Research Paper

Incorporating Feedback Timing into Learning and Retention of Vocabulary and Morpho-Syntax: Focus on Iranian EFL Learners In a Computer-Mediated Communication Context

Zohreh G. Shooshtari[1]
Shahid Chamran University of Ahvaz, Ahvaz, Iran
Khalil Bavizade
Shahid Chamran University of Ahvaz, Ahvaz, Iran
Alireza Jalilifar
Shahid Chamran University of Ahvaz, Ahvaz, Iran

Abstract

This study explored the impact of corrective feedback timing on the learning and retention of lexical items and morpho-syntax markers following a quasi-experimental design. Sixty Iranian EFL learners from three same-level intermediate online classes were initially homogenized and pigeonholed into three equal groups depending on the feedback timing they received: the immediate CF (IC), short-term CF (SC) and delayed CF (DC). Participants in IC received corrective feedback while performing the task (online CF), the learners in SC after completing the task, and in DC a few days later. A *one-way* ANOVA and a post-hoc test were administered to analyze the learners' performance in written and oral pre, post and delayed post-tests. The results demonstrated a significant effect for immediate CF in learning of target structures. On GJT, the immediate CF group performed significantly better than the other two groups. As for OPT, a significant difference existed between immediate and delayed CF groups. The results also revealed IC

---

[1] Corresponding author: z.gshooshtari@scu.ac.ir

group statistically outperformed both SC and DC groups on GJT as well as OPT tests in retention of target features. The benefit observed for the IC group might be due to the memory benefits experienced by the immediate CF group while making a cognitive comparison, accessible as long as the L2 learner could hold a representation of the propositional message.

1. Introduction

Foreign language researchers in general and cognitive psychologists in particular have become increasingly keen on the crucial role of corrective feedback (CF) in second language (L2) learning. According to Swain (as cited in Rahimi & Dastjerdi, 2012), the act of providing CF during oral production can facilitate the process of L2 learning. This topic has received considerable attention in SLA research studies, focusing either on conversational interactions (Amini & Afshari, 2019; Arroyo & Yilmaz 2018; Brown, 2016; Ellis, 2006; Li, 2017a; Li et al., 2016; Loewen & Erlam, 2006; Lyster & Ranta, 1997) or on L2 writing (Bitchener & Storch, 2016; Evans et al., 2010; Lee, 2013; Mackey, 2015; Manchón & Cerezo, 2018).

Several scholars have defined CF as responses to learner's utterances containing an error (Ellis, 2006; Li, 2018; Nakata, 2014; Nassaji, 2016). Li (2019) states that "CF has been at the core of the theory, pedagogy, and research of SLA and its significance is evidenced by the large body of research" (p. 1). Li et al., (2016) point out that previous empirical studies on CF have indicated that different variables including feedback type (e.g., Ellis et al., 2006), type of linguistic target (e.g., Yilmaz, 2012), length of treatment, (e.g., Revesz & Han, 2006) communication mode (e.g., Yilmaz &

G. Shooshtari et al.

Yuksel, 2011) cognitive individual differences (e.g., Yilmaz & Granena, 2010) and task characteristics (e.g., Revesz & Has, 2006) have so far played an important role in changing the level of effectiveness of CF on SLA. However, one more possible feedback-related issue which has not received enough attention and has remained under -researched is feedback timing due to its theoretical and pedagogical significance; that is, whether CF should be supplied immediately after an error is made or delayed after the completion of a communicative task (Li, 2017). Arroyo and Yilmaz (2018) assume that an investigation into the role of CF timing is pedagogically crucial since it can supply the correct answer to the question of whether CF should be supplied during performing a task or be postponed for a later time at the end of task performance or even a few days later.

Despite all that has so far been done, research on feedback timing has come up with mixed and inconclusive findings. For instance, Carpenter and Vul (2011) have pointed out that when the correct answer is supplied immediately after the response, L2 learners may confuse their erroneous production with the correct response and might learn false information. On the contrary, some researchers (e.g., Doughty 2011; Lee, 2013) claim that effective CF is an immediate one because it leads to a better memory performance; in other words, memory performance would naturally act better in this circumstance. Therefore, these scholars came to this conclusion that educational research points to an advantage for immediate CF over delayed feedback in language teaching and learning process. Nakata (2014) argues that language learners generally prefer immediate over delayed feedback since the use of immediate CF has positive effects on L2 learners' motivation and can be more desirable for them. To elucidate these mixed and diverse findings, Henderson (2018) argues that more empirical studies are required into the timing of feedback. Although, the incorporation of CF timing has

recently been probed in some studies (e.g., Amini & Afshari, 2019; Arroyo & Yilmaz, 2018; Chaudron, 1998; Dekeyser, 2007; Li et al., 2016; Nakata, 2014; Nassaji, 2016) it seems there are still some more questions which have remained unanswered. Little, if any, is known as how CF timing may simultaneously influence both lexical items and morpho-syntax learning and retention (Nakata, 2015). This study investigates the influence of incorporating feedback timing on Iranian EFL learners' learning and retention of vocabulary and morpho-syntax markers in a synchronous computer-mediated communication (SCMC) context to contribute to the ongoing research on oral corrective feedback (OCF).

2. Literature review
2.1 The importance of CF timing in SLA

The issue of feedback timing is a crucial instructional aspect of oral CF. CF timing refers to when errors in L2 are treated (Quinn, 2021). According to Quinn (2014), the issue of timing was marginalized in SLA literature when communicative-based approaches such as CLT were at their best in the 1980s. Nassaji (2016) believes the improvements in form-focused instruction introduced new theoretical dimensions to the issue of feedback timing. In this regard, Li (2016) assumes that L2 instructors usually confront the question of whether CF should be supplied immediately after an error is committed or be postponed for a later time. He claims that errors are developmental, and in this sense, errors are similar to those made by children acquiring their L1. Therefore, it takes time to internalize linguistic knowledge. Li (2016) believes that "it may be advisable to assist language learners only with errors which are not repairable via their own internal resources and which persist over time" (p. 197). Some scholars like Willis and Willis (2007) are exclusively in favor of supplying CF during the post-task stage because, in this way, form-focused instruction is contextualized, and L2 learners will not

be predisposed to concentrate on a particular linguistic structure during the task. Willis and Willis (2007) discuss that when linguistic forms are addressed in a pre-task stage, language learners' consequent obsession with form can undermine the primary focus on meaning, which is of overarching importance in a task-based or communicative approach. Regarding what errors to target in the educational settings, Hedge (2000) suggests that L2 teachers should respond to errors which are due to a lack of knowledge about a linguistic item, rather than mistakes which are non-systematic and occur as a result of factors such as fatigue or anxiety. Li (2014) claims that this is easier said than done since it is difficult for L2 teachers to distinguish errors from mistakes in spontaneous classroom discourse. This researcher argues that language teachers are highly recommended to correct only global errors, those which cause communication problems not local ones which do not. This recommendation, however, prioritizes the conversational function of CF and seems to neglect its generally recognized pedagogical importance, which is to supply opportunities for exposure to negative as well as positive evidence and the consolidation of L2 linguistic knowledge.

2.2 Theoretical justification of feedback timing

Most SLA theories do not make overt claims about the ideal time to supply CF (Li, 2018). However, their perspectives on CF timing can be derived by examining their claims about the role of CF in language learning or the mechanism through which CF affects L2 development. Hendrickson (1978) believes that in the audio-lingual method, teachers were expected to correct the error immediately after an error had been committed preventing bad habits from becoming entrenched. But this practice was later challenged by some new approaches like communicative language teaching (CLT) in the 1980s. In CLT, as Freeman (2003) discusses, L2 teachers are highly recommended to avoid immediate correction in the classroom context.

Scrivener (2005) argues that "the general position in mainstream language pedagogy was if the objective is accuracy, then immediate CF is likely to be useful; if the aim is fluency, an immediate correction that diverts from the flow of speaking is less appropriate and applicable" (p. 299).

Some theoretical positions in SLA view immediate CF as not only desirable but also facilitative of interlanguage development. Immediate CF as Morris et al., (1977) assume is usually supported by some theories. One assumption is taken from Long's (1997) Interaction Hypothesis and the priority he has given to focus-on-form. He states that immediate CF can create a window of opportunity, that is, supplying CF immediately allows learners to make a cognitive comparison between their intended message, their error, and an accurate model provided in the CF. Li et al. (2016) state this three-way comparison should occur immediately, or at least within a window of somewhere around 40 seconds, while all three elements are still recent in learners' working memory. The second theory in SLA which takes side with immediate CF is Transfer Appropriate Processing (TAP). In this theory, memory performance is not only determined by the depth of processing, but also by the relationship between how information is initially encoded and how it is later retrieved. Li (2019) argues that the context in which the learned knowledge is retrieved and applied must match the context in which the knowledge is acquired; that is, knowledge is not transferable between domains. Lightbown (2008) discusses in favor of this theory claiming that "we can better remember what we have learned in the classroom if the cognitive processes which are active during learning are similar to those which are active during retrieval" (p. 27). In immediate CF, L2 learners are required to process receptively and/or productively those linguistic parts required to express what they want to utter, which activates

learning processes in a context where they are mainly concentrating on communicating (Li et al., 2016).

Regarding delayed feedback, the first theoretical framework that may account for its effectiveness is the Limited Capacity Model proposed by Skehan (1998). This model maintains that the limited nature of human processing renders it difficult for learners to focus on meaning and form simultaneously (Ellis, 1997; Skehan, 1998; VanPatten, 1990). Thus, as Henderson (2018) claims it is likely that delaying and separating CF from a communicative task where participants must also focus on meaning, prevents overloading the attentional resources and enhances noticing, processing, and usability of the CF. The second theory which offers an account for the benefits of delayed CF is the Reactivation and Reconsolidation Theory proposed by Nader and Einarsson (2010). This view does not predict an advantage of delayed CF over immediate CF but it assumes that when representation in long term memory is activated, it becomes liable and susceptible to change. Applied to CF, when a learner is reminded of a previously learned structure; that is, incorrect (e.g., through error repetition), that incorrect mental representation is reactivated. Then, a new stimulus (e.g., a recast) can become part of the incorrect representation, resulting in a new target-like representation to be stored in long-term memory, leading to L2 learning. Therefore, as Harmer (2007) assumes, delaying CF also has practical value in SLA because it is often considered as a natural occurrence in L2 classrooms. The other practical view in this regard is Preparatory Attention (Ellis, 2006; La Berde et al., 2000). Quinn (2014) points out that this theory proposes that learning is enhanced when complete attention is paid to the task at hand rather than divided between multiple tasks. This theory suggests that when L2 learners do a difficult task they will realize they lack crucial information and thus are more likely to pay attention to this

information when it is subsequently made available to them. In this way the link between retrospective and prospective memory is established, making learning possible (McDaniel et al., 1998).

2.3 Taxonomy of CF strategies

In CF research, initially Lyster and Ranta (1997) distinguished six different types of feedback as follows: recasts, explicit correction, elicitation, clarification requests, repetition, and metalinguistic clues. Li (2019) claims that recasts and explicit correction, which supply the correct form, are called reformulations and the remaining four CF types which withhold the correct form and encourage self-correction are collectively referred to as prompts. Goo and Mackey (2013) assume that reformulations are more effective for learning new linguistic forms while prompts are better at consolidating previously learned forms. Lyster and Ranta (cited in Li, 2019) have discussed that CF can also be divided into explicit and implicit depending on whether learners' attention is drawn to linguistic forms or not. Based on this taxonomy, recasts, elicitation, repetition, and clarification are relatively implicit while explicit correction and metalinguistic feedback are more explicit. Li (2016) declares that recasts and repetition give feedback implicitly, and it is up to the L2 learner to notice that an error was committed, while the other types are explicit in showing that an error occurred. The interpretation of the distinction is related to the setting of the feedback. For instance, an implicit recast may be argued to be explicit in formal classroom settings. Moreover, intonation and visual cues accompanying CF delivery should be taken into consideration. Li (2016) argues that in analyses of CF in classroom settings, recasts turned out to be by far the most frequent technique for error correction because they do not stop the flow of communication. In another taxonomy, Ellis and Shintani (2014) have arranged CF strategies in forms of whether they are input-providing (e.g., recasts & explicit correction)

G. Shooshtari et al.

or output-providing (e.g., clarification requests & elicitation) and also according to whether the corrective force of the feedback is explicit (e.g., explicit correction & elicitation) or implicit (e.g., recasts & clarification requests).

Lyster (2004) relying on the notions of input and output, also categorized CFs as either recast (input-providing CF) or a number of other strategies all aimed at eliciting output from the students (output-providing CF). This taxonomy was in turn modified in Lyster and Ranta's (2007) study where two broad CF categories were suggested. Reformulations comprised recast and explicit correction supplied. In prompts, on the other hand, which includes a wide range of different CF strategies such as elicitation, metalinguistic clues, clarification request and repetition, the correct form is withheld and clues are instead supplied to prompt learners to retrieve the correct form from their existing knowledge (Lyster & Saito, 2010). In the latest classification of CFs, while acknowledging the input and output providing categorization of CF, Sheen and Ellis (2011) maintained that CF can be implicit (e.g., recast) or explicit (e.g., explicit correction). Ellis (2010) linked the implicit-explicit distinction to the arguments about input saliency and noticing in language teaching. He continued that there is no consensus as to which of the CF classifications presented is the most theoretically valid. It is worth noting that in this empirical study, the type of OCF which was operationalized was a combination of repetition of the learner's error followed by an explicit correction. A repetition of an error was selected to precede the explicit correction in an attempt to make the CF move more salient. This hybrid CF as Henderson (2018) believes is to be effective and salient to L2 learners and can be employed effectively.

2.4 The role of retention in SLA

Undoubtedly, one of the crucial aspects of language learning is the retention of previously learned materials. Richards and Platt (1992, p. 457) define retention as "the ability of the learners to remember things after a while". Bartle (cited in Kess, 1992, p. 188) states that "retention is not a reproduction of the important ideas, but rather than that, it is an influential reconstruction". Here a distinction should be made between short-term recall and long-term retention. Short-term or immediate recall is the ability of the learners to remember the learned materials immediately at the end of the treatment in each session. This is based on Laufer's (2007) definition for short term recall. She assumes that "it is measured immediately after performing a task which is expected to entail retaining some information, after a short intervention" (p. 29). For instance, a language teacher may teach some lexical or grammatical items then he can assess the recall of them immediately or make the learners do another task for about 10-20 minutes then tests the target features. In the case of long-term or delayed retention, there is no straightforward definition of how much time should elapse between initial learning and the time we can test L2 learners on the target items. As far as long-term retention is concerned in this study, we adopt the definition provided by Laufer (2007). She states that "some people administer a test a week or two later, some a month or even three months, some people repeat a measurement several times to check how much learners retain at different points of time" (p. 30). In this empirical study, the long-term or delayed retention was operationalized as administering the delayed post-test to participants quite after three weeks from the time the treatment ends to see whether retention of the target features have taken place in an SCMC mode or not.

2.5 Related research on CF timing

The effectiveness of feedback timing has always been the subject of remarkable empirical works in instructed SLA in the last few years (Nassaji, 2016). Yilmaz (2016) claims that CF timing studies have been conducted either in a computerized mode or in a face-to-face (FTF) context. Research findings suggest that the timing of feedback can significantly affect L2 learning (Lee, 2013; Li, 2010; Lyster et al., 2013; Nakata, 2015). The previously done studies have so far yielded diverse and contradictory results; some works (e.g., Amini & Afshari, 2019; Henderson, 2018) revealed no significant difference between immediate and delayed CF whereas some others (e.g., Henshaw, 2011; Nakata, 2015; & Quinn, 2014) indicated that both immediate and delayed CF are identically effective on SLA. It should be mentioned that some research (e.g., Farmani et al., 2017; Fu & Nassaji, 2016; Shabani & Safari, 2016) have demonstrated advantages for immediate CF whereas the results of a few works have indicated that delayed CF is more beneficial for L2 learning than immediate CF (e.g., Butler et al., 2007; Rahimi & Dastjerdi, 2012). Summary of research findings on corrective feedback timing is as follows:

Identical impact of both immediate delayed CF on SLA	Arroyo & Yilmaz, 2018; Lavolette (2014) and & Li et al. (2016); Nakata (2014); Vamosfadrani (2006)
The effectiveness of immediate CF over delayed CF on SLA	Farmani et al., 2017; Fu Nassaji (2016); Shabani & Safari (2016a);
The effectiveness of delayed CF over immediate CF on SLA	Rahimi & Dastjerdi (2012) Butler et al. (2007)
No difference between delayed CF and immediate CF	Henshaw (2011); Nakata, (2015); Quinn (2014)

The above brief review on the effectiveness of immediate and delayed CF on the acquisition of L2 forms have demonstrated conflicting views and mixed results. Therefore, more empirical studies are required into the timing of feedback to clarify these inconsistent findings. The present study

following the same avenue of research aims at examining the impact of CF timings on Iranian EFL learners' learning and retention of lexical items and morpho-syntax markers in a computerized context.

The present empirical work seeks to answer the following research questions:

1. Do different timings (immediate, short-term & delayed) lead to differential learning and retention of vocabulary and morpho-syntax in a synchronous computer-mediated communication context among Iranian EFL learners?
2. Is there any significant difference between different timings on learning and retention of vocabulary and morpho-syntax in a synchronous computer-mediated communication context among Iranian EFL learners?

3 Method

3.1 Study design

A pretest-posttest-delayed-posttest quasi-experimental design was adopted to conduct this research. Three intact groups (immediate CF group, short-term CF group and delayed CF group) were chosen from three same-level intermediate online classes. The control group was not included since in some previous related investigations (e.g., Ammar & Spada, 2007; Li et al., 2016) participants in the control group were deliberately deprived of CF which can be considered as a deficiency. The study lasted for eighteen sessions, around ninety minutes a session, twice a week, in about three months.

3.2 Participants

Sixty participants in Ahvaz, south west of Iran, were accessible for the present study who were selected based on non-random convenience sampling. The Preliminary English Test (PET) test was administered to examine the sample population's homogeneity. Eighty-two participants were initially recruited from among whom sixty learners were chosen and pigeonholed into three intact groups, depending on the treatment they

received: the immediate CF group (IC), short-term CF group (SC) and delayed CF group (DC). Each group included twenty participants. Touchstone series (McCarthy et al., 2006) were taught at the target language institute. All the participants had already studied and passed course books such as First Friends (Lanuzzi, 2015), Get Ready (Hopkins, 2003) and English Time (Rivers & Tayama, 2015). They studied Touchstone 4 during the semester.

3.3 Instruments

Language proficiency test: The Preliminary English Test (2020) was used to ensure the participants' homogeneity. The test includes a reading section with multiple-choice items, matching, and cloze test, a writing section, a listening part, including a recorded text and an oral interview part.

Pre-test, post-test and delayed post-test: These tests assessed the participants' knowledge of the target structure and consisted of three types of tasks: an untimed Grammaticality Judgement Test (GJT), Elicited Imitation Test (EIT) and Oral Picture Description (OPD) tasks as outcome measures. The pretest, post-test, and delayed post-test were devised by the researcher, following the related CF studies in designing measures in instructed L2 development (Ellis, 2005; Li et al., 2016). The same items of these three tests were randomly scrambled to minimize the potential of order effects.

The GJT included thirty items, twenty of which were related to the target structures of Touchstone 4 and ten were distractors relating to the same morpho-syntax the participants were taught. Among the thirty items, twenty-five items were ungrammatical, and five items grammatical. In scoring the participants' responses, one point was given if an ungrammatical sentence was assessed to be ungrammatical and the correct form was supplied by the learner immediately. The internal reliability for the GJT, indexed by Cronbach's alpha, was .9 for the pre-test and 0.95 for both immediate and

delayed post-tests. In taking ELT, L2 learners were required to judge the veracity of a sentence and then repeat it orally within the time allocated. These utterances were different from the ones used in each lesson's exercises. The EIT consisted of thirty statements of which twenty were related to the target structure of Touchstone 4 and ten were distractors. Twenty-five of the target items in EIT were ungrammatical, and five grammatical. In scoring EIT, one point was given to a correct response and since self-corrections may involve the use of explicit knowledge, only the first attempt was immediately scored. A reliability analysis showed that the Cronbach's alpha for the EIT was 0.70, 0.76, and 0.78 for the pre-test, post-test, and delayed post-test. The OPD task consisted of pictures different in order from those in the treatment task but it contained the same target items. Below the pictures, there were some English questions   asking the participants to name and describe the pictures through a PowerPoint file. The Cronbach's alpha value for the OPD was 0.73, 0.75, and 0.76 for the pre-test, post-test, and delayed post-test, respectively.

To validate the tests items, three experienced Ph.D. English instructors were asked to precisely comment on all pre-planned sentences in terms of grammaticality and wording. The finalized piloted items were composed by the researcher, drawing on the learners' English textbook. The immediate post-test was assigned two days after 18 sessions of instruction to assess the target feature knowledge as well as the durability of learning from feedback timing whereas the delayed post-test was exactly administered three weeks after treatment to assess the retention on the target structure. The delayed post-test was administered without prior notice so that the participants would not go over the target structure between the immediate post-test and delayed post-test period.

3.4 Procedure

At first, the researcher assigned each of the intact classes to one treatment condition: the immediate CF group (IC), short-term CF group (SC) and delayed CF group (DC). Because of pandemic (COVID-19) in Iran, each online class session met for about ninety minutes a day, two days a week. At the onset of the study, PET was administered and the participants whose scores fell 2 standard deviations above and below the mean were chosen. The pre-test was administered in Week Three. Then, three distinctive treatments were employed for the groups. In each session, the researcher focused on new target items by utilizing a PowerPoint file on Skyroom. The main task was a summary retelling task which required participants to study the reading passage and then take turns in retelling the text individually. Learners' errors in vocabulary or morpho-syntax, provided the teacher with the opportunities of CF. The type of CF was a combination of repetition of the learner's error followed by an explicit correction. The three groups of the study were assigned the same materials and instruction, but received different CF timing. In IC group, erroneous utterances were immediately corrected during each summary retelling task. The participants in SC were provided with teacher's comments on their erroneous speech at the end of the session. The participants in DC were reminded of the sentential contexts of their deviation and their committed errors in the next online session and then received the same CF. The first post-test was administered to all groups two days later and after three weeks, the second posttest (delayed posttest) was assigned to all groups to measure possible occurring retention. All steps were recorded via SD screen recorder app or X-recorder, transcribed and strictly coded for further precise analysis. Descriptive statistics and a one-way ANOVA test were used to check any within and between group differences. The pre-test and first post-test scores of the three groups as well as the first and the

delayed post-test scores were compared to see the possible impacts of different types of feedback timing on learning and retention of the target structures.

4. Results

To investigate the differential effect of feedback timing on the learners' learning and retention of the vocabulary and morpho-syntax items, the possible variance was examined from pre-test to post-test and ultimately to delayed post-test. First, the results of PET including the test of homogeneity of the variances, and ANOVA are presented.

Table 1
ANOVA test for comparing the groups' scores on PET

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	57.6	2	28.8	.16	.85
Within Groups	10104	57	177.26		
Total	10161.6	59			

As shown in Table 1, no statistically significant differences among the groups in relation to language proficiency test was found $(p < .85)$.

First research question

To answer the first research question, the three groups' mean scores on the pre-test were submitted to *a one-way* ANOVA test.

Table 2
Descriptive statistics for the groups' scores on pre-test (GJT, EIT & OPD)

	Grammatically Judgment Test			**EIT and OPD Tests**	
	N	Mean	SD	Mean	SD
immediate CF group	20	7.42	2.41	15.60	3.61
short-term CF group	20	7.85	3.55	15.30	4.24
delayed CF group	20	7.65	2.62	14.25	3.78

Table 2 presents the three groups' mean scores on the pre-test which indicates approximately similar performance on ELT and OPD tests.

G. Shooshtari et al.

Table 3

Normality test one sample Ks for pre-test scores of GJT and OPD

	Tests of Normality							
	Kolmogorov-Smirnova			Shapiro-Wilk				
	Statistic	df	Sig.	Statistic	df	Sig.	Skewness	Kurtosis
Pre-immediate GJT	.128	20	.21	.958	20	.510	.145	-.78
Pre-short-term GJT	.148	20	.08	.913	20	.071	1.21	1.78
Pre-delayed GJT	.147	20	.301	.944	20	.290	.406	-.59
Pre-immediate OPD	.096	20	.898	.970	20	.763	-.299	-.517
Pre-short OPD	.120	20	.345	.955	20	.443	.797	.63
Pre-delayed OPD	.124	20	.298	.944	20	.288	.628	.003

To answer the research question, the data was checked to ensure normal distribution. As seen in the table 3, the Skewness values ranged from −.29 to 1.2, and the Kurtosis values from −.78 to 1.7, indicating normal distribution. Also, test of homogeneity of variances for GJT, EIT and OPD revealed that the level of significance is set at 0.31 and 0.36 for GJT and 0.83 and 0.84 for EIT and OPD. So, the condition of homogeneity of variance in all groups was also established.

Table 4

The results of ANOVA for GJT, EIT and OPD on pre-test

ANOVA					
pretest (Grammatically Judgment Test)					
	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	.533	2	.267	.032	.969
Within Groups	481.65	57	8.45		
Total	482.18	59			
pretest (EIT & ODP Tests)					
	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	20.1	2	10.05	.66	.51
Within Groups	862.75	57	15.13		
Total	882.85	59			

Incorporating Feedback …

Since the level of significance was set at 0.96 and 0.51 in written and oral tests, insignificant difference was observed among the three groups in the pre-test phase.

Table 5
Descriptive statistics for GJT, EIT and OPD on post-test scores

	Grammatically Judgment Test			EIT and OPD Tests	
	N	Mean	SD	Mean	SD
immediate CF group	20	19.20	4.61	33.85	8.31
short-term CF group	20	15.20	3.48	28.7	8.13
delayed CF group	20	15.25	4.11	26.25	7.98

Table 5 presents the mean scores on the GJT, EIT and OPD tests, which indicate improved performance in each group from pre-test to post-test. Also, the immediate CF group outperformed both the short-term CF and delayed CF groups.

Table 6
Normality test for post-test scores of GJT and OPD

	Tests of Normality							
	Kolmogorov-Smirnov[a]			Shapiro-Wilk				
		df	Sig.	Statistic	df	Sig	Skewness	Kurtosis
Post immediate GJT	.167	20	.144	.950	20	.361	-.042	.125
Post short-term GJT	.223	20	.060	.853	20	.066	1.69	1.81
Post delayed GJT	.208	20	.094	.857	20	.097	.738	-.923
Post immediate OPD	.148	20	.201	.963	20	.600	-.020	-.886
Post short-term OPD	.211	20	.090	.901	20	.063	-.632	-.388
Post delayed OPD	.174	20	.114	.963	20	.598	.004	-.255

As it is clear in the Table 6, the scores related to the post-test stage are normal and the data can be used for parametric analysis.

G. Shooshtari et al.

Test of homogeneity of variances for GJT, EIT and OPD post-tests revealed that   the level of significance based on mean and trimmed mean on GJT is set at 0.34 and 0.30 for GJT and 0.89 and 0.89 or oral post-test. This indicates that the variance of all three groups was homogeneous and statistically insignificant on post-test. To check for any significant difference within each of the groups, pairwise comparison of the means was done as presented below.

Table 7
Pairwise Comparisons of Mean Scores for Immediate CF group (GJT)

(I) factor1	(J) factor1	Mean Difference (I-J)	Std. Error	Sig.[b]	95% Confidence Interval for Difference[s]	
					Lower Bound	Upper Bound
1	2	-8.767*	.643	.000	-10.35	-7.183
	3	-12.583*	.678	.000	-14.25	-10.913
2	1	8.767*	.643	.000	7.183	10.351
	3	-3.817*	.650	.000	-5.418	-2.215
3	1	12.583*	.678	.000	10.913	14.254
	2	3.817*	.650	.000	2.215	5.418

As shown in Table 7, *the immediate CF group* (GJT) was compared in the written test in three stages: pre-test, post-test and delayed post-test, and a meaningful difference was observed among the three stages ($p \leq 0.5$).

Table 8
Pairwise Comparisons of Mean Scores for Short-Term CF Group (GJT)

Measure: **short-term CF group (GJT)**						
(I) factor1	(J) factor1	Mean Difference (I-J)	Std. Error	Sig.[b]	95% Confidence Interval for difference[s]	
					Lower Bound	Upper Bound
1	2	-7.350*	1.148	.000	-10.363	-4.337
	3	-11.050*	1.040	.000	-13.780	-8.320
2	1	7.350*	1.148	.000	4.337	10.363
	3	-3.700*	.886	.002	-6.025	-1.375
3	1	11.050*	1.040	.000	8.320	13.780
	2	3.700*	.886	.002	1.375	6.025

As seen in Table 8, *the short-term CF group* (GJT) was compared in the written test in three stages; pre-test, post-test and delayed post-test and

meaningful difference was observed from pretest to posttest and delayed posttest ($p \leq 0.5$).

Table 9
Pairwise comparisons of mean scores for delayed CF group (GJT)

(I) factor1	(J) factor1	Mean Difference (I-J)	Std. Error	Sig.[b]	95% Confidence Interval for difference	
					Lower Bound	Upper Bound
1	2	-7.600*	.796	.000	-9.69	-5.51
	3	-11.400*	1.159	.000	-14.44	-8.35
2	1	7.600*	.796	.000	5.51	9.69
	3	-3.800*	1.370	.036	-7.39	-.2
3	1	11.400*	1.159	.000	8.35	14.4
	2	3.800*	1.370	.036	.2	7.39

According to Table 9, *the delayed CF group* (GJT) was compared in the written test in three stages; pre-test, post-test and delayed post-test and a meaningful difference was observed among the three steps ($p \leq 0.5$).

Table 10
Pairwise Comparisons of Mean Scores for Immediate CF group (OPT)

(I) factor1	(J) factor1	Mean Difference (I-J)	Std. Error	Sig.[b]	95% Confidence Interval for difference[s]	
					Lower Bound	Upper Bound
1	2	-14.550*	1.108	.000	-17.281	-11.819
	3	-18.683*	1.163	.000	-21.549	-15.817
2	1	14.550*	1.108	.000	11.819	17.281
	3	-4.133*	1.472	.020	-7.761	-.506
3	1	18.683*	1.163	.000	15.817	21.549
	2	4.133*	1.472	.020	.506	7.761

Based on the Table 10, *the immediate CF group* (OPT) was compared in three stages of pre-test, post-test and delayed post-test in the oral test and a significant difference was observed between all three stages ($p \leq .5$).

Table 11
Pairwise Comparisons of Mean Scores for Short-Term CF group (OPT)

(I) factor1	(J) factor1	Mean Difference (I-J)	Std. Error	Sig.[b]	95% Confidence Interval for difference	
					Lower Bound	Upper Bound
1	2	-13.400*	1.81	.000	-18.17	-8.62
	3	-15.850*	1.77	.000	-20.5	-11.19
2	1	13.400*	1.81	.000	8.626	18.17
	3	-2.450	2.24	.863	-8.33	3.43
3	1	15.850*	1.77	.000	11.19	20.5
	2	2.450	2.24	.863	-3.43	8.33

As shown in Table 11, the *short-term CF group* (OPT) was compared in the oral test in three stages: pre-test, post-test and delayed post-test, and no significant difference was reported between the second and third stages ($p = .863 \geq 0.5$)., but there was a significant difference between the first stage and the other two stages ($p \leq 0.5$).

Table 12
Pairwise Comparisons of Mean Scores for Delayed CF Group (OPT)

(I) factor1	(J) factor1	Mean Difference (I-J)	Std. Error	Sig.[b]	95% Confidence Interval for difference[s]	
					Lower Bound	Upper Bound
1	2	-12.000*	1.814	.000	-16.76	-7.239
	3	-18.150*	1.884	.000	-23.09	-13.20
2	1	12.000*	1.814	.000	7.23	16.76
	3	-6.150	2.533	.076	-12.8	.50
3	1	18.150*	1.884	.000	13.2	23.09
	2	6.150	2.533	.076	-.5	12.8

Based on Table 12 *the delayed CF group* (OPT) was compared in three stages of pre-test, post-test and delayed post-test in the oral test and no significant difference was reported between second and third steps. The rest of the steps have been significantly different from each other.

Second research question: part one

To answer the first part of second research question (detecting any significant difference between different timings on learning of vocabulary

and morpho-syntax) the three groups' mean scores on the post-test were submitted to a *one-way* ANOVA.

Table 13

The Results of ANOVA for GJT, EIT and OPD on Post-Test

ANOVA					
Post-test (Grammatically Judgment Test)					
	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	210.700	2	105.35	6.267	.003
Within Groups	958.150	57	16.81		
Total	1168.850	59			

Post-test (EIT & ODP Tests)					
	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	601.9	2	300.95	4.535	.015
Within Groups	3782.5	57	66.36		
Total	4384.4	59			

As shown in Table 13, since the level of significance set at 0.003 and 0.015 which are less than 0.05, there is a statistically significant difference among the three groups' performances in the post-test, but to locate the difference, follow-up tests and multiple comparisons were used.

Table 14

Multiple Comparisons for GJT

Bonferroni				
(I) groups	(J) groups	Mean Difference (I-J)	Std. Error	Sig.
immediate CF group	Short-term CF	4.00	1.29	.009
	delayed CF group	3.95	1.29	.000
short-term CF group	delayed CF group	1.75000	3.93517	1.000

Table 15 indicates that on GJT test, the mean difference between the immediate CF group and the other two groups is statistically significant ($p=$ 009,001$\geq$0.05), and by examining the mean difference between short-term CF and delayed CF groups, no meaningful difference is observed. As for OPT, as shown in Table 15 the significant difference is between immediate and delayed CF group but the mean difference between immediate and short-term CF group is insignificant ($p = .151 \geq .05$).

Table 15

Multiple Comparisons for OPT

Bonferroni

(I) groups	(J) groups	Mean Difference (I-J)	Std. Error	Sig.
immediate CF group	Short-term CF	5.15	2.57	.151
	delayed CF group	7.600	2.57	.014
short-term CF group	delayed CF group	2.45	2.57	1.000

Second research question: part two

To detect any significant difference between different timings on retention of vocabulary and morpho-syntax, the three groups' mean scores on the delayed post-test were submitted to a *one-way* ANOVA.

Table 16

Descriptive Statistics for the Three Groups' Scores on Delayed Post-Test

	Grammatically Judgment Test			EIT and OPD Tests	
	N	M	SD	M	SD
Immediate CF Group	20	23.15	3.82	37.65	9.80
Short-Term CF Group	20	18.90	3.59	31.15	6.80
delayed CF group	20	19.05	5.83	32.40	9.24

Incorporating Feedback …

As shown in Table 16, the delayed posttest mean scores and standard deviations in the GJT for the immediate CF group are ($M = 23.15$, $SD = 3.82$) for the short-term CF group ($M = 18.90$, $SD = 3.59$) and for the delayed CF group ($M = 19.05$, $SD = 5.83$). Statistics related to EIT and OPD tests for the immediate CF group are ($M = 37.65$, $SD = 9.80$) the short-term CF group are ($M = 31.15$, $SD = 6.80$) and for the delayed CF group ($M = 32.40$, $SD = 9.24$). These results indicate that the immediate CF group performed better than the other two groups.

Table 17
Normality Test One Sample Ks for Delayed Post-Test Scores of GJT and OPT

	Tests of Normality							
	Kolmogorov-Smirnov[a]			Shapiro-Wilk				
	Statistic	df	Sig.	Statistic	df	Sig.	Skewness	Kurtosis
Delayed post immediate GJT	.188	20	.063	.91	20	.08	-.884	.85
Delayed post short GJT	.156	20	.070	.95	20	.378	-.362	-.72
Delayed post delayed GJT	.156	20	.073	.93	20	.196	.917	.91
Delayed immediate ODP	.158	20	.068	.94	20	.278	-.001	-1.05
Delayed short ODP	.150	20	.069	.95	20	.389	.29	-.54
Delayed post delayed ODP	.181	20	.084	.942	20	.259	.771	.479

Regarding normal distribution, the Skewness values ranged from −.01 to .91, and the Kurtosis from −1.05 to .91, indicating normal distribution of the delayed post-test scores.

Test of homogeneity of variances for two delayed posttests (GJT, EIT & OPD) was run. The results revealed that since the level of significance based

G. Shooshtari et al.

on mean and trimmed mean set at .14 and .14 for the written test and 0.14 and 0.15 for the oral test, it can be argued that the condition of homogeneity of variance has been established in all the three groups.

Table 18

The Results of ANOVA for GJT, EIT and OPD on Delayed Post-Test

ANOVA

Delayed Posttest (Grammatically Judgment Test)					
	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	232.63	2	116.317	5.66	.006
Within Groups	1171.3	57	20.549		
Total	1403.93	59			
Delayed Posttest (EIT & ODP Tests)					
	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	475.833	2	237.917	3.334	.043
Within Groups	4067.900	57	71.367		
Total	4543.733	59			

As can be seen in Table 18, since the level of significance is set at .006 and $0.043 \leq 0.05$, there is a statistically significant difference between the performances of the three groups in the delayed post-test. To locate the difference, follow-up multiple comparisons were conducted.

Table 19

Multiple Comparisons for GJT And OPT on Delayed Post-Test Scores

Multiple Comparisons for GJT				
Dependent Variable: delayed posttest				
Bonferroni				
(I) groups	(J) groups	Mean Difference (I-J)	Std. Error	Sig.
immediate CF group	Short-term CF	4.25	1.43	.000
	delayed CF group	4.10	1.43	.000
short-term CF group	delayed CF group	-.15	1.43	1
Multiple Comparisons for OPT				

Incorporating Feedback …

Dependent Variable: delayed posttest
Bonferroni

(I) groups	(J) groups	Mean Difference (I-J)	Std. Error	Sig.
immediate CF group	Short-term CF	6.500	3.33532	.054
	delayed CF group	5.25	3.33532	.163
short-term CF group	delayed CF group	-1.250	3.33532	1

As shown in Table 19, a meaningful difference existed between the immediate and the other two CF groups in their performance on GJT (*p*:.000 ≤ 0.05) in the delayed posttest phase. That is, the participants in the immediate CF group performed better in the retention of target structures. As for OPT no significant difference was found between immediate and delayed CF groups (*p*:163 ≥ .0.05), but a significant difference was observed between immediate and short-term groups (*p* < .05).

5. Discussion

The results of the first research question revealed overall positive impact of feedback timing on the participants' learning and retention of the target structures since significant improvement of mean scores for the three groups from pre-test to post-test and delayed post-test was observed. In spite of timing, the corrective feedback was also an effective factor in both written and oral performance of the three groups. Amini and Ashrafi (2019) claim that the observed effectiveness of CF timing on target features can be argued in terms of three crucial elements in CF research conditions, namely, the type of CF, the duration of the treatment and the type of outcome measures. The type of CF employed, in the present study, was a hybrid CF package composed of learner's error repetition followed by an explicit correction technique. Some scholars (e.g., Li, 2013; Mackey & Philp, 1998) have argued that explicit CF appears to be more effective than implicit CF in the

teaching and learning process whereas some other researchers (e.g., Lyster et al., 2013) have assumed that "a variety of CF types is probably more effective than consistent use of only one type (p. 21)". They continue to claim that to date the results of research are not conclusive in this concern.

According to Amini and Ashrafi (2019) the duration of the treatment is one of the most controversial issues that should be taken into serious consideration when it comes to discussing the effectiveness of CF timing on the target structures. In this regard, Li et al., (2016) have cited, "it is crucial to expose L2 learners to longer instructional treatments since the improvement of L2 implicit and explicit knowledge requires both extensive and intensive exposure to linguistic input" (p. 12). Thus, it is rational to claim that the duration of CF treatment is a crucial topic that must be regarded when interpreting the results of an investigation related to the effectiveness of CF timing. This assertion is also consistent with Ellis's (2018) claim who believes that longer instructional treatment and adequate exposure to linguistic input can result in better development and enhancement of explicit and implicit knowledge in the acquisition of new items. Therefore, we can assume that in contrast to some previously conducted studies (e.g., Farrokhi et al, 2018; Rahimi & Dastjerdi, 2012) increasing the length of treatment to eighteen sessions, as done in the study, led to better enhancement in learning and retention of newly learnt items.

Concerning the results of the first research question, a significant improvement from pre to delayed post-test in the written test was observed but in oral production tests the improvement was from pre to post-test. Arroyo and Yilmaz (2018) discuss that since the instructions on the GJT asked participants to regard the accuracy of test items and to correct ungrammatical sentences, the task conditions were conducive to directing focal attention toward the target features and to consciously inspecting the

rules by which the forms operate whereas the instructions in the oral production test asked language learners to describe distinctions between two objects as fast as they could. These instructions might have created conditions more favorable for language learners' attention to be primarily on meaning and message creation and less favorable for them to reflect consciously on the target structures. These observations indicate that when task conditions favored a deliberate reflection on forms, all groups were able to utilize their knowledge. As far as the outcome measurement instrument is concerned, the results of this work are in line with some of the previously conducted research (Farmani et al., 2017; Shabani & Safari, 2016). Most of the empirical studies have highlighted the acquisition measurement tool as a source of variation. For instance, Arroyo and Yilmaz (2018) have cited that the developmental gains of Spanish L2 forms due to immediate CF were significantly higher than delayed CF when the acquisition was measured with an oral production task. However, no differential effect was seen when the outcome was only measured with an untimed GJT.

The first part of second research question aimed at detecting significant changes among different timings on the learning of vocabulary and morpho-syntax markers in a computerized mode. Results of ANOVA showed that the immediate CF group's performance on post-test was significantly higher than those of the short-term CF and delayed CF groups. Multiple comparison results indicated that on GJT, the mean difference between short-term CF group and delayed CF group was insignificant. As for OPT, immediate CF outperformed other groups but the mean difference between short-term CF group and delayed CF group was insignificant.

The benefit observed for the immediate CF group on the post-test might be due to the memory benefits experienced by the immediate CF group while making a cognitive comparison. Doughty, 2011 (as cited in Arroyo &

Yilmaz, 2018; Henderson, 2019) assumes that a cognitive comparison would occur within a cognitive window which is accessible for about 40 seconds after the corrective feedback as long as the L2 learner could hold a representation of the propositional message, his/her own non-target like utterance, and the word carrying the CF. However, the findings that immediate CF group outperformed short-term CF group and delayed CF group indicates that taking advantage of CF requires more than a visual comparison between the two forms on the computer screen. Arroyo and Yilmaz (2018, p. 22) claim that "instead of a visual comparison, it is possible that language learners compare the memory traces of their erroneous productions and the information they extracted from the CF". If this was the case, the short delay between the error and CF in the immediate condition might have permitted memory traces of the erroneous utterances to remain active in L2 learners' working memory until they received the feedback. These researchers have discussed that in the delayed feedback condition, however, memory traces of errors, which might have been active immediately after the error was committed, may have decayed by the time L2 learners received the delayed CF, and seeing their errors again on a computer screen in a decontextualized way might not have been enough to activate the memory of having made the error. In other words, language learners may not have been capable of linking the error re-presented to them in the delayed CF stage to the error they made during their task performance.

Moreover, the advantage found for the immediate CF over the short-term CF and delayed CF in this study might be related to the more favorable conditions for hypothesis testing in the immediate-feedback condition. Even though the three groups had equal chances to produce output during the treatment, there were distinctions between the groups as to whether the CF could influence these opportunities. The immediate CF group participants,

depending on when exactly they received the first CF during the task, had some of their production opportunities before the feedback and some after the feedback. In these opportunities after the feedback, they could utilize the information supplied in the CF to form and test new hypotheses about how the target form works. The short-term CF and delayed CF groups, however, had all of their production opportunities before they received CF and thus could not put their newly formed hypotheses to the test through output production. Fu and Li (2020) state that some theoretical views in SLA (e.g., skill acquisition theory) regard immediate CF as both favorable and facilitative of interlanguage development. Li et al., (2019) discuss that the differentiated acquisition from immediate CF seen in this work is in line with the claims proposed by Interaction hypothesis which gives priority to focus-on-form and TAP theory which claim that memory performance is not only determined by the depth of processing, but also by the relationship between how information is initially encoded and how it is later retrieved. Different types of CF employed as treatment involved implicit processing while the output measurement (GJT) involved explicit knowledge of target structures. Accordingly, significant developmental distinctions were expected. Thus, it can be assumed that different timings might have resulted in different outcomes and might have led to better learning of target features.

With respect to the impact of the feedback timing on the retention of the target structures, the results revealed that statistically, immediate CF group outperformed both short-term CF and delayed CF groups on GJT as well as OPT delayed post-test scores. As discussed before, a number of SLA theories like the Interaction Hypothesis, focus-on form, transfer appropriate processing and skill acquisition theory lend support to immediate CF while some other theories in cognitive psychology such as preparatory attention, memory theory, reactivation and reconsolidation theory predict that delayed

feedback is more effective. This study demonstrated that immediate corrective type of feedback to a great extent contributed to the development of target features. To justify the superiority of the immediate CF in the development of target structures, it should be mentioned that some distinctions were made between errors and feedback in terms of feedback timing. In immediate CF, the CF was supplied right away following the error but in the short-term CF and delayed CF groups, CF was provided after task performance or in the next online session. The main difference lay in the contextual nature of the CF. In the immediate CF group, language learners received CF on the errors they had just committed as they struggled to reconstruct the narratives; they had the chance to edit their utterances during each summary retelling task while in the short-term CF the participants did not receive any instant forms of CF during summary retelling task. Rather, the instructor took notes of his learners' erroneous speech and at the end of the session, they were supplied with CF. In the delayed condition, the participants were reminded of the sentential contexts of their deviation and their committed errors in the online session and then received the same CF. In the immediate condition, the CF was connected continuously to the language learners' attempts to retell the story and to produce newly learnt items correctly. That is, the L2 learners had the opportunity to use the CF they had received when producing language involving the new vocabulary or morphosyntax markers as they continued to re-tell the story whereas in the short-term CF and delayed CF groups, the language learners did not produce their own sentences and only used the corrected sentences which were presented by their teacher.

Skill acquisition theory can assist to explain why the contextualized condition was more effective in the retention of target structures. DeKeyser (1998) insisted on the importance of language learners using their declarative

knowledge as a crutch to support their attempts to communicate. In the current study, this knowledge helped to embed the declarative representations of the target structure more deeply in the learners' memories, which were therefore better sustained over time. The results of this study differ from other studies that have investigated immediate and delayed feedback. In particular, a few scholars (e.g., Henshaw, 2011; Nakata, 2015; Quinn, 2014; Rahimi & Vahid Dastjerdi, 2012) found that immediate CF was equally effective as a delayed and post-delayed CF or attested an advantage to the delayed CF in the retention of the target items. Overall, the findings of this investigation are in line with a growing number of empirical investigations (e.g., Farmani et al., 2017; Fu & Nassaji, 2016; Shabani & Safari, 2016; Shintani & Aubrey, 2016; Siyyari, 2005). Some scholars (e.g., Oxford, 1990; Rivers, 1981) claim that once a mental link is created as a result of applying a certain language learning strategy, the presented items would stick to language learners' minds and become significantly memorable. In such a case one can claim that retention is automatically guaranteed.

6. Conclusion

This study began with the assumption that incorporating various types of CF timing could enhance EFL learners' learning and retention of lexical items and morpho-syntax markers in a SCMC context. The instruction lasted about ninety days, and the researchers employed diverse types of CF timings in a computerized mode. The results indicated that any between-group differences in the participants' post-test scores might be attributable to the instructional treatments. Post-hoc test indicated that both at the time of post-test and delayed posttests the immediate CF group scored significantly higher and the differences were associated with large effect sizes. This suggests that applying immediate CF for EFL learners was pedagogically effective regarding the learning and retention of the target structures. Izadpanah et al.,

(2023) assume that both types of oral and written CF have similar effects on the retention of target structures. It is hoped that the findings contribute to CF research by examining an underexplored variable, CF timing, and in an increasingly relevant context, namely, online language learning.

To conclude, the current investigation revealed that short-term CF and delayed CF were not as effective as immediate CF on a summery retelling task, requiring L2 learners to be accurate while their primary attention was on meaning (like the oral production test). Because gains on these types of tasks have been viewed as a better indicator of L2 acquisition (Doughty, 2011; Ellis, 2016), this result can be taken to mean that short-term CF and delayed CF might not be a good alternative to immediate CF in computerized environment. However, given the practical importance of delayed CF for practitioners, due to the existence of contexts where the provision of immediate CF is not feasible because of limited human resources, it is highly recommended that future research investigates the factors that increase the effectiveness of short-term and delayed corrective feedback. For instance, it might be possible to achieve such effectiveness when L2 learners are supplied with production opportunities after the CF stage or when a more salient feedback strategy is selected (e.g., explicit correction, metalinguistic feedback). Ellis (2016) states that investigation into the role of such factors in moderating the effectiveness of delayed feedback would eventually help the scholars determine the conditions under which short-term or delayed feedback could be an alternative to immediate CF. Additionally, it is recommended that the same types or other types of CF be also targeted to find out the association between variables under investigation and also to figure out if there is any superiority of one or more types of CF over the others for all learners using different feedback timings.

References

Ammar, A., & Spada, N. (2007). One size fits all? Recasts, prompts, and L2 learning. *Studies in Second Language Acquisition*, *28*, 543-574.

Amiri, D., & Ashrafi, S. (2019). Effect of feedback timing and willingness to communicate on the acquisition of simple past form. *Modern Research in English Language Studies*, *6*(2),27-52.

Arroyo, D. C., & Yilmaz, Y. (2018). An open for replication study: The role of feedback timing in synchronous computer-mediated communication. *Language Learning*, *68*(4), 942-972.

Bitchener, J., & Storch, N. (2016). *Written corrective feedback for L2 development*.Multilingual Matters.

Brown, H. D. (2015). *Principles of language learning and teaching* (6th Ed.). McGraw-Hill.

Brown, D. (2016). The type and linguistic foci of oral corrective feedback in the L2 classroom: A meta-analysis. *Language Teaching Research, 20*(4), 436-458.

Butler, A.C., Karpicke, J.D., & Roediger, H.L. (2007). The effect of type and timing of feedback   on learning from multiple-choice tests. *Journal of Experimental Psychology: Applied*, *13*, 273-281.

Carpenter, S. K., & Vul, E. (2011). Delaying feedback by three seconds benefits retention of face name pairs: The role of active anticipatory processing. *Memory & Cognition*, *39*, 1211-1221.

Chaudron, C. (1998). A descriptive model of discourse in the corrective treatment of learners' errors. *Language Learning*, *27*, 29-46.

DeKeyser, R. (1998). *Beyond focus on form: Cognitive perspectives on learning and practicing second language grammar*. Cambridge: Cambridge University Press.

DeKeyser, R. (2007). *Practice in a second language.* New York: Cambridge University Press.

Ellis, R. (1997). A framework for investigating oral and written corrective feedback. *Studies in Second Language Acquisition, 32,* 335-349.

Ellis, N. C. (1997). Vocabulary acquisition, word structure, collocation, word-class, and meaning. In Schmitt, N., & M. McCarthy (Ed.), *Vocabulary: Description, acquisition and pedagogy* (pp. 122–139). Cambridge: Cambridge University Press.

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

Ellis, R. (2006). Researching the effects of form-focused instruction on L2 acquisition. *AILA Review* 19, 18-41.

Ellis, R., S. Li & Y. Zhu (2016). The effects of pre-task explicit instruction on the performance of a focused task. *System* 80, 38-47.

Evans, N.W., Hartshorn, K. J., McCollum, R. M., & Wolfersberger, M., (2010). Contextualizing corrective feedback in second language writing pedagogy. *Language Teaching Research*, *14*, 445-463.

Farmani, R., Akbari, O., & Ghanizadeh, A. (2017). The impact of immediate and delayed error correction on Iranian EFL learners' motivation. *European Journal of Foreign LanguageTeaching, 2(3),* 76-88.

Farrokhi, F., Zohrabi, M., & Chehr Azad, M. H. (2018). Corrective feedback and Iranian EFL learners' spoken complexity and accuracy. Teaching English Language, *12*(2), 117-143.

Fu, M., & Li, S. (2020). The effects of immediate and delayed corrective feedback on L2 development. *Study Second Language Acquisition Journal*. 44, 2–34.

Larsen-Freeman, D. (2003). Teaching language from grammar to grammaring. Heinle/Thomson.

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

Harmer, J. (2007). *The practice of English language teaching*. Pearson Longman.

Hedge, T. (2000). *Teaching and learning in the language classroom*. Oxford University Press.

Hendrickson, J. (1978). Error correction in foreign language teaching: Recent theory, research, and practice. *The Modern Language Journal,* 62(8), 387–398.

Henderson, C. (2018). The effect of feedback timing on L2 Spanish vocabulary acquisition insynchronous computer-mediated communication. *The Modern Language Journal*, *93*,139-328.

Henshaw, F. (2011). Effects of feedback timing in SLA: A computer assisted study on the Spanish subjunctive. Implicit and explicit language learning: *Conditions, processes, & knowledge in SLA* (pp. 85-100). Washington, DC: Georgetown University Press.

Izadpanah, J., Sadeghi, F., & Akbarpour, L., (2023). The effect of explicit corrective feedback onEFL learners' retention of grammar. *Journal of Studies in Learning and Teaching English.* 12(1), Ser. 23, 99-122.

Karpickle, J. D., & Roediger, H.L. (2007). Expanding retrieval practice promotes short-term retention, but equally spaced retrieval enhances long-term retention. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *33*, 704-719.

Kess, A. D., (1992). Retention of vocabulary over time: Investigating the role of mnemonicassociation. *System*, 8*(2)*, 21-35.

Lanuzzi, S. (2015). *First Friends*. Oxford: Oxford University Press.

Laufer, B. (2007). Possible change in attitude toward vocabulary acquisition research. *International review of Applied Linguistics in Language Teaching*, *24*, 73-7.

La Berde, D, Auclair, L., & Seiroff, E. (2000). Preparatory attention: Experiment and theory. Consciousness and Cognition, 9, 396–434.

Lee, I. (2013). Revisiting teachers' feedback in EFL writing from sociocultural perspectives. *TESOL Quarterly, 48*, 201-213.

Lee, I. (2013b). Research into practice: Written corrective feedback.

Language Teaching, *46*, 108-119.

Li, S. (2014). The effectiveness of corrective feedback in SLA: A meta-analysis. *Language Learning*, *60*(2), 309-365.

Li, S. (2016). The effectiveness of corrective feedback in SLA: A meta-analysis. *Language Learning*, *60*, 309-365.

Li, S. (2017a). The effects of cognitive aptitudes on the process and product of L2 interaction: A synthetic review. In L. Gurzynski-Weiss (ed.), Expanding individual difference research in the interaction approach. John Benjamins, 41-70.

Li, S. (2018). Corrective feedback. In J. Liontas et al. (eds.), *The TESOL encyclopedia of English language teaching*. Blackwell, 1-10.

Li, S. (2019). Corrective feedback. In J. Lintas et al. (eds.), *The TESOL encyclopedia of English language teaching*. Blackwell, 1–10.

Li, S., Zhu, Y., & Ellis, R. (2016). The effects of the timing of corrective feedback on the acquisition of a new linguistic structure. *The Modern Language Journal*, *100*, 276-295

Lightbown, P. M. (2008). Transfer appropriate processing as a model for classroom second language acquisition. In Han, Z.H. (Ed.), *Understanding second language process* (pp. 27-44). Clevedon: Multilingual Matters.

Long, M. (1997). The role of the linguistic environment in second language acquisition. In W.Ritchie & T. Bhatia (Eds.), *Handbook of second language acquisition* (pp. 413-468). Academic Press.

Long, M. H. (2007). The role of the linguistic environment in second language acquisition. In W. Ritchie & T. Bhatia (Eds.), *Handbook of second language acquisition* (pp. 413-468). Academic Press.

Lyster, R. (2004). Differential effects of prompts and recasts in form-focused instruction. Studies in Second Language Acquisition, 26, 399–432. 27, 79–103.

Lyster, R. & L. Ranta (2007). Corrective feedback and learner uptake: Negotiation of form in communicative classrooms. *Studies in Second Language Acquisition* 19.1, 37-66.
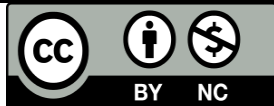
Lyster, R. & K. Saito (2010). Oral feedback in classroom SLA: A meta-analysis. Studies in Second Language Acquisition *32*, 265-302.

Lyster, R., Saito, K., & Sato, M. (2013). Oral corrective feedback in second language classrooms. *Language Teaching*, *46*(1), 1-40.

Mackey, A., & Philp, J. (1998). Conversational interaction and second language development: Recasts, responses, and red herrings? *Modern Language Journal,* 82, 338-356.

McDaniel, M. A., Robinson-Riegler, B., & Einstein, G. O. (1998). Prospective remembering: Perceptually driven or conceptually driven processes? *Memory & Cognition, 26*(1), 121–134.

Mackey, A. (2015.). *Conversational interaction in SLA: A collection of empirical studies*. Oxford University Press.

Mackey, A., & Goo, J. (2007). Interaction research in SLA: A meta-analysis and research synthesis. In A. Mackey (Ed.), *Conversational interaction in SLA: A collection of empirical studies* (pp. 408-452). Oxford University Press.

Manchón, R. M., & Cerezo, L. (2018). Writing as language learning. *The TESOL encyclopedia of English language teaching* (pp. 1-6). Wiley.

McCarthy, M., McCarten, J., & Sandiford, H. (2006). *Touchstone*. Cambridge University Press.

Morris, C. D., Bransford, J. D., & Franks, J. J. (1977). Levels of processing versus transfer appropriate processing. *Journal of Verbal Learning & Verbal Behavior, 16*(5), 519–533.

Nader, K., & Einarsson, E. O. (2010). Memory reconsolidation: An update. *Annals of the New York Academy of Sciences, 1191*(1), 27-41.

Nakata, T. (2014). Effects of feedback timing on second language vocabulary learning: Does delaying feedback increase learning? *Language Teaching Research, 9*(4), pp. 416-434.

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

Nakata, T. (2015). Optimizing second language vocabulary learning from flashcards (Unpublished doctoral dissertation). Victoria University of Wellington, New Zealand.

Quinn, P. (2014). *Delayed versus immediate corrective feedback on orally produced passive errors in English*. (Unpublished doctoral dissertation). University of Toronto, Toronto, Canada.

Quinn, P. (2021). Corrective feedback timing and second language grammatical development: research, theory, and practice. *The Cambridge handbook of corrective feedback in second language learning and teaching*. eds. H. Nassaji and E. Kartchava (Cambridge: Cambridge University Press), 322–340.

Rahimi, A., & Dastjerdi, H. (2012). Impact of immediate and delayed error correction on EFL learners‟ oral production: CAF. *Mediterranean Journal of Social Sciences, 3*(1), 45-54.

R´ev´esz, A., & Han, Z. (2006). Task content familiarity, task type and efficacy of recasts. *Language Awareness*, *15*, 160–179.

Richards, J. C., & Platt, J. (1992). *Approaches and methods in language teaching*. Oxford: Oxford University Press.

Rivers, M. (1981). Teaching foreign language skills. Chicago University Press.

Rivers, S., & Tayama, S. (2015). *English Time*. Oxford University Press.

Scrivener, J. (2005). *Learning teaching:* A guidebook for English language teachers (2nd ed.). Macmillan.

Shabani, K. & Safari, F. (2016a). The effect of immediate and delayed error correction on accuracy of EFL learners' oral production. *Journal of Studies in Education*, *6*(3), 93-112.

Sheen, Y. & Ellis, R. (2011). Reexamining the role of recasts in second language acquisition. *Studies in Second Language Acquisition*, *28*(4), 575–600.

Siyyari, M. (2005). A comparative study of the effect of implicit and delayed, explicit focus on form on Iranian EFL learners' accuracy of oral production. Unpublished MA thesis. Iran University of Science and Technology, Tehran, Iran.

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

VanPatten, B. (1990). Attending to form and content in the input. *Studies in Second Language Acquisition*, *12*, 287-301.

Willis, D. & Willis, J. (2007). *Doing task-based teaching*. Oxford: Oxford University Press.

Yilmaz, Y. (2018). Task effects on focus on form in synchronous computer-mediated communication. *The Modern Language Journal*, *95*, 115-132.

Yilmaz, Y., & Granena, G. (2010). The effects of task type in synchronous computer-mediated communication. *Recall Journal*, *22*, 20-38.

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

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

G. Shooshtari et al.

Yilmaz, Y. (2016). The linguistic environment, interaction and negative feedback. *Brill Research Perspectives in Multilingualism and Second Language Acquisition*, *1*, 45-86.