Ellis et al. 2006; Havranek & Cesnik, 2003; Yilmaz, 2012, 2013) or shown no difference between the two (e.g., Kim & Mathes, 2001; Loewen & Erlam, 2006; Loewen & Nabei, 2007; Zhao & Ellis, 2020). Note that these studies have focused on the accuracy of particular forms. No studies have examined the effect of explicit versus implicit CF on L2 Fluency, and therefore, there is no empirical evidence for fluency. Also, note that in some of these studies, CF has been provided during traditional face-to-face interaction (e.g., Carroll, 2001; Carroll & Swain, 1993; Ellis et al. 2006; Havranek & Cesnik, 2003; Loewen & Nabei, 2007; Yilmaz, 2013; Zhao & Ellis, 2020) while some have delivered feedback during online communication (e.g., Bryfonski & Ma, 2020; Loewen & Erlam, 2006; Monteiro, 2014; Saito & Akiyama, 2017; Yilmaz, 2012). Section 2.6.1 concerns the former and Section 2.6.2 concerns the latter.

2.6.1 Face-to-Face Studies

One of the earliest experimental studies by Carroll and Swain (1993) compared the effectiveness of metalinguistic feedback with that of recasts in the acquisition of dative verbs by adult low-intermediate ESL learners. The study revealed that both explicit and implicit CF improved learners' accuracy. However, metalinguistic feedback was shown to have a more significant effect than implicit recasts. Ellis et al. (2006) also compared learners who received metalinguistic feedback with those receiving recasts on the acquisition of English past tense. The study designed an oral imitation test to measure implicit knowledge, and a grammatical judgment test and a metalinguistic knowledge test

to measure explicit knowledge. Finding that explicit feedback resulted in superior performance on both the oral imitation and the grammatical judgment tests, Ellis et al. concluded that explicit feedback benefits not only explicit but also implicit knowledge. More recently, Yilmaz's (2013) study compared three groups of explicit feedback, implicit feedback, and mixed feedback (a combination of explicit and implicit feedback) on oral production tasks. The results revealed that the explicit feedback group and the mixed feedback group outperformed the implicit feedback group, whereas there was no difference between the explicit and mixed feedback groups. Such positive results for explicit feedback have been attributed to the saliency of such feedback. Carroll (2001), for example, concluded that for learners to benefit from CF, they should be able to recognize the corrective nature of feedback so that their focus could be shifted from meaning to form.

In contrast, Kim and Mathes's (2001) comparison of metalinguistic feedback with recasts did not yield any significant differences between the two in facilitating learners' language development, although the participants expressed a preference for explicit feedback. Loewen and Nabei's (2007) experimental study also showed that learners benefited from recasts and metalinguistic feedback to a very similar extent. Goo (2012) also compared the effects of recasts with metalinguistic feedback on the accurate use of English *that*-trace through a written production test and a grammaticality judgment test. The results revealed no significant differences. Besides, the positive effect of recasts in language learning has been confirmed in some of the previous laboratory studies (e.g., Iwashita, 2003; Nassaji, 2009, 2017; Philp, 2003). However, it has been shown that such effects may depend on certain factors such as learner proficiency and recast

characteristics. Nassaji (2009), for example, argued that the efficacy of recasts varies according to how explicitly they are provided. In a more recent classroom study, Zhao and Ellis (2020) operationalized implicit CF as single-move recasts and explicit CF as dual-move (combined with repletion) recasts, which contained intonational stress on the corrected form. The results of an elicited imitation test, designed to measure implicit knowledge, and an untimed grammaticality judgment test, designed to measure explicit knowledge, showed no significant differences between the two types of recasts in the accurate use of regular past tense.

2.6.2 Corrective Feedback in Online Communication

With technology becoming more popular in L2 education, mobile and computer-delivered CF has attracted increasing attention. Some have reported its positive role in language development (e.g., Bower & Kawaguchi, 2011; Loewen & Erlam, 2006; Monteiro, 2014; Rassaei, 2019; Sauro, 2009; Shintani, 2016), and some have argued that it may be even more effective than face-to-face feedback (e.g., Xu et al., 2017; Yilmaz & Yuksel, 2011). Xu et al. (2017), for example, have argued that mobile assisted CF alleviates learners' communication and social anxiety and is, therefore, less threatening than face-to-face feedback (see also High & Caplan, 2009 for a discussion of anxiety in online communication). They also state that due to time constraints, not all learners can receive CF in the classroom context while mobiles and computers have made it easy to provide online feedback on a one-on-one basis. Wang (2006) also posits that the one-on-one nature of online communication provides more opportunities for negotiation of meaning and form. However, technological issues such as background noise, unstable

internet connection, low video and audio quality, lack of technology knowledge, and so on may be problematic during online communication.

One of the early studies of computer-mediated explicit and implicit CF by Loewen and Erlam (2006) provided either recast or metalinguistic feedback to 31 classroom students of English during two communicative tasks and administered a timed and an untimed grammaticality judgment test to measure learning at end of the study. The results showed no differences in the effectiveness of the feedback types. In another study with high-intermediate and advanced learners, Sauro (2009) compared computer-delivered recasts with metalinguistic feedback on the removal of the zero article with abstract uncountable nouns. The results revealed no differences between the two feedback types, but the metalinguistic feedback group showed an immediate superiority over the no-feedback group. Monteiro (2014), too, found no differences between these two feedback types and reported that recasts and metalinguistic feedback both improved low-intermediate learners' explicit and implicit knowledge of English article past tense. More recently, Bryfonski and Ma (2020) compared computer-delivered implicit CF with explicit CF on the acquisition of Mandarin lexical tone by elementary learners. The study found that learners who received recasts improved more significantly in tone production compared to the metalinguistic explanation group. Yilmaz (2012) also compared the posttest performances of learners who received explicit correction with those who received recasts during Turkish language communicative tasks. In contrast to the previous studies, the posttests revealed that learners benefited more from explicit feedback, especially on measures of oral production and comprehension.

One reason why recasts have been as effective as explicit CF in most of the computer-mediated studies may be because computer-mediated recasts can be more salient than face-to-face recasts. Computers and mobiles have made it possible to communicate synchronously not only orally but also in writing (i.e., text chat), and when learners have the opportunity to see, read, and reread CF, they may notice the provided feedback more easily (Sauro, 2009; Yilmaz & Yuksel, 2011). One type of online communication which shares more characteristics with face-to-face communication is *videoconferencing* (Monteiro, 2014; Saito & Akiyama, 2017). For example, such contextual cues as gestures, facial expressions, eye contact, etc., which are common in traditional face-to-face interaction, are also present in videoconferencing. Further, video-based feedback, which is provided orally without the use of text chat, may resemble face-to-face feedback more than text-based CF in terms of saliency.

2.6.3 Explaining the Need for the Present Study

Several gaps in the existing literature have provided the rationale for the present study. First, as can be seen in Sections 2.6.1 and 2.6.2, research into computer-mediated explicit and implicit feedback types comprises a much smaller proportion of literature compared to traditional face-to-face studies, and thus more research into online CF is needed. More specifically, less is understood about the effectiveness of online *videoconferencing* CF as most of the studies of computer-mediated CF have delivered feedback using text chat (e.g., Bryfonski & Ma, 2020; Erlam & Philp, 2006; Sauro, 2009; Yilmaz, 2012) rather than videoconferencing.

Second, in most of both face-to-face (e.g., Carroll & Swain, 1993; Ellis et al., 2006; Kim & Mathes's, 2001; Loewen & Erlam, 2006) and computer-delivered (e.g., Bryfonski & Ma, 2020; Monteiro, 2014; Yilmaz, 2012) studies, the supposedly implicit feedback *recasts* were not provided fully implicitly. Research has shown that recast implicitness varies depending on such factors as feedback length (e.g., Sheen, 2006), context (e.g., Lyster & Mori, 2006), prosodic emphasis (e.g., Nassaji, 2007), number of feedback moves (e.g., Sheen, 2006), the intensity of focus (Erlam & Loewen, 2010), and their frequency of distribution (i.e., the number of feedback moves provided) (e.g., Nassaji, 2017). In Ellis et al. (2006) and Loewen and Erlam (2006), for example, not all of the recast moves were implicit: Both of the studies exemplified recasts as a single-word reformulation of an incorrect utterance, but short recasts are rather explicit (e.g., Sheen, 2006). Similarly, Monteiro (2014) and Yilmaz's (2012) studies of online CF provided recasts through partial reformulations. More importantly, these studies have all provided recasts *intensively* on a particular linguistic form, but intensive recasts are less implicit than extensive ones (Erlam & Loewen, 2010). These shortcomings call for a study that carefully controls the implicitness of recasts.

Third, most of the previous studies (e.g., Carroll & Swain, 1993; Ellis et al., 2006; Kim & Mathes's, 2001; Loewen & Erlam, 2006; Loewen & Nabei 2007; Sauro, 2009) have operationalized explicit feedback as metalinguistic feedback, while very few have used explicit feedback in the form of explicit correction (e.g., Yilmaz, 2012). Thus, more research is needed to compare the effectiveness of explicit correction versus recasts. In addition, the use of explicit correction, instead of metalinguistic feedback, can avoid the possible confusion caused by metalinguistic feedback, because as proposed by Nassaji

(2007), *metalinguistic clues* and *metalinguistic feedback* have always been used interchangeably. However, they are not the same. While the former does not provide the correct form, the latter occurs "in conjunction with explicit correction" (p. 526). This makes the interpretation and comparison of the results of the previous studies using these feedback types (e.g., Carroll & Swain, 1993; Ellis et al., 2006; Loewen & Erlam, 2006) difficult.

Most importantly, previous studies have all focused on the acquisition or accurate use of a particular linguistic form. No studies have examined the effectiveness of explicit versus implicit CF in learners' *overall* accuracy of L2 performance. For example, face-to-face studies of Ellis et al. (2006), Goo (2012), and Zhao and Ellis (2020) focused particularly on learners' accurate use of past tense -*ed*, *that*-trace filter, and 3rd person -*s,* respectively. In computer-delivered studies, too, only specific accuracy has been addressed (e.g., Bryfonski & Ma, 2020; Monteiro, 2014; Yilmaz, 2012). Thus, it is still unclear whether explicit CF and implicit CF have differential effects on learners' overall accuracy. In the meantime, accuracy is only one aspect of language performance. Not many have investigated how feedback explicitness/implicitness may affect learners' fluency, another important component of language performance. A relevant study by Sato and Lyster (2012) compared recasts with prompts on measures of overall accuracy and fluency. The study did not find any differences between the two feedback types and concluded that CF generally helped accuracy but neither helped nor hindered fluency. However, Sato and Lyster did not focus on the effect of feedback explicitness/implicitness on accuracy and fluency. For example, they stated that "all CF happened to be relatively explicit because of its pragmatically unnatural occurrence in the

context of peer interaction," and therefore, their feedback types were distinct "along the dimension of input providing and output providing feedback" (p. 602) rather than the dimension of explicitness. In another relevant study, Saito and Akiyama (2017) reported that EFL learners who received recasts from native speakers in weekly videoconferencing improved significantly in their fluency and lexico-grammar. Saito and Akiyama compared a recast group who participated in task-based interaction with a no-feedback group (rather than an explicit feedback group) who did not participate in any interaction; thus, it is not clear whether gains in fluency were due to the feedback or due to the interaction itself. All in all, no study has compared implicit CF with explicit CF on the overall accuracy and fluency of L2 speech yet.

2.7 Research Questions

To address the issues explained in Section 2.6.3, the study reported in this thesis examined the effects of explicit versus implicit videoconferencing CF on L2 accuracy and fluency. A secondary purpose of the study was to investigate whether videoconferencing CF had any general effects irrespective of its explicit and implicit types. More specifically, the study addressed the following questions:

1. What effects does videoconferencing CF in the form of implicit recasts and explicit correction have on the accuracy of L2 speech?
2. Are the effects on accuracy differential for recasts and explicit correction?
3. What effects does videoconferencing CF in the form of implicit recasts and explicit correction have on the fluency of L2 speech?

4. Are the effects on fluency differential for recasts and explicit correction?

Chapter 3-Methods

3.1 Participants

In total, 78 Iranian learners of English volunteered in response to recruitment notices posted on social networks (e.g., Facebook and Instagram), but only 52 who had similar proficiency levels, were selected. Four participants withdrew by the end of the study, and thus, 48 (29 female and 19 male) completed all of the research events. The participants all reported that they were taking English classes at the time of data collection and that their proficiency level was estimated as low-intermediate (comparable to A2-B1 in Common European Framework of Reference for Languages) by the language institutes they were enrolled in. The average time they reported to have attended a language institute was 26 months, ranging from 11 to 36 months. Except for one female participant who had a 12-day stay in London, none had ever traveled to or stayed in an English-speaking country. All had started learning English after the age of 11. The participants lived in five different provinces in Iran and shared the same L1, Farsi. Their average age was 26.4 (*SD* = 5.53), ranging from 20 to 38. All signed consent forms and finalized their preferred days and hours to meet with me online. During the online meetings, the participants were based in Iran and I in Canada.

3.2 Operationalization of Implicit and Explicit CF

Similar to previous studies (e.g., Adams, Nuevo, Egi, 2011; Bryfonski & Ma, 2020; Ellis et al., 2006; Granena & Yilmaz, 2019; Monteiro, 2014; Nassaji 2017; Sato &

Loewen, 2018; Yilmaz, 2012), the present study operationalized implicit CF as recasts, defined as feedback moves that provide target-like forms for learners through reformulating their erroneous utterances (Goo & Mackey, 2013; Long, 2007; Nassaji, 2017). However, in contrast to such studies, I tried to deliver highly implicit recasts: The recasts (a) did not contain any prosodic emphasis (Nassaji, 2007), (b) were not short (Sheen, 2006), (c) were not provided intensively on only particular errors (Erlam & Loewen, 2010), and (d) were not accompanied with any other feedback moves (Erlam & Loewen, 2010; Sheen, 2006). Recasts can be illustrated by the following examples from the data:

Example 3

Student (S): The boy is laughing to his friend (wrong preposition; word-choice).

Researcher (R): Oh, the boy is laughing at his friend.

Example 4

S: I can see a couple of boys who they are waiting (syntactic error).

R: Okay, so you can see a couple of boys who are waiting.

Explicit CF, on the other hand, was operationalized as explicit correction, defined as explicit provision of the correct form, together with a direct indication that an utterance is erroneous (Nassaji, 2015, 2016b; Yilmaz, 2012). Explicit correction can be illustrated by the following examples from the data:

Example 5

S: He is more happier than others (syntactic error).

R: *More happier* is incorrect. You should say *happier*.

Example 6

S: I think one of them is overcoming the others (word-choice error).

R: *Overcome* is wrong. You should say *outrun.*

3.3 Research Design

The study was an experimental study in which all of the research events occurred during online video-based communication. The study had one independent and two dependent variables. The independent variable was videoconferencing CF with two levels, namely explicit feedback and implicit feedback. The dependent variables were the accuracy and fluency of L2 speech. Pretest-treatment-posttest data collection was used to examine the immediate effects of explicit and implicit CF on learners' accuracy and fluency of L2 speech performance. A delayed posttest was also administered to see if such effects were sustained during the study (see Section 3.8 for data collection procedures). Participants were randomly and equally assigned to two experimental groups that received either implicit or explicit CF, and a control group that did not receive any feedback of any kind. Each group consisted of 16 participants. Table 1 summarizes the design.

Table 1

Research Design

	n	Treatment	Dependent variables
Explicit feedback group	16	Videoconferencing explicit correction	Accuracy and fluency of L2 speech
Implicit feedback group	16	Videoconferencing implicit recasts	Accuracy and fluency of L2 speech
Control group	16	Videoconferencing with no feedback	Accuracy and fluency of L2 speech

3.3.1 Videoconferencing and Audio-recordings

I chose WhatsApp to conduct the study. WhatsApp is a widely used social networking application in Iran, which allows for fast file-transferring and simultaneous audio and video conferencing free of charge. WhatsApp is easy to work with and operates on smartphones with regular internet connections. All of the participants reported that they had access to WhatsApp and knew how to work with it. The majority stated that they used it regularly for their daily communication purposes. Although all of the researcher-participant interactions occurred via videoconferencing, only the audio was recorded because it was sufficient for the study and was less threatening to the participants. Audio was recorded using Voice Memos, an application which uses the built-in microphone in iPhone to record audio.

3.4 Tasks

Eleven picture description tasks were used in this study: one in the pretest, eight in the treatment sessions (four in each session), one in the immediate posttest, and one in the delayed posttest. During such tasks, the participants were given six pictures and were asked to construct and report a meaningful description of the event using their own words. The picture tasks (Appendix A) used in the pretest, immediate posttest, and delayed posttest were Bicycle, Race, and Tiger, respectively, which were originally designed by Heaton (1966). These tasks were chosen because they all consisted of an equal number of frames (i.e., six) and had the same type of sequential structure (i.e., tight) as well as the same number of characters and locations. de Jong and Vercellotti (2016) reported that these three tasks were similar in the complexity of storyline (some intentional reasoning) and derived similar performances on measures of accuracy and fluency. These picture tasks were originally black and white, but de Jong and Tillman (2018) have recently added colour to them to make them clearer to learners. We, accordingly, used de Jong and Tillman's version of the pictures. For the treatment sessions, similar but different sets of pictures were adopted from Heaton (1966) and Mayer (1967) whose pictures have been used by other SLA researchers as well (e.g., Kormos & Trebits, 2012; Tavakoli & Foster, 2008) (see Section 3.7 for procedures).

3.5 Measurement of Accuracy and Fluency

The audio-recordings of the pre and posttest performances were all transcribed using normal arthrography. The transcriptions were checked against the audio files twice to make sure that they were accurate and complete. To measure accuracy, each

transcription was divided into T-units and clauses, and errors were identified. A T-unit is defined as "a main clause plus any other clauses which are dependent upon it" (Foster et al., 2000, p. 360). As for fluency, the syllables produced by the participants were counted.

Note that different researchers have offered different inventories of accuracy and fluency measures (e.g., Ellis, 2005, 2008; Ellis & Barkhuizen, 2005; Iwashita et al., 2008; Michel, 2017; Polio, 2001; Vercellotti, 2017). Ellis (2005, 2008) argues that the existence of a plethora of such measures complicates the comparison of results across the studies, but Ellis also supports the use of more than one measure for each construct to strengthen assessment validity. Particularly in the case of CF, Nassaji (2020) has encouraged to use multiple measures to (a) "capture different levels of acquisition", (b) "cross-validate findings across studies and measures" and (c) consider and compensate for the "variability" in the performance of learners (p. 18). Accordingly, I used two measures for each construct.

In line with the purpose of the study, I used global measures of accuracy rather than specific measures such as correct uses of past-tense, articles, verb-subject agreement, and so on. Overall accuracy was measured by calculating the proportion of error-free clauses to all clauses and error-free T-units to all T-units and then multiplied by 100. T-units, rather than C-units, were used because performances on the pre and posttest tasks were not communicative, and therefore, there were very few incomplete sentences (see Foster et al., 2000 for a discussion of C- and T-units). To identify errors, I followed Foster and Skehan's (1996) definition of error as syntactic, morphological, and word-order deviance from target norms. However, because previous descriptive studies have

shown that teachers provide CF also on learners' phonological errors (e.g., Lyster, 1998b; Lyster & Ranta, 1997; Nassaji, 2010; Sheen, 2006), I counted pronunciation errors as well. Lexical (i.e., word choice) errors were also incorporated as long as they made the message incomprehensible.

As for fluency, speech rate (SR) measures were used, which are thought to reflect learners' underlying cognitive processing and their amount of control over and access to proceduralized and automatized knowledge (Michel, 2017; Segalowitz, 2010; Tavakoli & Skehan, 2005). I used both pruned and unpruned SRs. Unpruned SRs were obtained by the number of syllables in each speech, divided by the number of seconds used to do the task, and then multiplied by 60. As for pruned SRs, the same procedure was followed, but this time, all repeated, reformulated, and replaced syllables, words, and phrases (the preceding ones) were left out. According to Wendel (1997) and Yuan and Ellis (2003), these two measures account for both the amount of speech and the duration of pauses.

Table 2

Identification of T-units, Clauses, Errors, and Syllables

	Identification
T-units	A T-unit is defined as "a main clause plus any other clauses which are dependent upon it" (Foster et al., 2000, p. 360). *Example:* When the boy saw the driver, he started to laugh (composed of two clauses and one T-unit).
Independent clauses	A clause that has a subject and verb and is meaningful on its own. *Example:* he started to laugh.

Dependent clauses	A grammatical structure consisting of a verb which cannot stand on its own and which depends on another clause to be meaningful. *Example:* the first clause in *when he saw the driver, he started to laugh.*
Errors	Syntactic, phonological, morphological, or word-order deviances from target norms, as well as incorrect word choice which obscures the meaning.
Syllables	The number of vowels (e.g., A, E, I, O, U) in each word, but with the followings in mind: a. Count y if it is pronounced as a vowel (e.g., pretty). b. Do not count silent vowels (e.g., e in table). c. Note that diphthongs or triphthongs make only one sound (e.g., flour).

In summary, the following measures were used:

Accuracy:

- percentage of error-free clauses
- percentage of error-free T-units

Fluency:

- number of syllables produced per minute
- number of meaningful syllables produced per minute

Also, the definitions and examples of T-units, clauses, errors, and syllables are presented in Table 2.

3.6 Learners' Monitoring Behaviour

At the beginning of the study, I did not plan to analyze learners' monitoring behaviour. However, at the time of data transcription, I noticed that the explicit correction group exhibited a large amount of error self-repair behaviour, especially on the immediate posttest. Such behaviour was rarely observed in the recast and control groups, and I thought it could be a good indication of the extent to which feedback constituted an accuracy-demanding context and drew learners' attention to form. Consequently, I coded and counted the episodes of error self-repair on the immediate and delayed posttests.

In line with Levelt (1983), error self-repair behaviour was defined as the self-correction of (a) "an erroneously activated word", (b) an "inappropriate syntactic structure", (c) "a wrong morpheme", or (d) or "a wrong phoneme" (Kormos, 1999, p. 317). According to van Hest (1996) and Kormos (1998), a self-repair episode contains a *reparandum* (i.e., an erroneous form) and a *reparatum,* which replaces the reparandum. These definitions were used to code the error self-repair episodes on the posttests. Such episodes can be illustrated by the following examples from the data:

Example 7

P: The boy is rest under, sorry, is resting under the tree,

or

P: He run [pause] ran very fast.

3.7 Procedures

All of the research events occurred online during WhatsApp videoconferencing. At the beginning of each online meeting, the participant and I checked that our internet

connection was good and stable. On the first day of data collection, each participant met individually with me online, and the pretest was administered. I transferred (via WhatsApp) the first picture-description task to each participant and asked them to describe the set of pictures orally in their own words. To capture more authentic and spontaneous speech, the participants were not allowed to take notes and were given not more than one minute to plan their talk. During the pretest, none of the participants received any feedback of any kind. Thus, the descriptions were monologue in nature, and the participants did not have the privilege of instant exposure to linguistic input. Based on a small pilot study with four learners from a similar background, the time limit for task completion was set to 3 minutes, which was the maximum time spent on the task. Each learner's description was audio-recorded and organized according to the group they came from. Finally, the descriptions of each group were first transcribed and then rated on measures of speech accuracy and fluency based on the guidelines in Section 3.5.

Immediately after the pretest, each participant engaged in dyadic videoconferencing interaction with the researcher, where s/he was treated according to the feedback group s/he came from. In the first treatment session, each participant completed four picture-description tasks (in 40-45 min), but unlike the pretest, I communicated with the participant and provided either recast or explicit correction when an error was committed. To make the treatment tasks interactive, I asked multiple questions about the content of the pictures (e.g., "what is the tall boy doing in the first picture?" or "How is the kid feeling in the third picture?"), and provided acknowledgments (e.g., "I think you are right.", "I see.", "I think I agree with you.") or CF in response to the participants. All of the three groups were given the same tasks and

time limits throughout the study. During the treatment tasks, the researcher, who had native-like proficiency, ensured that learners from the implicit CF condition received implicit recasts, those from the explicit CF condition received explicit correction, and the no-feedback participants received feedback only on content and not language. Note that all of the feedback moves for the experimental groups were delivered orally without using the text-chat option. Text chat was avoided because it could provide the advantage of visual feedback saliency. Thus, particularly in the implicit CF group, text chat could decrease the implicitness of recasts. Furthermore, the feedback was provided on a one-on-one basis to isolate the treatment effect. This was done because when CF is provided to a group, there is no guarantee that each person receives as much feedback as others (Nassaji, 2007; Yilmaz, 2012).

Treatment for each participant continued for another session, where each engaged in four more picture-description tasks with the researcher. The second treatment session occurred one day after the first one and followed the same procedures. The reason for keeping the time gap between the two treatment sessions as short as possible was to minimize the effect of any type of learning from outside of the study. Furthermore, different picture tasks were used across the treatment sessions to ensure that the feedback conditions received feedback on a variety of errors as the diversity of the pictures would make the participants use a variety of L2 structures from task to task. Using different picture tasks also avoided the extraneous learning caused by task repetition (see Bygate 2001; Gass & Mackey, 1999). The treatment sessions were audio-recorded, and 15 minutes of each session (i.e., the first five minutes, five minutes from the middle, and the

last five minutes) were analyzed later to make sure that the participants were treated exactly according to the group they came from.

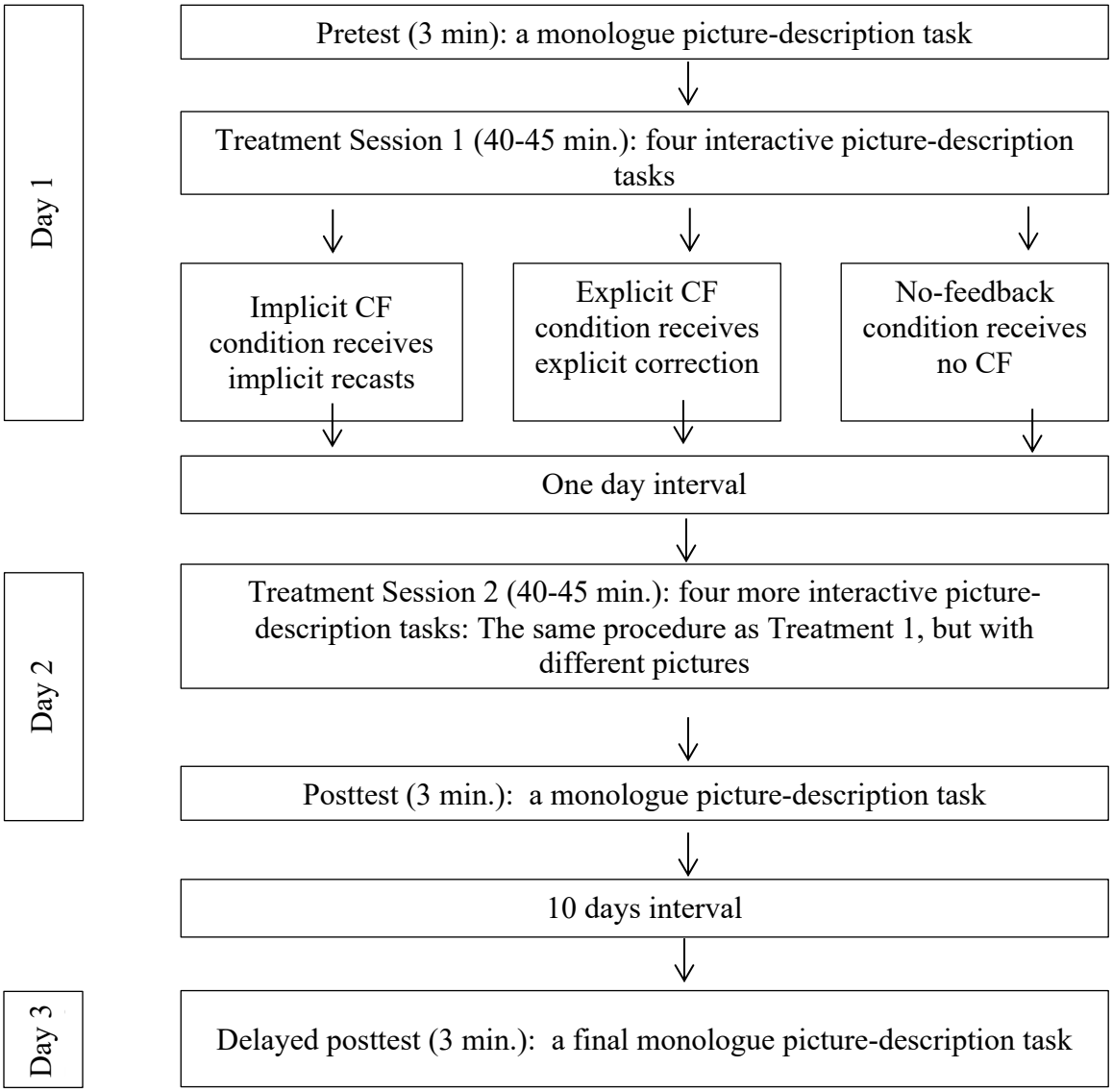

Figure 2. Data collection procedures.

Finally, on the same day of the second treatment session, the immediate posttest, and 10 days later, the delayed posttest was administered. The procedures for the posttests

were the same as the pretest, but again to neutralize the effect of task repetition on the participants' performances, different picture-description tasks were used. Figure 2 summarizes the procedures of data collection.

3.8 Inter-Rater Reliability

A trained assistant (a language teacher with nine years of experience and with a bachelor's degree in applied linguistics) independently rated 31.25% of the data (15 participants, 5 from each group), and the inter-rater reliability for all of the measures of accuracy and fluency was calculated using Cronbach's alpha. Table 3 shows the scores. As can be seen, the reliability scores for the fluency measures are higher than the accuracy measures. It is worth noting that it is quite difficult to obtain high inter-rater reliability on accuracy. For example, it is controversial whether accuracy should be evaluated solely based on the formal rules of a language or also based on the informal but acceptable usages (Ellis, 2008; Housen & Kuiken, 2009; Housen et al., 2012; James, 2013; Pallotti, 2009). Particularly, Cumming et al. (2005) have argued that 76% of inter-rater agreement on accuracy is sufficient.

Table 3

Reliability Scores

Measures	Alpha
Syllables per minute	.94
Meaningful syllables per minute	.91
Error-free clauses	.83
Error-free T-units	.82

3.9 Statistical Analysis of the Data

First, descriptive statistics (means and standard deviations) for the pretest, immediate posttest, and delayed posttest were calculated, followed by a series of one-way ANOVAs to ensure that the three groups were similar on the pretest. To answer the research questions, a series of two-way repeated-measures ANOVAs, with feedback type (explicit versus implicit) as the between-subject factor and time (pretest, immediate posttest, and delayed posttest) as the within-subject factor, were run. Alpha level was set at .05 for all of the analyses. The study used ANOVA, rather than MANOVA, because the measures of accuracy and fluency are considered independent and distinctive (Norris & Ortega, 2009; Skehan, 2009). ANCOVA was not used because the groups were not statistically different on the pretest. To address the effects of videoconferencing CF in general on accuracy and fluency (i.e., the first and third research questions), planned comparisons were performed on the pretest and the immediate and delayed posttests. Planned comparisons make it possible to treat two groups (in this case, recasts and explicit correction) as one single group and compare their weighted mean with the mean of another group (i.e., control) (Pituch et al., 2013). To address the differential effects of recasts and explicit correction on accuracy and fluency (i.e., the second and fourth research questions), a series of Bonferroni pairwise comparisons of the three groups were administered. Where needed, effect size values were also calculated using partial eta-squared (η_p^2). Cohen's (1988) criteria were used to interpret the effect size scores (.01–.05: small effect, .06–.13: moderate effect, and .14 or greater: large effect).

Chapter 4 – Results

4.1 Treatment Data

Before the pretest-posttest analyses, the treatment interactions were analyzed to obtain the distribution of explicit correction and implicit recasts, as well as the frequency of learner repair following each feedback type. As can be seen in Table 4, in total, 480 feedback moves (explicit correction = 229, implicit recasts = 251) were provided to the participants, with explicit correction leading to a significantly higher percentage of learner repair (92.1%) than recasts (21.1%). These results indicate that the two feedback groups received an almost similar amount of CF, but the explicit feedback participants corrected a higher number of their errors after receiving feedback.

Table 4

Feedback and Uptake Frequencies

	Feedback provided			Total learner Repair
	Session 1	Session 2	Total	
EC (*n* = 16)	131	98	229	211 (92.1%)
IR (*n* = 16)	133	118	251	53 (21.1%)
Total	264	216	480	278 (56.7%)

Notes. EC = explicit correction and IR = implicit recasts.

4.2 Self-Repair (Monitoring) Behaviour

As noted earlier, at the outset of the study, there was no plan to analyze learners' self-repair behaviour. However, as we observed many self-repair episodes in the explicit correction group, especially on the immediate posttest, we decided to count and report the number of such episodes as well. Self-repair behaviour could indicate the extent to which the two different feedback types drew the learners' attention to accuracy. As can be seen in Table 5, the participants in the explicit correction group produced more self-repair episodes (79) than the recast group (19) and the control group (21) on the immediate posttest. However, the self-repair behaviour of the explicit correction group decreased on the delayed posttest (29).

Table 5

The Frequency of Learners' Self-Repair Behaviour

	Error self-repair episodes		
	Immediate posttest	Delayed posttest	Total
EC (*n* = 16)	79	29	108
IR (*n* = 16)	19	16	35
NF (*n* = 16)	21	11	32
Total	119	56	175

Notes. EC = explicit correction, IR = implicit recasts, and NF = No-Feedback.

4.3 Pretests

As stated earlier, pretests on accuracy and fluency were given to the three groups to ensure that they were statistically similar at the onset of the study. First a series of one-way ANOVAs were used to compare the students' performance on the pretests. As Table 6 displays, based on the results of one-way ANOVAs, the three groups exhibited similar

performances on the measures of accuracy – that is, the percentage of error-free clauses, $F(2,45) = .774$, $p = .46$, and error-free T-units, $F(2,45) = 2.25$, $p = .11$.

Table 6

Descriptive and One-Way ANOVA Results for Accuracy Pretests

Measure	Groups	*M*	*SD*	*F*	*Sig.*
EFC	IR (n = 16)	22.3	7.5	.774	.467
	EC (n = 16)	20.9	8.7		
	NF (n = 16)	19.1	5.6		
EFT	IR (n = 16)	12.9	4.2	2.25	.116
	EC (n = 16)	11.6	4.8		
	NF(n = 16)	9.8	3.3		

Notes. EFC = error-free clauses, EFT = error-free T-units, IR = implicit recasts, EC = explicit correction, and NF = no feedback.

Table 7

Descriptive and One-Way ANOVA Results for Fluency Pretests

Measure	Groups	*M*	*SD*	*F*	*Sig.*
SPM	IR (n = 16)	45.9	9.5	.468	.629
	EC (n = 16)	43.5	9.6		
	NF (n = 16)	46.3	7.4		
MSPM	IR (n = 16)	41.1	9.2	.961	.390
	EC (n = 16)	38.3	8.8		
	NF (n = 16)	42.5	7.5		

Notes. SPM = syllables per min, MSPM = meaningful syllables per min, IR = implicit recasts, EC = explicit correction, and NF = no feedback.

Similar results were obtained for the fluency measures. The three groups were not statistically different in terms of the number of syllables per minute, $F(2,45) = .468$, $p =$

.62, and the number of meaningful syllables per minute, $F(2,45) = .961$, $p = .39$. Table 7 shows the results.

4.4 Tests of Accuracy

The first two research questions concerned the effects of videoconferencing CF in the form of explicit correction and implicit recasts on the overall accuracy of L2 speech. The descriptive statistics for the two accuracy measures are presented in Table 8.

Table 8

Descriptive Statistics for Accuracy Measures

		Pretest		Immediate posttest		Delayed posttest	
Measures	Groups	*M*	*SD*	*M*	*SD*	*M*	*SD*
EFC	IR (*n* = 16)	22.3	7.5	26.6	7.9	23.7	9.5
	EC (*n* = 16)	20.9	8.7	30.1	8.0	28.5	7.0
	NF (*n* = 16)	19.1	5.6	20.5	9.8	21.4	7.2
EFT	IR (*n* = 16)	12.9	4.2	14.1	3.3	13.0	5.4
	EC (*n* = 16)	11.6	4.8	17.1	5.6	15.6	3.3
	NF (*n* = 16)	9.8	3.3	11.9	5.9	11.7	5.0

Notes. EFC = error-free clauses, EFT = error-free T-units, IR = implicit recasts, EC = explicit correction, and NF = no feedback.

Further statistical analyses revealed that the distribution of data for both correct clauses and correct T-units was not significantly different from normal on any of the tests. Table 9 shows the normality results for the two measures of accuracy.

Table 9

Tests of Normality for Error-Free Clauses and Error-Free T-units.

		Shapiro-Wilk								
		Pretest			Immediate posttest			Delayed posttest		
Measures	Groups	*Stat.*	*df*	*Sig.*	*Stat.*	*df*	*Sig.*	*Stat.*	*df*	*Sig.*
EFC	IR (*n* = 16)	.975	16	.91	.979	16	.95	.961	16	.67
	EC (*n* = 16)	.914	16	.13	.958	16	.61	.947	16	.44
	NF (*n* = 16)	.929	16	.23	.915	16	.13	.979	16	.95
EFT	IR (*n* = 16)	.974	16	.90	.972	16	.87	.955	16	.57
	EC (*n* = 16)	.952	16	.52	.949	16	.48	.979	16	.95
	NF (*n* = 16)	.936	16	.30	.921	16	.17	.942	16	.37

Notes. EFC = error-free clauses, EFT = error-free T-units, IR = implicit recasts, EC = explicit correction, and NF = no feedback.

As can be seen in Tables 10 and 11, the assumptions of Sphericity and homogeneity of variances were not violated, either.

Table 10

Results of Mauchly's Test of Sphericity for the Measures of Accuracy

Measure	Mauchly's W	Approx. Chi-Square	*df*	*Sig.*
EFC	.972	1.24	2	.53
EFT	.993	.31	2	.85

Notes. EFC = error-free clauses, EFT: error-free T-units.

Table 11

Results of Levene's Test of Equality of Error Variances for the Measures of Accuracy

measure		*F*	*df1*	*df2*	*Sig.*
EFC	Pretest	1.04	2	45	.36
	Immediate posttest	.89	2	45	.41
	Delayed posttest	1.37	2	45	.26
EFT	Pretest	.92	2	45	.40
	Immediate posttest	2.60	2	45	.09
	Delayed posttest	1.97	2	45	.15

Notes. EFC = error-free clauses, EFT: error-free T-units.

The two measures of accuracy were independently analyzed using two-way mixed model repeated measures ANOVAs with feedback type (explicit, implicit, and no feedback) as the between-subject factor and time (pretest, immediate posttest, and delayed posttest) as the within-subject factor. Sections 4.4.1 and 4.4.2 report the results.

4.4.1 Error-free Clauses

As for error-free clauses, the results of the two-way mixed model repeated measures ANOVA revealed a main effect for time, $F(2,90) = 10.22$, $p < .001$, with a large effect size ($\eta_p^2 = .18$), indicating that the percentage of participants' error-free clauses increased from the pretest to the posttests. There was also a main interaction effect for time X groups, $F(4,90) = 2.56$, $p < .05$, with a moderate effect size ($\eta_p^2 = .10$), as well as a main effect for group, $F(2,45) = 3.52$, $p < .04$, with a moderate effect size ($\eta_p^2 = .13$). These results suggest that the three groups were significantly different from

each other and that they improved differently over the study. Table 12 displays the mixed model ANOVA results, and Figure 3 illustrates the development of each group over time.

Table 12

Tests of within- and between- Subject Effects for Error-Free Clauses

	Mean square	*F*	*Sig.*	Partial eta squared
Time (Sphericity assumed)	320.8	10.22	.000	.18
Time X groups (Sphericity assumed)	80.4	2.56	.044	.10
Group	463.5	3.52	.038	.13

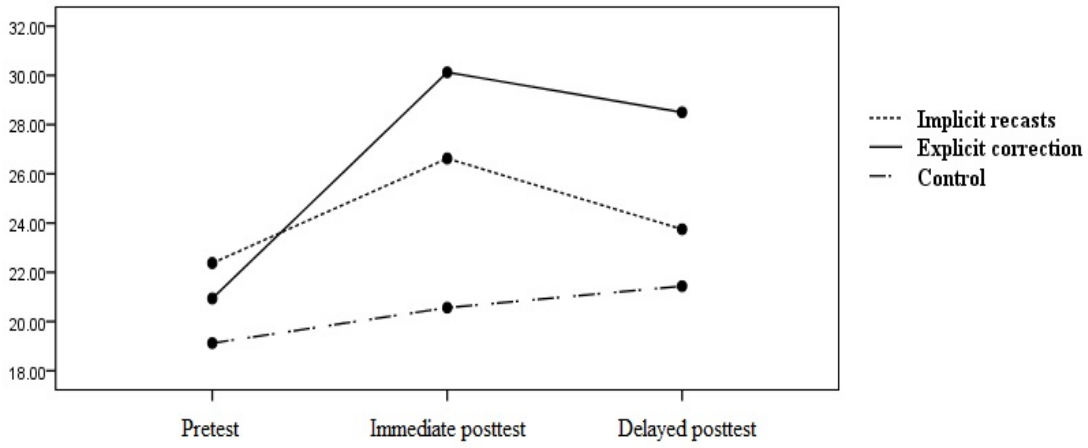

Figure 3. Group means for tests of error-free clauses.

To examine whether CF, irrespective of its explicitness and implicitness, had any effects on the percentage of correct clauses, the two feedback groups were considered as one group and their weighted mean was compared with the mean of the control group. Planned comparisons revealed that although the feedback group and the control group were not statistically different on the pretest, t (45) = 1.11, p = .27, the feedback group

significantly outperformed the control group on the immediate posttest, t (45) = 2.94, p < .006. The feedback group continued to perform better than the control group on the delayed posttest, but the difference was not statistically significant, t (45) = 1.90, p = .063. These findings suggest that CF, regardless of its explicitness/implicitness, had a positive effect on L2 accuracy, but such effect became insignificant over the study.

To examine whether implicit recasts and explicit correction had differential effects on learners' error-free clauses, a series of Bonferroni pairwise comparisons among the experimental groups and the control group were conducted. The analyses indicated the following pairwise differences on the posttests. On the immediate posttest, no significant difference was found between the two excremental groups (p = .77). However, although the explicit correction group significantly outperformed the control group (p < .01), the recast group did not (p = .16). Similar results were obtained on the delayed posttest, where no significant difference was found between the explicit and implicit feedback groups (p = .30) and between the implicit feedback and the control groups (p = .89). However, the explicit feedback group continued to outperform the control group (p < .05). These results suggest that explicit correction had a positive effect on improving the percentage of error-free clauses on both the immediate and delayed posttests.

4.4.2 Error Free T-units

Learners' accuracy in terms of error-free T-units was analyzed following the same statistical procedures as error-free clauses. As can be seen in Table 13, the mixed model ANOVA results revealed a main effect for time, F(2,90) = 8.46, p < .001, and for group, F(2,45) = 3.90, p < .03, showing learners' over-time improvement in accuracy as well as

statistical differences between the groups. However, in contrast to error-free clauses, no interaction effect was found, $F(4,90) = 2.00$, $p = .10$, suggesting that the three groups' production of correct T-units did not improve differently during the study. In addition, the effect size was large for time ($\eta_p^2 = .15$) and group ($\eta_p^2 = .14$) but moderate for time by group ($\eta_p^2 = .08$). Figure 4 shows the development of the three groups on correct T-units measure.

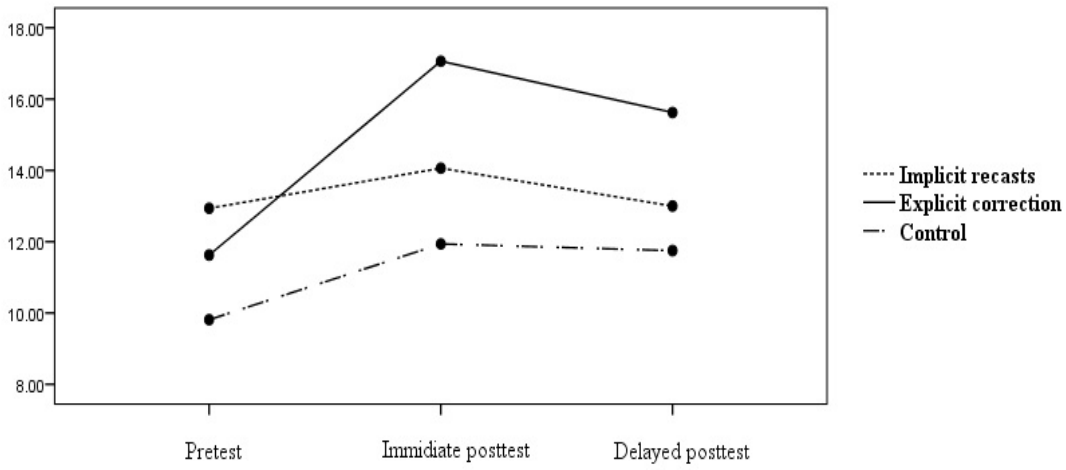

Figure 4. Group means for tests of error-free T-units.

Table 13

Tests of within- and between- Subject Effects for Error-Free T-units

	Mean square	*F*	*Sig.*	Partial eta squared
Time (Sphericity assumed)	105.5	8.46	.000	.15
Time X groups (Sphericity assumed)	25	2.00	.100	.08
Group	158	3.90	.027	.14

Planned comparisons indicated no statistically significant difference between the feedback group (comprised of the explicit correction group and the implicit recast group) and the control group on the pretest. On the immediate posttest, however, the feedback group produced significantly more accurate T-units than the control group, t (45) = 2.31, $p < .03$. On the delayed posttest, the feedback group continued to outperform the no-feedback group, but not statistically significantly, t (45) = 1.78, $p = .08$. The results suggest a positive effect for CF, regardless of its explicitness/implicitness, on learners' production of correct T-units.

Also, Bonferroni pairwise comparisons of the explicit correction group, the implicit recast group, and the control group revealed the following results. On the immediate posttest, no significant difference was found between the two experimental groups ($p = .31$) and between the recast group and the control group ($p = .73$). A significant difference, however, was found between the explicit correction group and the control group ($P < .03$). On the delayed posttest, the explicit correction group performed better than the control group, but, in contrast to the previous measure of accuracy, the difference was not statistically significant ($p = .07$). No differences were found between the recast group and the control group and between the recast group and the explicit correction group ($p = .30$). The results indicate the positive effect of explicit correction on L2 accuracy.

4.5 Tests of Fluency

The third and fourth research questions asked about the immediate and over-time effects of explicit and implicit CF on the overall fluency of L2 speech. Table 14 displays the descriptive statistics for the three groups' scores on the measures of fluency.

Table 14

Descriptive Statistics for the No. of Learners' Syllables and Meaningful Syllables per min

		Pretest		Immediate posttest		Delayed posttest	
Measures	Groups	*M*	*SD*	*M*	*SD*	*M*	*SD*
SPM	IR (*n* = 16)	45.9	9.5	46.3	7.8	46.8	9.3
	EC (*n* = 16)	43.5	9.6	38.8	6.8	41.8	8.0
	NF (*n* = 16)	46.3	7.4	46.8	9.3	48.3	9.0
MSPM	IR (*n* = 16)	41.1	9.2	41.4	9.0	42.0	8.8
	EC (*n* = 16)	38.3	8.8	32.0	8.8	37.0	8.5
	NF (*n* = 16)	42.5	7.5	42.9	8.5	43.4	7.6

Notes. SPM = syllables per min, MSPM = meaningful syllables per min, IR = implicit recasts, EC = explicit correction, and NF = no feedback.

Further statistical analyses confirmed the normality of data distribution for both the number of syllables per minute and the number of meaningful syllables per minute. Table 15 presents the normality results for the two fluency measures.

Table 15

Tests of Normality for the Measures of Fluency

		Shapiro-Wilk								
		Pretest			Immediate posttest			Delayed posttest		
Measures	Groups	*Stat.*	*df*	*Sig.*	*Stat.*	*df*	*Sig.*	*Stat.*	*df*	*Sig.*
SPM	IR (*n* = 16)	.912	16	.12	.977	16	.93	.947	16	.45
	EC (*n* = 16)	.941	16	.36	.960	16	.66	.945	16	.41
	NF (*n* = 16)	.974	16	.90	.938	16	.32	.976	16	.92
MSPM	IR (*n* = 16)	.927	16	.22	.943	16	.38	.958	16	.62
	EC (*n* = 16)	.935	16	.29	.936	16	.30	.946	16	.42
	NF (*n* = 16)	.977	16	.91	.918	16	.15	.910	16	.11

Notes. SPM = syllables per min, MSPM = meaningful syllables per min, IR = implicit recasts, EC = explicit correction, and NF = no feedback.

As Tables 16 and 17 show, analyses also indicated that the assumptions of Sphericity and homogeneity of variances were not violated for any of the fluency measures.

Table 16

Results of Mauchly's Test of Sphericity for the Measures of Fluency

Measure	Mauchly's W	Approx. Chi-Square	*df*	*Sig.*
SPM	.947	2.40	2	.30
MSPM	.994	.31	2	.86

Notes. SPM = syllables per min, MSPM = meaningful syllables per min.

Table 17

Results of Levene's Test of Equality of Error Variances for Fluency Measures

measure		*F*	*df1*	*df2*	*Sig.*
	Pretest	.76	2	45	.47
SPM	Immediate posttest	1.45	2	45	.24
	Delayed posttest	.13	2	45	.87
	Pretest	.37	2	45	.70
MSPM	Immediate posttest	.003	2	45	.99
	Delayed posttest	.086	2	45	.92

Notes. SPM = syllables per min, MSPM = meaningful syllables per min.

The two fluency measures were distinctively analyzed using two-way mixed model repeated measures ANOVAs with feedback type (explicit, implicit, and no

feedback) as the between-subject factor and time (pretest, immediate posttest, and delayed posttest) as the within-subject factor. The next two Sections report the results.

4.5.1 The Number of Syllables per Minute

As can be seen in Table 18, the two-way repeated measures ANOVA results for this measure of fluency revealed no effect for time, $F(2,90) = 1.23$, $p = .295$, and for time by group, $F(4,90) = 1.19$, $p = .31$. No effect for group was found either although the p value was approaching significance, $F(2,45) = 2.88$, $p = .067$. In addition, the effect size values were small for time ($\eta_p^2 = .02$) and for time by group ($\eta_p^2 = .05$), but moderate for group ($\eta_p^2 = .11$). These results indicate no statistically significant over-time effects of CF on fluency measured by the number of learners' syllables per minute. Figure 5 illustrates the performance of each group over time.

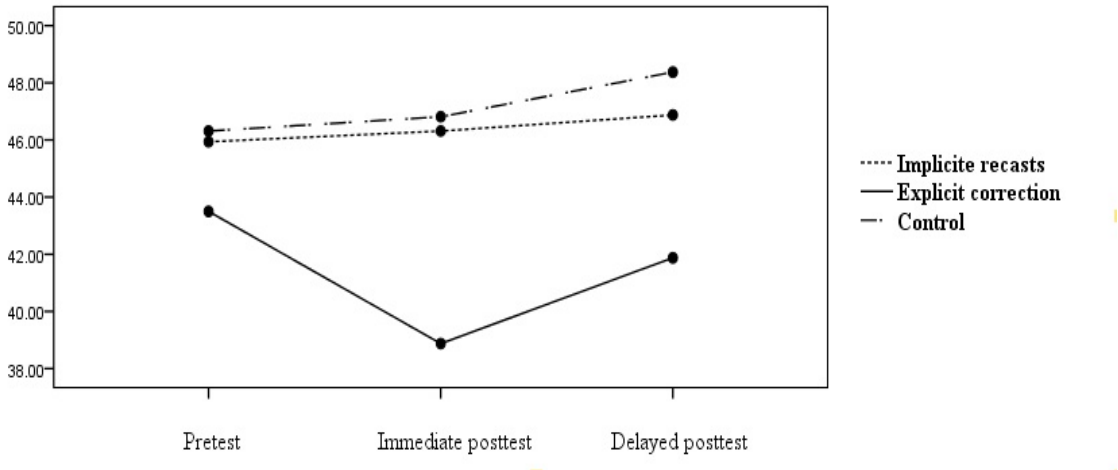

Figure 5. Group means for the number of syllables per min.

Table 18

Tests of within- and between- Subject Effects on Syllables per min.

	Mean square	*F*	*Sig.*	Partial eta squared
Time (Sphericity assumed)	37.5	1.23	.295	.02
Time X groups (Sphericity assumed)	36.3	1.19	.316	.05
Group	466.2	2.88	.067	.11

As for the third research question (i.e., the effects of CF in general on speech fluency), a series of planned comparisons were conducted to compare the mean of the feedback group, comprised of the recast group and the explicit correction group, with the control group. The results revealed no difference on the pretest, t (45) = -.58, p = .56, the immediate posttest, t (45) = -1.70, p = .09, and the delayed posttest, t (45) = -1.48, p = .14. These results suggest that when CF was provided in the form of both recasts and explicit correction, it did not have any negative or positive immediate or delayed effect on L2 fluency.

To answer the fourth research question (i.e., the differential effects of explicit and implicit CF), Bonferroni comparisons of the three groups were conducted. The following results were found. On the immediate posttest, both the recast group (p < .04) and the control group (p < .03) outperformed the explicit correction group, but there was no difference between the recast group and the control group (p = 1). The results indicate that explicit correction had a negative immediate effect on fluency whereas recasts did not. On the delayed posttest, however, no significant pairwise difference was found among the three groups. The results of the delayed posttest suggest that explicit correction did not hinder the development of L2 fluency over the duration of the study.

I also used paired sample *t*-tests to examine within-subject differences between the pretest and the immediate posttest. No significant difference was found between the pretest and the immediate posttest in the recast group, t (30) = -.202, p = .84, and in the control group t (30) = -.218, p = .83. However, the explicit correction group performed significantly better on the pretest compared to the posttest, t (30) = 2.44, p < .03. Overall, these results indicate that on the measure of syllables per minute, explicit correction had a negative immediate effect on fluency whereas recasts did not.

4.5.2 The Number of Meaningful Syllables per Minute

Learners' number of meaningful syllables per minute was analyzed using the same statistical procedures as the previous measure of fluency (i.e., number of syllables per min). The results, however, were different for this measure. This time, the two-way repeated measures ANOVA indicated a main effect for time F(2,90) = 11.19, p < .001, for time by group F(4,90) = 11.76, p < .001, and for group F(2,45) = 3.26, p < 05. Also, the effect size was large for time (η_p^2 = .19), time by group (η_p^2 = .34), but moderate for group (η_p^2 = .12). These results show that as far as the number of *meaningful* syllables was concerned, (a) learners' fluency changed from the pretest to the posttests, (b) the three groups were different form each other, and (c) the groups improved differently over time. Table 19 presents the results, and Figure 6 illustrates the performance of each group over time.

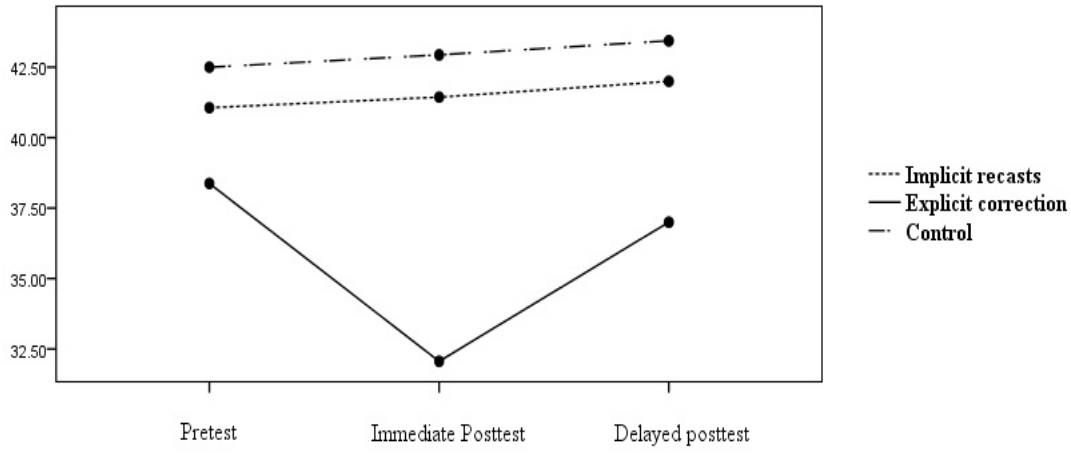

Figure 6. group means for the number of meaningful syllables per min.

Table 19

Tests of within- and between-Subject Effects on Meaningful Syllables per Minute.

	Mean square	*F*	*Sig.*	Partial eta squared
Time (Sphericity assumed)	59.1	11.19	.000	.19
Time X groups (Sphericity assumed)	62.1	11.76	.000	.34
Group	684.2	3.26	.047	.12

Planned comparisons revealed that on this measure of fluency, the control group significantly outperformed the feedback group (including participants from the explicit correction group and the recast group) on the immediate posttest, t (45) = -2.29, $p < .03$, but not on the pretest, t (45) = -1.06, $p = .29$, and the delayed posttest t (45) = -1.54, $p = .13$. These results demonstrate that videoconferencing CF had a negative short-term effect on fluency, even though the effect was not sustained to the end of the study.

Post hoc Bonferroni comparisons examined whether the effects were different for implicit recasts and explicit correction. On the immediate posttest, the recast and control groups were statistically similar ($p = 1$). However, the explicit correction group produced a significantly lower number of meaningful syllables compared to both the recast group ($p < .02$) and the control group ($p < .004$). These results suggest explicit correction hindered fluency, but recasts did not. On the delayed posttest, no pairwise difference was found among the three groups.

Similar to the previous measure of fluency, pretest-posttest differences were also examined using paired *t*-tests. The following results were found. The control group performed similarly on the pretest and the immediate posttest, $t\ (30) = -.953$, $p = .35$. The recast group performed better on the immediate posttest compared to the pretests although not statistically significantly, $t\ (30) = -1.86$, $p = .08$. In contrast, the explicit correction group performed significantly better on the pretest compared to the immediate posttest $t\ (30) = 26.67$, $p < .001$. The results indicate that while recasts assisted fluency, explicit correction had a negative immediate effect on fluency.

4.6 Summary of the Results

The research questions of the study reported in this thesis concerned the immediate and over-time effects of videoconferencing explicit correction and implicit recasts on the accuracy and fluency of L2 speech. In the case of accuracy, when the explicit and implicit feedback groups were considered as one single group (i.e., the feedback group), they produced a significantly higher percentage of correct clauses and correct T-units compared to the control group on the immediate posttest. The

effectiveness of CF, irrespective of its explicitness, however, did not remain statistically significant on the delayed posttest. On the immediate posttest, although the explicit correction group significantly outperformed the control group on both of the accuracy measures, the recast group did not. On the delayed posttest, the explicit correction group continued to perform better than the control group, with the difference being statistically significant on the measure of correct clauses but insignificant on the measure of correct T-units.

As for fluency, different results were obtained. On the measure of syllables per minute, the feedback group and the control group exhibited statistically similar performances on all of the tests. These results suggest that on the measure of syllables per minute, CF does not have any negative effect on fluency. However, on the measure of *meaningful* syllables, despite similarities on the pretest and the delayed posttest, the control group significantly outperformed the feedback group. Given that the analyses revealed some positive effects for recasts but negative effects for explicit correction on both of the measures of fluency, it can be argued that much, if not all, of the negative effect for CF came from explicit correction. Although there was no difference between the recast and control groups, a comparison of the pre and posttest data indicated that recasts assisted fluency while explicit correction had a negative effect on fluency (even though it became insignificant during the study). No effect was found for the control group. Table 20 summarizes the results.

Table 20

Summary of the Finings

Research Questions	Findings
(1) What effects does videoconferencing CF in the form of implicit recasts and explicit correction have on the accuracy of L2 speech?	Videoconferencing CF, regardless of its type had positive effects on L2 accuracy compared to no feedback.
(2) Are the effects on accuracy differential for recasts and explicit correction?	Explicit correction was more effective than implicit recasts.
(3) What effects does videoconferencing CF in the form of implicit recasts and explicit correction have on the fluency of L2 speech	Videoconferencing CF had a negative immediate effect on fluency, but the negative effect came mainly from explicit correction, not recasts.
(4) Are the effects on fluency differential for recasts and explicit correction?	While recasts seemed to assist fluency, explicit correction had a negative effect on L2 fluency. The negative effect of explicit correction was not sustained during the study.

Chapter 5 – Discussion and Conclusion

5.1 Discussion

As stated in Chapter 2, there has been strong theoretical and empirical support for the use of CF, including computer-mediated CF, in L2 learning and teaching. An important question, raised over the past few decades, is whether more explicit or more implicit types of CF are more effective. To address this issue, the study reported in this thesis examined the effects of videoconferencing explicit correction and implicit recasts on two important aspects of L2 proficiency, namely the accuracy and fluency of L2 speech. The results of the study are discussed below regarding each of the research questions.

5.1.1 The First Research Question

The first research question asked whether videoconferencing CF would have positive effects on overall accuracy compared to no feedback. To address this, the study treated the two experimental groups as one group (i.e., feedback group) and then compared their weighted mean with the mean of the control group on the posttests. Based on the results, videoconferencing CF, irrespective of its explicitness/implicitness, had beneficial effects on overall accuracy. The effectiveness of CF was observed on both of the measures of accuracy (i.e., error-free clauses and error-free T-units) and both the immediate and delayed posttests, although the effect was only significant on the immediate test. The findings of this study, together with those of Monteiro (2014) and

Saito and Akiyama (2017), touch on the effectiveness of videoconferencing feedback, suggesting that, similar to traditional face-to-face CF and other types of computer-mediated CF (e.g., text-based), CF, mediated by videoconferencing, helps learners notice their errors and improve their accuracy. Of course, this study did not compare videoconferencing feedback with face-to-face feedback or with other types of computer-mediated feedback. Yet, given that the learners' accuracy improved from the pretest to the posttests, it could be argued that videoconferencing feedback benefitted the learners.

Besides, building on both face-to-face (e.g., Carroll, 2001; Carroll & Swain, 1993; Ellis et al. 2006; Havranek & Cesnik, 2003; Loewen & Nabei, 2007; Yilmaz, 2013; Zhao & Ellis, 2020) and computer-mediated studies (e.g., Bryfonski & Ma, 2020; Loewen & Erlam, 2006; Monteiro, 2014; Sauro, 2009; Yilmaz, 2012) that have pointed to the beneficial effect of CF on *specific* accuracy (i.e., accurate use of particular structures), the present study demonstrates that CF is effective in the development of *overall* accuracy as well. The results are particularly in line with the face-to-face study of Sato and Lyster (2012) in which CF, irrespective of its type, resulted in gains in overall accuracy of EFL learners. It is worth noting that whereas the treatment intervention in Sato and Lyster continued for around 10 weeks (1.5 hr per week), the feedback groups of the current study received treatment during only two sessions of 40-45 minutes, and they still improved in overall accuracy. This observation confirms the effectiveness of CF, even if provided short-term, and adds support to the strong theoretical case that has been made for CF (e.g., Long, 1991, 1996; Schmidt, 2001; Swain, 1993, 1995).

5.1.2 The Second Research Question

The second research question concerned the differential effects of videoconferencing explicit correction and implicit recasts on the overall accuracy of L2 speech. Based on the results of the mixed model repeated measures ANOVAs and post hoc tests, although there was no difference between the explicit correction group and the implicit recast group on any of the measures and tests, only the explicit correction group outperformed the control group on both the immediate and delayed posttests. This finding challenges the assumption that implicit CF is as effective as explicit CF in the development of L2 accuracy while giving further support to the line of research advocating feedback explicitness (e.g., Carroll, 2001; Carroll & Swain, 1993; Ellis et al. 2006; Havranek & Cesnik, 2003; Yilmaz, 2012, 2013). The results are particularly at odds with Saito and Akiyama's (2017) study which confirmed the effectiveness of videoconferencing recasts in accuracy development. One possible explanation for the differences in results could be the larger number of treatment sessions and the higher frequency of recasts in Saito and Akiyama. Saito and Akiyama delivered feedback over 10 sixty-minute sessions while in our study, the treatment for each participant was complete in less than 95 minutes. Consequently, the higher frequency of their recasts might have made recasts more noticeable and salient, and the longer duration of treatment might have provided more uptake and practice opportunities for their participants.

Also, the findings contrast those of the face-to-face studies that have found that recasts significantly improve learners' accuracy (e.g., Goo, 2012; Kim & Mathes, 2001; Loewen & Nabei, 2007; Nassaji, 2017; Zhao & Ellis, 2020). A possible reason for the discrepancy in the findings may be the greater noticeability of the recasts delivered by

these studies. To deliver highly implicit recasts, the present study controlled explicit feedback characteristics such as short length, prosodic emphasis, the intensity of focus, and so on. Such features, however, have seldom been controlled by previous studies. As an example, Nassaji's (2017) study used a mixture of short and long recasts, but research has shown that short recasts are relatively explicit (sheen, 2006). Consequently, his recasts could be more salient to learners than the highly implicit ones in this study. The differences in results may also be attributed to the use of online videoconferencing in the present study. Similar to Monteiro (2014), I noticed several technology-related issues that could prevent the students from noticing the interlocutor's feedback. For example, there were times when the internet connection became unstable and the quality of audio and video declined. There were also times when the participants were distracted by for example their smartphones freezing, running out of battery, or even receiving a call or message from a third party. Such problems could at times make the already-implicit recasts even less noticeable and salient to participants.

There are three possible explanations for the effectiveness of explicit correction over implicit recasts. First, the greater noticeability and salience of explicit correction over recasts could provide more opportunities for learners to conduct cognitive comparisons and notice the gaps between their interlanguage and the target language. As the analyses of the interaction data revealed, explicit correction resulted in a significantly higher percentage of learner repair (92.1%) than did recasts (21.1%). This indicates that the corrective intent of recasts was much less noticeable to learners compared to explicit correction. For example, on many occasions, it was observed that learners perceived recasts as comments on meaning rather than language and responded with

acknowledgments such as “yes, right, exactly,” and so on, rather than correcting their errors. In contrast, explicit correction drew learners’ attention more directly and clearly to their errors, enabling them to compare their erroneous output with the correct input provided while learning from the differences between the two.

A second explanation could be provided from a knowledge automatization or skill acquisition perspective (Anderson, 1993, 1996, 2005, 2007; Anderson & Lebiere, 1998; Anderson et al., 2004; DeKeyser, 2007). It can be argued that explicit correction provided learners with the relevant declarative knowledge required for the processes of knowledge proceduralization and automatization, which account for error-free language production. Recasts, on the other hand, were provided as implicitly as possible, thus transmitting minimum declarative knowledge. During the treatment sessions, explicit correction transferred declarative knowledge verbally to learners by giving them explicit information about the location of their errors and about the best way those errors could be corrected (i.e., X is incorrect; you need to say Y). Of course, because the study was not longitudinal and because leaners did not engage in long-term repeated practice, it is probably not safe to assume that all of the declarative knowledge had become proceduralized and that all of the proceduralized knowledge had become automatized by the time of the posttests. However, it can be argued that at least part of the acquired declarative knowledge might have become proceduralized, as proceduralization by itself, unlike automatization, is not a time-consuming process and can be done after only a few trials and instances (DeKeyser, 2007). In the current study, explicit correction was repeatedly (n = 229) provided by the interlocutor and was repeatedly uptaken (n = 211) by learners over eight speaking tasks, and therefore, learners received arguably adequate

practice to proceduralize part of their declarative knowledge. Consequently, the explicit correction group might have drawn on their partly declarative and partly procedural knowledge to produce more correct clauses and T-units on the posttests.

A third explanation could be provided by the effect of explicit correction on the monitoring and formulation stages of Levelt's (1989) model of language production. According to Skehan (2009), in contexts where accuracy is important, learners pay more attention to the formulation stage of language production. In accuracy-demanding contexts, speakers retrieve more appropriate lemmas, put lemmas together more carefully, and perform more monitoring to avoid errors. An example of such contexts could be the researcher-participant interactions in the explicit correction group of the present study. The presence of a corrective interlocutor, who provided abundant explicit and direct feedback, might have constituted an accuracy-demanding context, which made learners avoid errors by attending more to their accuracy. In the recast group, however, the corrective intent of the interlocutor might have been implicit and vague to learners, and therefore, the context might have demanded a lower level of accuracy. Particularly on the immediate posttest, I noticed a good deal of post-articulatory monitoring among the explicit correction participants. For example, data analyses revealed that, in contrast to the recast and control group, the explicit correction group frequently self-repaired their errors even though they did not receive any feedback on the posttests (see Section 4.2). The explicit correction group also exhibited a large amount of feedback-seeking behaviour, frequently asking about the grammaticality of their utterances (e.g., "was my sentence correct?" or should I say ran or run?") or seeking metalinguistic comments (e.g., "what is the past tense of drive?") on their output. Self-repair and feedback-seeking

behaviours were rarely observed in the recast group and the control group, and they could be a good indication of the explicit correction group's tendency to avoid errors. It is worth noting that although the immediate posttest was monologue and feedback-free, it was conducted only minutes after the treatment, and as a result, the accuracy-demanding context that was formed in the treatment sessions might have continued to exist at the time of the immediate posttest. However, on the delayed posttest, which occurred ten days after the treatment, the monitoring behaviour of the explicit correction group declined. This can further explain why the superiority of the explicit feedback group over the control group was more significant on the immediate posttest ($p < .01$, on correct clauses; $p < .03$, on correct T-units) than was on the delayed posttest ($p < .05$, on correct clauses; $p = .07$, on correct T-units).

5.1.3 The Third and Fourth Research Questions

To avoid being repetitive, the third and fourth research questions are discussed together. The questions concerned the effects of videoconferencing CF on L2 fluency, and whether such effects were different for the explicit and implicit types of feedback. The results revealed that explicit correction and recasts as one group had a negative effect on L2 fluency only on the immediate posttest and only on the measure of meaningful syllables. However, given that the recast group and the control group outperformed the explicit correction group on the immediate posttest, and given that recasts were found to have positive effects on fluency (although insignificantly), it can be argued that the negative immediate effect for videoconferencing CF came mainly from explicit correction. The results of the two-way mixed model repeated measures ANOVAs,

together with post hoc tests, showed that explicit correction negatively influenced L2 fluency on the immediate posttest, but not on the delayed posttest.

The immediate negative impact of explicit correction on fluency can be explained by Skehan's trade-off hypothesis (1998, 2009), which states that paying more attention to one area of language performance (e.g., form or accuracy) may negatively influence the other (e.g., meaning or fluency). The reason for this argument is that performance in accuracy and fluency involves attention and working memory, but our attentional resources are selective and our working memory capacities are limited. Consequently, it is fairly difficult to perform excellently in both accuracy and fluency as learners need to prioritize their selective attention. In the explicit correction group, as stated earlier, the presence of an interlocutor, who directly and explicitly pointed to learners' errors, made learners perform more monitoring and attend more to their accuracy (see Section 5.1.2). Committing more attention to accuracy, however, might have been at the expense of fluency, thus resulting in lower fluency scores on the immediate posttest. As attention to accuracy decreased by the time of the delayed posttest (as indicated by a decrease in self-repair and feedback-seeking behaviour), the fluency scores increased in the explicit correction group. On the other hand, implicit recasts were more meaning-focused and constituted a less accuracy-demanding context compared to explicit correction. The feedback-free interactions in the control group were even more meaning-oriented and, as such, learners in the recast group and the control group had probably greater attentional resources available to process the forms fluently. Consequently, they produced a more fluent speech than did the explicit correction group.

Another explanation for these findings can be provided from a knowledge automatization perspective, which states that fluency depends significantly on procedural and automatized knowledge (e.g., Anderson, 1993, 1996, 2005, 2007; Anderson & Lebiere, 1998; Anderson et al., 2004; DeKeyser, 2007; Segalowitz et al., 1998; Segalowitz, 2010). In this study, explicit correction developed learners' declarative knowledge to which access was slow and flexible. The study did not provide longitudinal practice, and by the immediate posttest, not all of the learners' declarative knowledge had become proceduralized and automatized, and therefore, learners had little, if any, automatic and fast access to the acquired structures. However, by the time of the delayed posttest, the explicit correction group had probably proceduralized a larger amount of their declarative knowledge and thus performed as fluently as the other groups. Of course, this argument needs to be considered with caution as no further practice was provided to learners after the immediate posttest, and arguably without any practice, declarative knowledge may remain declarative (Anderson, 1993, 1996, 2005, 2007; DeKeyser, 2007). However, note that at the time of the study, all of the learners were enrolled in language schools, where they had the opportunity to use and practice the forms that they had acquired from the present study. Consequently, it can be argued that the explicit correction participants continued to proceduralize their declarative knowledge between the immediate posttest and the delayed posttest.

5.2 Implications

5.2.1 Theoretical Implications

Previous studies of CF (except Sato & Lyster, 2012) have mainly drawn on Long's (1985, 1991, 1996) interaction hypothesis, Schmidt's (1995, 2001) noticing hypothesis, and Swain's (1993, 1995) output hypothesis to explain the effectiveness of CF. While these theories constitute a legitimate and powerful framework, I would argue that CF can be viewed from other perspectives as well. The present study is one of the first and few to explain the effectiveness of CF from an information-processing perspective of language performance (e.g., Levelt, 1998, 1999; Skehan, 1998, 2009) and a knowledge automatization model of language development (e.g., Anderson, 1993; 1996, 2005, 2007; Anderson & Lebiere, 1998; Anderson et al., 2004; DeKeyser, 2007; Segalowitz et al., 1998; Segalowitz, 2010).

5.2.2 Pedagogical Implications

The study has particular implications for videoconferencing language classes, as all of the research events were conducted using WhatsApp videoconferencing. However, given that videoconferencing interaction shares many characteristics with face-to-face interaction (Monteiro, 2014; Saito & Akiyama, 2017), the study can have implications for face-to-face pedagogy as well.

One of the concerns of language teachers – whether teaching online or face-to-face – is what type of feedback is more effective for their students. This concern becomes more highlighted in FL contexts, where, compared to second language contexts, there are fewer opportunities for interaction and feedback, and where the classroom and the

teacher are one of the few sources of learning. The present study compared the two ends of the feedback explicitness continuum in an FL context and found that explicit correction resulted in a higher level of accuracy than implicit recasts in such a context. However, another concern that many teachers may have is whether more explicit and direct types of CF impede the development of L2 fluency. Many teachers, for example, are suspicious of error correction because they believe it interferes with the flow of communication (Basturkmen et al., 2004). While this study revealed the negative immediate effects of explicit correction on L2 fluency, it found that in the long term, explicit correction did not hinder fluency. However, another finding of the study was that recasts had some positive effects on L2 fluency. Thus, overall, the study encourages the use of both explicit and implicit feedback to enhance not only accuracy but also fluency. Based on the results, the study suggests language teachers use a combination of explicit and implicit feedback rather than relying on only one type.

5.2.3 Implications for videoconferencing CF

The study has also particular implications for online feedback. The findings of the study are significant as technology is gathering momentum in L2 education and as distant/online learning is becoming more popular worldwide. The study suggests that similar to traditional face-to-face and similar to other types of computer-mediated CF (e.g., text-based), videoconferencing feedback is helpful. However, the study faced several issues related to videoconferencing. The study recommends that teachers be cautious about and consider the following challenges when delivering videoconferencing feedback. First, there were times when the internet connection was not stable, and as a

result, the quality of audio and video suffered. This could result in unpleasant background noises, audio-video mismatches, audio latency, and interruptions in audio and video. Other issues such as smartphone freezing, running out of battery, and receiving calls and messages from other parties could cut off the conversation as well. Such technological problems could prevent the interlocutor from detecting the learner's error and the learner from understanding the interlocutor's feedback. Thus, teachers should be aware of these issues and make sure their students have access to high-speed internet and those technological problems are controlled. Another issue with videoconferencing feedback was the existence of many distractors in learners' environment. For example, on multiple occasions, it was observed that learners were distracted by someone entering their room, someone turning up the TV, someone calling out their names, and so on. The main reason for these distractors was that most of the participants were at their houses at the time of the data collection. Teachers could prevent these issues by preparing a less distracting place for their learners.

5.3 Conclusions, limitations, and Directions for Future Research

The purpose of the study reported in this thesis was to examine the effects of videoconferencing CF on the overall accuracy and fluency of L2 speech, and whether the effects were differential for the explicit and implicit types of such feedback. The present study, together with Monteiro (2014) and Saito and Akiyama (2017), indicated that videoconferencing CF, irrespective of its type, was effective in the development of accuracy. However, the effect was much more significant for explicit correction than implicit recasts. Consequently, consistent with many studies (e.g., Carroll, 2001; Carroll

& Swain, 1993; Ellis et al., 2006; Muranoi, 2000; Nagata, 1993; Rosa & Leow, 2004; Yilmaz, 2012, 2013), the present study suggests that the effectiveness of CF in accuracy development may depend on how directly and explicitly it draws the learner's attention to form. The study further showed that while implicit correction helped fluency, explicit correction negatively influenced fluency because it drew learners' selective attention to accuracy. However, the effect was short-term and became insignificant during the study as learners went through language proceduralization. Thus, similar to Sato and Lyster (2012), I argue that as the declarative knowledge obtained from feedback is proceduralized, CF, even in its explicit and direct form, is unlikely to obstruct the development of fluency. However, in contrast to Sato and Lyster, I conclude that due to the limitedness of working memory capacity and the selectiveness of attentional resources, CF, especially explicit CF, may have a negative impact on fluency *immediately* after the interaction.

The study contributes to SLA research in several ways. First, it is the first to examine the effects of explicit and implicit CF on L2 fluency. This is important in terms of the argument that language proficiency is not restricted to accuracy only and that achieving high levels of fluency is also favourable. The study, therefore, opens a new line of research into feedback explicitness: its relationship with speech fluency. The study is also novel in that it uses global measures of accuracy, in contrast to many studies that have examined CF on specific measures (i.e., one or two particular structures). Given that the study was not longitudinal and that the results were still significant, I argue that explicit CF is effective even if it is provided over a short period. The study also contributes to research into computer-mediated CF, as it is one of the few to deliver CF

via videoconferencing. Video-based online language classes are becoming more and more popular worldwide, but insufficient research has been conducted in this area. For example, while many have investigated written forms (e.g., using text chat) of online feedback, only a few have examined the effectiveness of videoconferencing feedback. The present study suggests that CF mediated by videoconferencing is effective and encourages further research in this area.

The results of the study, however, should be interpreted with caution. First, the finding that explicit correction did not impede the development of L2 fluency during the study was attributed to the proceduralization of the declarative knowledge gained from feedback. It was speculated that proceduralization occurred as a result of the practice that was provided by both the study and the participants' use of the language outside the study. However, the study did not empirically measure proceduralization. Future research could use tests of declarative and procedural knowledge (e.g., timed and untimed grammaticality judgment tests, Ellis, et al., 2006) to examine whether and how much of the declarative knowledge obtained from explicit feedback is proceduralized.

Another limitation of the study is that it was not longitudinal. There were only two treatment sessions, and the intervals between the research events were fairly short. Although the results were significant, a more longitudinal study could have helped obtain a bigger picture of how explicit CF and implicit CF differed. Nassaji (2017), for example, has suggested that the interval between the immediate and delayed posttests be much longer than two weeks to "show whether learners can incorporate the forms into their long-term interlanguage system" and "whether learners are able to transfer their

knowledge to new contexts" (p. 365). Consequently, the present study could be replicated using a more longitudinal design.

Besides, this study compared the effects of two different types of CF, explicit correction and implicit recasts, on accuracy and fluency. Future research can investigate these effects for the explicit and implicit subtypes of the same feedback. For example, it is worthwhile to examine whether more explicit and more implicit recasts (e.g., short versus long recasts) or more implicit and more explicit prompts (e.g., elicitations versus clarification requests) have differential effects on accuracy and fluency. Also, given that the study delivered CF using videoconferencing, future studies could examine whether similar results would be obtained for face-to-face or other types of computer-mediated feedback. For example, the present study did not use such options of WhatsApp as text chat and voice mail. An interesting question for future research then is whether such options can influence the effectiveness of explicit and implicit feedback in accuracy and fluency development.

Another open topic for future research is the role of individual differences in the effectiveness of explicit and implicit feedback. As noted earlier, one of the variables that may affect the effectiveness of CF is working memory capacity (Goo, 2012; Sagarra, 2007; Yilmaz, 2013). Yilmaz (2013), for example, found a relationship between the effectiveness of explicit correction in accuracy development and learners' working memory capacity. Consequently, future research could examine whether the variable of working memory capacity interacts with the effect of explicit feedback and implicit feedback on L2 fluency.

Finally, the study did not examine the effectiveness of feedback regarding each type of error. Consequently, future research could examine how the type of error interacts with the effectiveness of explicit and implicit CF. More specifically, future research could investigate (a) what types of errors are committed more or less than other types, and (b) for what types of errors explicit versus implicit CF are more or less effective.

References

AbuSeileek, A., & Abualsha'r, A. (2014). Using peer computer-mediated corrective feedback to support EFL learners' writing. *Language Learning & Technology*, *18*(1), 76-95.

Adams, R., Nuevo, A. M., & Egi, T. (2011). Explicit and implicit feedback, modified output, and SLA: Does explicit and implicit feedback promote learning and learner–learner interactions?. *The Modern Language Journal*, *95*(1), 42–63.

Anderson, J. R. (1993). *Rules of the mind*. Hillside, NJ: Erlbaum.

Anderson, J. R. (1996). *The architecture of cognition*. Cambridge, MA: Harvard University Press.

Anderson, J.R. (2000). *Learning and memory: An integrated approach*. New York: Wiley.

Anderson, J. R. (2005). Human symbol manipulation within an integrated cognitive architecture. *Cognitive Science*, *29*(3), 313-341.

Anderson, J. R. (2007). *How can the human mind occur in the physical universe?* New York, NY: Oxford University Press.

Anderson, J. R., Bothell, D., Byrne, M. D., Douglass, S., Lebiere, C., & Qin, Y. (2004). An integrated theory of the mind. *Psychological Review*, *111*(4), 1036–1060.

Anderson, J. R., & Lebiere, C. (1998). *The atomic components of thought*. Mahwah, NJ: Lawrence Erlbaum.

Ayoun, D. (2001). The role of negative and positive feedback in the second language acquisition of the passé composé and imparfait. *The Modern Language*

Journal, *85*(2), 226–243.

Basturkmen, H., Loewen, S., & Ellis, R. (2004). Teachers' stated beliefs about incidental focus on form and their classroom practices. *Applied Linguistics*, *25*(2), 243–272.

Bot, K. D. (1992). Applied linguistics. *Applied Linguistics*, *13*(1), 1–24.

Bower, J., & Kawaguchi, S. (2011). Negotiation of meaning and corrective feedback in Japanese/English eTandem. *Language Learning & Technology*, *15*(1), 41-71.

Brumfit, C. (1984). *Communicative methodology in language teaching: The roles of fluency and accuracy*. Cambridge: Cambridge University Press.

Bryfonski, L., & Ma, X. (2020). Effects of implicit versus explicit corrective feedback on mandarin tone acquisition in a SCMC learning environment. *Studies in Second Language Acquisition*, *42*(1), 61–88.

Bui, G., & Skehan, P. (2018). Complexity, fluency and accuracy. In: J. Liontas (Ed.), *TESOL encyclopedia of English language teaching*. Hoboken, NJ: Wiley-Blackwell.

Bygate, M. (2001). Effects of Task repetition on the structure and control of oral language. In M. Bygate, P. Skehan & M. Swain (Eds), *Researching pedagogic tasks: second language learning, teaching and testing*, pp. 23-48. Harlow: Longman.

Carroll, S. E. (2001). *Input and evidence. The raw material of second language acquisition.* Amsterdam: J. Benjamins.

Carroll, S., & Swain, M. (1993). Explicit and implicit negative feedback: An empirical study of the learning of linguistic generalizations. *Studies in Second Language Acquisition*, *15*(3), 357–386.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum.

Cumming, A., Kantor, R., Baba, K., Erdosy, U., Eouanzoui, K., & James, M. (2005). Differences in written discourse in independent and integrated prototype tasks for next generation TOEFL. *Assessing Writing*, *10*(1), 5–43.

De Jong, N., & Perfetti, C. A. (2011). Fluency training in the ESL classroom: An experimental study of fluency development and proceduralization. *Language Learning*, *61*(2), 533-568.

De Jong, N., & Tillman, P. (2018). Grammatical structures and oral fluency in immediate task repetition. In M. Bygate (Ed.), *Learning language through task repetition* (pp. 43–70). Amsterdam: Benjamins.

De Jong, N., & Vercellotti, M. L. (2016). Similar prompts may not be similar in the performance they elicit: Examining fluency, complexity, accuracy, and lexis in narratives from five picture prompts. *Language Teaching Research*, *20*(3), 387–404.

DeKeyser, R. (2007). Skill acquisition theory. In B. VanPatten & J. Williams, *Theories in second language acquisition* (pp. 97–113). Mahwah, NJ: Erlbaum.

Derwing, T. M. (2017). L2 Fluency Development. In S. Loewen., & M. Sato (Eds.), *The Routledge handbook of instructed second language acquisition* (pp. 246–259). London: Taylor & Francis.

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

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

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.

Ellis, R. (2001). Introduction: Investigating form-focused instruction. *Language Learning*, *51*(1), 1-46.

Ellis, R. (2005). Planning and task-based performance: Theory and research. In R. Ellis (Ed), *Planning and task performance in second language* (pp. 3–34). Amsterdam: John Benjamins.

Ellis, R. (2008). *The study of second language acquisition*. Oxford: Oxford University Press.

Ellis, R. (2016). Focus on form: A critical review. *Language Teaching Research*, *20*(3), 405–428.

Ellis, R. & Barkhuizen, G. (2005). *Analyzing learner language*. Oxford: Oxford University Press.

Ellis, R., Basturkmen, H., & Loewen, S. (2001). Learner uptake in communicative ESL lessons. *Language Learning*, *51*(2), 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*(2), 339–368.

Erlam, R., & Loewen, S. (2010). Implicit and explicit recasts in L2 oral French

interaction. *Canadian Modern Language Review*, *66*(6), 877–905.

Foster, P., & Skehan, P. (1996). The influence of planning and task type on second language performance. *Studies in Second Language Acquisition*, *18*(3), 299–323.

Foster, P., Tonkyn, A., & Wigglesworth, G. (2000). Measuring spoken language: A unit for all reasons. *Applied Linguistics*, *21*(3), 354-375.

Foster, P., & Wigglesworth, G. (2016). Capturing accuracy in second language performance: The case for a weighted clause ratio. *Annual Review of Applied Linguistics*, *36*(1), 98-116.

Fotos, S. & Nassaji, H. (Eds.). (2007). *Form-focused instruction and teacher education: Studies in honour of Rod Ellis*. Oxford: Oxford University Press.

Freed, B. (1995). What makes us think that students who study abroad become fluent? In B. Freed (Ed.), *Second language acquisition in a study abroad context* (pp. 123-1481, Philadelphia: John Benjamins Publishing Company.

Freed, B. F. (2000). Is fluency, like beauty, in the eyes of the beholder? In H. Riggenbach (Ed.), *Perspectives on fluency* (pp. 243–265). Ann Arbor: University of Michigan Press.

Freed, B. F., Segalowitz, N., & Dewey, D. P. (2004). Context of learning and second language fluency in French: Comparing regular classroom, study abroad, and intensive domestic immersion programs. *Studies in Second Language Acquisition*, *26*(2)*,* 275–301.

Fu, T. & Nassaji, H. (2016). Corrective feedback, learner uptake, and feedback perception in aChinese as a foreign language classroom. *Studies in Second Language Learning and Teaching*, *6*(1), 159–181.

Gass, S. M. (2003). Input and interaction. In C. J. Doughty & M. H. Long (Eds), *Handbook of second language acquisition.* (pp. 224–225). Oxford: Blackwell.

Gass, S., Mackey, A., Alvarez-Torres, M. J., & Fernández-García, M. (1999). The effects of task repetition on linguistic output. *Language Learning*, *49*(4), 549-581.

Goo, J. (2012). Corrective feedback and working memory capacity in interaction-driven L2 learning. *Studies in Second Language Acquisition*, *34*(3), 445–474.

Goo, J. (2020). Research on the role of recasts in L2 learning. *Language Teaching*, *53*(3), 289-315.

Goo, J., & Mackey, A. (2013). The case against the case against recasts. *Studies in Second Language Acquisition*, *35*(1), 127-165.

Granena, G., & Yilmaz, Y. (2019). Corrective feedback and the role of implicit sequence-learning ability in l2 online performance. *Language Learning*, *69*(S1), 127-156.

Hammerly, H. (1991). *Fluency and Accuracy: Toward Balance in Language Teaching and Learning*. Multilingual Matters.

Havranek, G., & Cesnik, H. (2003). Factors affecting the success of corrective feedback. In S. Foster- Cohen & A. Nizegorodzew (Eds.), *EUROSLA Yearbook, Volume 1.* Amsterdam: Benjamins.

Heaton, J. B. (1966). *Composition through pictures*. Essex: Longman.

High, A. C., & Caplan, S. E. (2009). Social anxiety and computer-mediated communication during initial interactions: Implications for the hyperpersonal perspective. *Computers in Human Behavior*, *25*(2), 475–482.

Housen, A., & Kuiken, F. (2009). Complexity, accuracy, and fluency in second language acquisition. *Applied Linguistics*, *30*(4), 461–473.

Housen, A., Kuiken, F., & Vedder, I. (Eds.). (2012). *Dimensions of L2 performance and proficiency: Complexity, accuracy and fluency in SLA: Volume 32*. Amsterdam: John Benjamins.

Iwashita, N. (2003). Negative feedback and positive evidence in task-based interaction: Differential effects on L2 development. *Studies in Second Language Acquisition*, *25*(1), 1–36.

Iwashita, N., Brown, A., McNamara, T., & O'Hagan, S. (2008). Assessed levels of second language speaking proficiency: How distinct?. *Applied Linguistics*, *29*(1), 24–49.

James, C. (2013). *Errors in language learning and use: Exploring error analysis*. New York: Routledge.

Kim, H. R., & Mathes, G. (2001). Explicit vs. implicit corrective feedback. *The Korea TESOL Journal*, *4*(1), 57–72.

Kormos, J. (1998) *Self-repairs in the speech of Hungarian learners of English*. PhD dissertation, Eotvos University, Budapest.

Kormos, J. (1999). Monitoring and self-repair in L2. *Language Learning*, *49*(2), 303–342.

Kormos, J. (2011). Speech production and the Cognition Hypothesis. In Robinson, P. (Ed.), *Second language task complexity: researching the cognition hypothesis of language learning and performance* (pp. 39–60). Amsterdam: Benjamins.

Kormos, J. (2014). *Speech production and second language acquisition*. New York: Routledge.

Kormos, J., & Dénes, M. (2004). Exploring measures and perceptions of fluency in the

speech of second language learners. *System*, *32*(2), 145-164.

Kormos, J., & Trebits, A. (2012). The role of task complexity, modality, and aptitude in narrative task performance. *Language Learning*, *62*(2), 439-472.

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

Krashen, S. (1982). *Principles and practice in second language acquisition.* Oxford: Pergamon Press.

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

Leeman, J. (2003). Recasts and second language development: Beyond negative evidence. *Studies in Second Language Acquisition*, *25*(1)*,* 37-63.

Lennon, P. (1990). Investigating fluency in EFL: A quantitative approach. *Language Learning*, *40*(3), 387–417.

Levelt, W. J. (1983). Monitoring and self-repair in speech. *Cognition*, *14*(1), 41-104.

Levelt, W. J. M. (1989). *Speaking: From intention to articulation.* Cambridge: MIT Press.

Levelt, W.J.M. (1992). Accessing words in speech production: Stages, processes and representations. *Cognition, 42*(1)*,* 1-22.

Levelt, W.J.M. (1993). Timing in speech production. With special reference to word form encoding. *Annals of the New York Academy of Sciences.*

Levelt, W. J. M. (1999). Producing spoken language: A blueprint of the speaker. In P. Hagoort & C. M. Brown (Eds.), *The neurocognition of language* (pp. 94–122). Oxford: Oxford University Press.

Lightbown, P. M., & Spada, N. (1990). Focus-on-form and corrective feedback in communicative language teaching. *Studies in Second Language Acquisition*, *12*(4), 429–448.

Llinares, A., & Lyster, R. (2014). The influence of context on patterns of corrective feedback and learner uptake: A comparison of CLIL and immersion classrooms. *The Language Learning Journal*, *42*(2), 181–194.

Loewen, S. (2004). Uptake in incidental focus on form in meaning-focused ESL lessons. *Language Learning*, *54*(1), 153-188.

Loewen, S. (2011). Focus on form. In: E. Hinkel (Ed.), *Handbook of research in second language teaching and learning: Volume II* (pp. 576–592). New York: Routledge.

Loewen, S. (2012). The role of feedback. In S. M. Gass & A. Mackey (Eds.), *The R Routledge handbook of second language acquisition*. New York: Routledge, 24–40.

Loewen, S., & Erlam, R. (2006). Corrective feedback in the chatroom: An experimental study. *Computer Assisted Language Learning*, *19*(1), 1–14.

Loewen, S., & Nabei, T. (2007). Measuring the effects of oral corrective feedback on L2 knowledge. In A. Mackey (Ed.), *Conversational interaction in second language acquisition* (pp. 361–377). New York: Oxford University Press.

Loewen, S., & Philp, J. (2006). Recasts in the adult English L2 classroom: Characteristics, explicitness, and effectiveness. *The Modern Language Journal*, *90*(4), 536–556.

Long, M. H. (1983). Native speaker/non-native speaker conversation and the negotiation

of comprehensible input1. *Applied Linguistics*, *4*(2), 126–141.

Long, M. (1991). Focus on form: A design feature in language teaching methodology. In K. de Bot, R. Ginsberg, & C. Kramsch (Eds.), *Foreign language research in cross-cultural perspective* (pp. 39–52). Amsterdam: John Benjamins.

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. (2000). Focus on form in task-based language teaching. In R. D. Lambert & E. Shohamy (Eds.), *Language policy and pedagogy: Essays in honor of A. Ronald Walton* (pp. 179–192). Philadelphia: Benjamins.

Long, M. H. (2007). Recasts: The story so far. In *Problems in SLA*. Mahwah, NJ: Erlbaum.

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

Lyster, R. (1998a). Recasts, repetition, and ambiguity in L2 classroom discourse. *Studies in Second Language Acquisition*, *20*(1), 51-81.

Lyster, R. (1998b). Negotiation of form, recasts, and explicit correction in relation to error types and learner repair in immersion classrooms. *Language Learning*, *48*(2), 183-218.

Lyster, R. (2004). Differential effects of prompts and recasts in form-focused instruction. *Studies in Second Language Acquisition*, *26*(03), 399-432.

Lyster, R., & Ranta, L. (1997). Corrective feedback and learner uptake: Negotiation of

form incommunicative classrooms. *Studies in Second Language Acquisition*, *19*(1), 37–66.

Mayer, M. (1967). *A boy, a dog, and a frog*. New York: Dial Books for Young Readers.

McDonough, K. (2005). Identifying the impact of negative feedback and learners' responses on ESL question 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* (pp. 323–338). New York: Oxford University Press.

McDonough, K., & Mackey, A. (2006). Responses to recasts: Repetitions, primed production, and linguistic development. *Language Learning*, *56*(4), 693-720.

Michel, M. (2017). Complexity, accuracy and fluency in L2 production. In S. Loewen & M. Sato (Eds.), *The handbook of instructed second language acquisition* (pp. 50–68). New York: Routledge.

Monteiro, K. (2014). An experimental study of corrective feedback during video-conferencing. *Language Learning & Technology*, *18*(3), 56-79.

Muranoi, H. (2000). Focus on form through interaction enhancement: Integrating formal instruction into a communicative task in EFL classrooms. *Language Learning*, *50*(4), 617–673.

Myles, F., & Mitchel, R. (2004). *Second language theories*. New York: Hodder Arnold.

Nagata, N. (1993). Intelligent computer feedback for second language instruction. *The Modern Language Journal*, *77*(3), 330–339.

Nassaji, H. (2007). Elicitation and reformulation and their relationship with learner repair

In dyadic interaction. *Language Learning*, *57*(4), 511–548.

Nassaji, H. (2009). Effects of recasts and elicitations in dyadic interaction and the role of feedback explicitness. *Language Learning*, *59*(2), 411-452.

Nassaji, H. (2010). The occurrence and effectiveness of spontaneous focus on form in adult ESL classrooms. *Canadian Modern Language Review*, *66*(6), 907-933.

Nassaji, H. (2011). Immediate learner repair and its relationship with learning targeted forms in dyadic interaction. *System*, *39*(1), 17–29.

Nassaji, H. (2015). The interactional feedback dimension in instructed second language learning: Linking theory, research, and practice. London: Bloomsbury.

Nassaji, H. (2016a). Research timeline: Form-focused instruction and second language acquisition. *Language Teaching*, *49*(1), 35–62.

Nassaji, H. (2016b). Anniversary article Interactional feedback in second language teaching and learning: A synthesis and analysis of current research. *Language Teaching Research*, *20*(4), 535–562.

Nassaji, H. (2017). The effectiveness of extensive versus intensive recasts for learning L2 grammar. *The Modern Language Journal*, *101*(2), 353-368.

Nassaji, H. (2020). Assessing the effectiveness of interactional feedback for L2 acquisition: Issues and challenges. *Language Teaching*, *53*(1), 3-28.

Nassaji, H., & Fotos, S. (2004). 6. Current developments in research on the teaching of grammar. *Annual Review of Applied Linguistics*, *24*(1), 126–145.

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. S. (2011). *Teaching grammar in second language classrooms: Integrating form-focused instruction in communicative context*. London: Routledge.

Nassaji, H., & Kartchava, E. (Eds.). (2017). *Corrective feedback in second language teaching and learning*. New York: Routledge.

Norris, J. M., & Ortega, L. (2009). Towards an organic approach to investigating CAF in instructed SLA: The case of complexity. *Applied Linguistics*, *30*(4), 555–578.

Pallotti, G. (2009). CAF: Defining, refining and differentiating constructs. *Applied Linguistics*, *30*(4), 590–601.

Panova, I., & Lyster, R. (2002). Patterns of corrective feedback and uptake in an adult ESL classroom. *Tesol Quarterly*, *36*(4), 573–595.

Philp, J. (2003). Constraints on" noticing the gap": Nonnative speakers' noticing of recasts in NS-NNS interaction. *Studies in Second Language Acquisition*, 99–126.

Pica, T. (1994). Research on negotiation: What does it reveal about second-language learning conditions, processes, and outcomes?. *Language Learning*, *44*(3), 493–527.

Polio, C. G. (2001). Review of second language development in writing: Measures of fluency, accuracy, and complexity. *Studies in Second Language Acquisition*, *23*(3), 423–425.

Pituch, K. A., Whittaker, T. A., & Stevens, J. P. (2013). *Intermediate statistics: A modern approach*. New York: Routledge.

Rassaei, E. (2019). Recasts during mobile-mediated audio and video interactions: learners' responses, their interpretations, and the development of English

articles. *Computer Assisted Language Learning*, 1-27.

Riggenbach, H. (Ed.). (2000). *Perspectives on fluency*. Ann Arbor: University of Michigan Press.

Rosa, E. M., & Leow, R. P. (2004). Awareness, different learning conditions, and second language development. *Applied Psycholinguistics*, *25*(2), 269.

Sarré, C., Grosbois, M., & Brudermann, C. (2019). Fostering accuracy in L2 writing: impact of different types of corrective feedback in an experimental blended learning EFL course. *Computer Assisted Language Learning*, 1-23.

Saito, K., & Akiyama, Y. (2017). Video-based interaction, negotiation for comprehensibility, and second language speech learning: A longitudinal study. *Language Learning*, *67*(1), 43–74.

Sato, M., & Loewen, S. (2018). Metacognitive instruction enhances the effectiveness of corrective feedback: Variable effects of feedback types and linguistic targets. *Language Learning*, *68*(2), 507–545.

Sato, M., & Lyster, R. (2012). Peer interaction and corrective feedback for accuracy and fluency development: Monitoring, practice, and proceduralization. *Studies in Second Language Acquisition*, *34*(4), 591-626.

Sauro, S. (2009). Computer-mediated corrective feedback and the development of L2 grammar. *Language Learning & Technology*, *13*(1), 96–120.

Schmidt, R. (1992). Psychological mechanisms underlying second language fluency. *Studies in Second Language Acquisition*, 357-385.

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. 1-65). Honolulu: University of Hawai'i Press.

Schmidt, R. (2001). Attention. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 3–32). Cambridge: Cambridge University Press.

Schmidt, R. W (2010). Attention, awareness, and individual differences in language learning. In W. M.Chan, S. Chi, K. N. Cin, J. Istanto, M. Nagami, J. W. Sew, T. Suthiwan, & I. Walker (Eds.), *Proceedings of CLaSIC 2010* (pp. 721–737). Singapore: National University of Singapore, Centre for Language Studies.

Schmidt, R., & Frota, S. (1986). Developing basic conversational ability in a second language: A case-study of an adult learner. In: R. Day (Ed.), *Talking to learn: Conversation in second language acquisition* (pp. 237–326). Rowley, MA: Newbury House.

Segalowitz, N. (2000). Automaticity and attentional skill in fluent performance. In H. Riggenbach (Ed.), *Perspectives on fluency*. Ann Arbor, MI: The University of Michigan Press, 200–219.

Segalowitz, N. (2010). *Cognitive bases of second language fluency*. New York: Routledge.

Segalowitz, N., & Freed, B. F. (2004). Context, contact, and cognition in oral fluency acquisition: Learning Spanish in at home and study abroad contexts. *Studies in Second Language Acquisition, 26*(2), 201–226.

Segalowitz, S. J., Segalowitz, N. S., & Wood, A. G. (1998). Assessing the development of automaticity in second language word recognition. *Applied Psycholinguistics, 19*(1), 53–67.

Sheen, Y. (2006). Exploring the relationship between characteristics of recasts and

learner uptake. *Language Teaching Research*, *10*(4), 361-392.

Shirani, R. (2019). Patterns of uptake and repair following recasts and prompts in an EFL context. *Studies in Second Language Learning and Teaching*, *9*(4), 607–631.

Shintani, N. (2016). The effects of computer-mediated synchronous and asynchronous direct corrective feedback on writing: a case study. *Computer Assisted Language Learning*, *29*(3), 517–538.

Skehan, P. (1996). A framework for the implementation of task-based instruction. *Applied Linguistics*, *17*(1), 38–62.

Skehan, P. (1998). *A cognitive approach to language learning*. Oxford: Oxford University Press.

Skehan, P. (2003). Task-based instruction. *Language Teaching*, *36*(1), 1-14.

Skehan, P. (2009). Modelling second language performance: Integrating complexity, accuracy, fluency, and lexis. *Applied Linguistics*, *30*(4), 510–532.

Skehan, P. (2018). *Second language task-based performance: Theory, research, assessment*. New York: Routledge.

Skehan, P., & Foster, P. (1999). The influence of task structure and processing conditions on narrative retellings. *Language Learning*, *49*(1), 93–120.

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 aren't enough. *Canadian Modern Language Review*, *50*(1), 158–164.

Swain, M. (1995). Three functions of output in second language learning. In G. Cook &

B. Seidhofer (Eds.), *For H. G. Widdowson: Principles and practice in the study of language* (pp. 125–144). Oxford: Oxford University Press.

Swain, M. (2005). The output hypothesis: Theory and research. In E. Hinkel (Ed.), *Handbook on research in second language teaching and learning*. Mahwah, NJ: Erlbaum.

Swain, M. & Carroll, S. (1987). The immersion observation study. In B. Harley, P. Allen, J. Cummins, & M. Swain (Eds.), *The development of bilingual proficiency: Final report. Volume II: Class- room treatment* (pp. 192–222). Toronto: Modern Language Centre, OISE.

Swain, M., & Lapkin, S. (1989). Canadian immersion and adult second language teaching: What's the connection?. *The Modern Language Journal*, *73*(2), 150–159.

Swain, M., & Lapkin, S. (1995). Problems in output and the cognitive processes they generate: A step towards second language learning. *Applied Linguistics*, *16*(3), 371–391.

Tavakoli, P., & P. Skehan. (2005). Planning, task structure, and performance testing in R. Ellis (Ed.): *Planning and task performance in a second language*. John Benjamins.

Tavakoli, P., & Foster, P. (2008). Task design and L2 performance. *Language Learning*, *58*(2), 429-473.

Tomlin, R. S., & Villa, V. (1994). Attention in cognitive science and second language acquisition. *Studies in Second Language Acquisition*, *16*(2), 183–203.

Towell, R., & Dewaele, J. M. (2005). The role of psycholinguistic factors in the

development of fluency amongst advanced learners of French. In J.M. Dewaele (Ed.), *Focus on French as a Foreign language: Multidisciplinary approaches* (pp. 210–239). Clevedon: Multilingual Matters.

Towell, R., Hawkins, R., & Bazergui, N. (1996). The development of fluency in advanced learners of French. *Applied Linguistics*, *17*(1), 84–119.

Van Hest, E. (1996) *Self-repair in L1 and L2 production*. Tilburg: Tilburg University Press.

Vercellotti, M. L. (2017). The development of complexity, accuracy, and fluency in second language performance: A longitudinal study. *Applied Linguistics*, *38*(1), 90-111.

Wang, Y. (2006). Negotiation of meaning in desktop videoconferencing-supported distance language learning. *ReCALL, 18*(1), 122–145.

Wendel, J. (1997). Planning and second language narrative production. Unpublished doctoral dissertation, Temple University, Japan.

Xu, Q., Dong, X., & Jiang, L. (2017). EFL learners' perceptions of mobile-assisted feedback on oral production. *TESOL Quarterly*, *51*(2), 408–417.

Yilmaz, Y. (2012). The relative effects of explicit correction and recasts on two target structures via two communication modes. *Language Learning*, *62*(4), 1134–1169.

Yilmaz, Y. (2013). The relative effectiveness of mixed, explicit and implicit feedback in the acquisition of English articles. *System*, *41*(3), 691–705.

Yilmaz, Y., & Yuksel, D. (2011). Effects of communication mode and salience on recasts: A first exposure study. *Language Teaching Research*, *15*(4), 457–477.

Yuan, F., & Ellis, R. (2003). The effects of pre-task planning and on-line planning on

fluency, complexity and accuracy in L2 monologic oral production. *Applied Linguistics*, *24*(1), 1-27.

Zhao, Y., & Ellis, R. (2020). The relative effects of implicit and explicit corrective feedback on the acquisition of 3rd person-s by Chinese university students: A classroom-based study. *Language Teaching Research*, 1–21. https://doi.org/10.1177/1362168820903343

Appendix A. The Picture Tasks Used on the Pretest, the Immediate Posttest, and the Delayed Posttest

Bicycle Picture Task

Race Picture Task

Tiger Picture Task