



Associations Between L2 Proficiency and Efficacy of Corrective Feedback Types

Yusuke Sato

Nagano University, Japan

This study examined the relative effectiveness of recasts, explicit correction, and metalinguistic prompts on developing English past progressive forms among second language (L2) learners, and whether the effectiveness of each feedback type was differently associated with L2 proficiency. A total of 197 Japanese learners of English as a foreign language were divided into recast, explicit correction, metalinguistic prompt, task-only, and test control groups. The first three groups performed tasks and received corrective feedback (CF). The task-only group performed the same tasks as the CF groups but did not receive CF. The test control group did not perform the tasks and solely took pretests and posttests. The results showed that explicit types of feedback (explicit correction and metalinguistic prompts) were more effective than implicit feedback (recasts). Further, simple regression analyses showed that L2 proficiency was significantly associated with gains in scores of the recast and metalinguistic prompt groups but not the explicit correction group. These results suggest that L2 proficiency differentially influences the effectiveness of different CF types.

Keywords: corrective feedback, second language proficiency, past progressive

Introduction

The effects of corrective feedback (CF) on second language (L2) learning have attracted the interests of both researchers and teachers, and studies have shown the efficacy of CF on L2 development (Li, 2010; Lyster & Saito, 2010; Mackey & Goo, 2007). Although various types of CF are employed in classrooms (Lyster & Ranta, 1997), several researchers have argued that learners' internal differences are related to the effectiveness of CF in L2 development (e.g., Ammar & Spada, 2006; Li, 2014; Mackey & Philp, 1998; Révész, 2012; Sagarra, 2007; Sheen, 2007). For instance, differences in L2 proficiency may affect the effectiveness of CF (Ammar & Spada, 2006; Li, 2014). Ammar and Spada (2006) found that the effectiveness of different types of CF (recasts or prompts) depended on learners' L2 proficiency. The results of their study suggested that high-proficiency learners benefited from both recasts and prompts, while low-proficiency learners benefited only from prompts. Despite the evidence that learners' proficiency may differently affect the effectiveness of different CF types on L2 learning, few studies have investigated this relationship. Clarifying the relationship would be useful not only for researchers who are interested in moderating roles of L2 proficiency on the benefits of CF but also for teachers who are willing to choose appropriate CF according to learners' proficiency levels. This study investigates the relative effectiveness of recasts, explicit correction, and metalinguistic prompts on L2 learning of the English past progressive forms and explores whether learners' general proficiency differentially moderates the effectiveness of the three types of CF on L2 development. Lyster and Saito (2010) argued



that a comparison of the three types of CF could clarify the processes needed for L2 development. Nonetheless, the associations between CF and L2 proficiency has not been investigated earlier, using three different types of CF in one study. Therefore, this study would further contribute to CF research in this respect.

Theoretical Issues

Many researchers have drawn on Long's (1996) interaction hypothesis to discuss the benefits of CF. In an interaction, the learner produces outputs, which may induce CF from the interlocutor. When the learner notices the difference between their output and the input (correct target forms) provided by the interlocutor's CF, it may lead to L2 development. Regarding this noticing effect (Schmidt, 1995, 2001), Sheen (2007) stated that if correct target forms are provided with metalinguistic information, awareness level rises from mere noticing to understanding. Moreover, Skill Acquisition Theory (DeKeyser, 2015, 2017) has been drawn on to discuss the relations between CF and the development of L2 procedural knowledge (e.g., Lyster & Saito, 2013). DeKeyser (2015) distinguished declarative and procedural knowledge and claimed that declarative knowledge turns into procedural knowledge when declarative knowledge is applied in actual language use. Lyster and Sato (2013) insisted that "opportunities for contextualized practice are needed to complement input-driven approaches designed to trigger noticing and awareness of target language features" (p. 71) in L2 classroom contexts. Therefore, if the learners are provided with practice opportunities to reformulate their incorrect utterances, they may proceduralize more of their declarative knowledge.

CF Types

Research has shown that teachers use various types of CF such as recasts, explicit correction, and metalinguistic prompts during class (Lyster & Ranta, 1997). Explicit correction refers to "the explicit provision of correct form" (Lyster & Ranta, 1997, p. 46). It may or may not be accompanied by metalinguistic explanation (Sheen & Ellis, 2011). Further, Lyster and Ranta (1997) operationalized recasts as the provision of "the teacher's reformulation of all or part of a student's utterance, minus the error" (p. 46) and metalinguistic prompts as the provision of "either comments, information, or questions related to the well-formedness of the student's utterance, without explicitly providing the correct form" (p. 47). Lyster and Saito (2010) argued that comparing the three feedback types is feasible and reveals the processes associated with L2 development, as the feedback types vary in several aspects. Explicit correction and metalinguistic prompts are classified as explicit CF while recasts are classified as implicit CF, depending on whether the corrective force of CF is covered. In addition, these feedback types can be divided into input-providing and output-prompting feedback (Sheen & Ellis, 2011). Input-providing feedback provides the learner with the correct form, while output-prompting feedback requires learners to correct their errors independently. In this dichotomy, the former includes recasts and explicit correction and the latter comprises metalinguistic prompts. Most studies have compared the effectiveness of recasts and prompts. In L2 classroom contexts, prompts are generally more effective than recasts (Ellis et al., 2006; Lyster, 2004). However, in foreign language (FL) classroom contexts, the results of research are mixed (Iizuka & Nakatsukasa, 2020; Li & Iwashita, 2019; Yang & Lyster, 2010), which necessitates further investigations into FL contexts (Li & Iwashita, 2019). In terms of comparisons between the two types of input-providing feedback, explicit correction is more beneficial than recast (Li, 2014; Yilmaz, 2012). However, as the studies were laboratory studies, the relative effectiveness of the two feedback types in classroom contexts remains unclear. To date, Sato (2021) has compared the relative efficacy of recasts, explicit correction, and metalinguistic prompts in a single study. The results suggest that the effectiveness of CF may differ depending on the type of knowledge (declarative or procedural) measured. However, more studies are necessary to clarify the relationships between CF and the development of divergent types of knowledge. In addition, although the effectiveness of CF may be moderated by

learners' internal differences, such as L2 proficiency, little attention has been paid to the influence of this individual difference variable on the effects of CF (Pawlak, 2021). In the following sections, the studies that examined the relations between the benefits of CF and L2 proficiency are reviewed.

CF and L2 Proficiency

Ammar and Spada (2006) examined whether L2 proficiency moderates the effectiveness of various types of CF (recasts and prompts) in the learning of third-person possessive determiners in English. The participants were classified into prompt, recast, and control groups. The results showed that low-proficiency learners whose pre-test scores were 50% or less benefited more from prompts, while high-proficiency learners whose pre-test scores were over 50% benefited equally from recasts and prompts. This finding suggests that the effectiveness of CF may be influenced by L2 proficiency. Moreover, whether L2 proficiency affects the effectiveness of CF depends on CF types. In their study, the effectiveness of recasts was influenced by L2 proficiency but that of prompts was not. In this regard, the authors argued that the implicitness of recast and the processing load that recasts required prevented low-proficiency learners from noticing the correct form. That is, they "seem to need techniques that explicitly signal the presence of errors and limit the processing effort required to notice the gap between their interlanguage system and the target language norm," while high-proficiency learners "with a greater knowledge of the target language forms might not need to be coached into noticing the correct form" (Ammar & Spada, 2006, p. 563).

Li (2014) examined the interactions between CF types, the nature of the linguistic target, and L2 proficiency. Learners of Chinese were divided into low- and high-proficiency groups based on their general L2 proficiency scores. Each group was further classified into recast, explicit correction (the provision of the correct target form and metalinguistic explanation), and control subgroups. The results showed that, on the one hand, explicit correction was more beneficial than recasts for low-proficiency learners. The differences were greater in the declarative knowledge measure (untimed grammaticality judgment test; UGJT) than the procedural knowledge measure (elicited imitation test; EIT). On the other hand, explicit correction and recasts were equally beneficial for high-proficiency learners. Thus, the study also suggests that whether L2 proficiency influences the effectiveness of CF may depend on CF type; compared to the efficacy of explicit correction, that of recasts was more influenced by L2 proficiency. However, if the target structure was salient, recasts were also effective for low-proficiency learners. Thus, the effects of CF on L2 knowledge may vary depending on the types of knowledge measured, learners' proficiency levels, and linguistic structures.

The results of the studies mentioned above establish that compared to implicit types, explicit types of feedback may be less influenced by L2 proficiency. This may be because explicit feedback usually includes metalinguistic information and may neutralize individual differences in L2 proficiency. One of the limitations of these studies is that they treated L2 proficiency dichotomously (high or low proficiency) rather than continuously. Thus, the whole picture of the relations between CF and L2 proficiency may be unclear. Moreover, dividing the participants into two groups inevitably led to small sample sizes, thus limiting the generalizability of the results. Furthermore, merely comparing the two types of CF did not reveal which dimension of CF was more related to L2 proficiency, explicitness/implicitness, or opportunities of input/output.

The Present Study

Sheen and Ellis (2011) claimed that explicit and output-prompting feedback is the most beneficial one for L2 learners in classrooms. However, in FL classroom contexts, the results are not necessarily in line with it (Iizuka & Nakatsukasa, 2020; Li & Iwashita, 2019), which necessitates further investigations in that context. Moreover, several studies have demonstrated that the effects of CF on L2 learning may differ according to the types of knowledge measured (Sato, 2021) or the learner's proficiency levels

(Ammar & Spada, 2006; Li, 2014). Their results confirm that compared to the effectiveness of explicit feedback, that of implicit feedback may be more influenced by L2 proficiency. However, as previous studies compared only two CF types (recasts vs. prompts, or recasts vs. explicit correction), the dimension of CF (explicitness/implicitness or opportunities of input/output) that is more associated with L2 proficiency remains unclear. Moreover, as the participants were divided into two groups (high and low proficiency groups), the sample sizes were relatively small, limiting generalizability. This study aims to examine the effects of the three different types of CF (recasts, explicit correction with metalinguistic information, and metalinguistic prompts) on the L2 development of English past progressive forms in an FL classroom setting. Comparing the three feedback types is feasible and makes a theoretical contribution to second language acquisition research (Lyster & Saito, 2010). Furthermore, this study explores whether L2 proficiency differently influences the effectiveness of the different CF types on L2 learning. The research questions are as follows:

1. What is the relative effectiveness of recasts, explicit correction, and metalinguistic prompts on the L2 development of English past progressive forms for learners in an FL classroom context?
2. Is L2 proficiency differentially related to the effectiveness of the three CF?

Regarding Research Question 1, as Sheen and Ellis (2011) insisted that explicit and output-prompting feedback is the most effective in classroom settings, metalinguistic prompts would be more beneficial for L2 learning than recasts and explicit correction. Moreover, as research has indicated explicit correction to be more effective than recasts (Li, 2014; Yilmaz, 2012), the former may be superior to the latter in this study. Although both recasts and explicit correction provide the correct target form, recasts may not be as effective as explicit correction, as recasts are implicit and the corrective intent of the feedback may be less noticeable, especially for low proficiency learners. Regarding Research Question 2, compared to the effectiveness of explicit feedback, that of implicit feedback may be more influenced by L2 proficiency (Ammar & Spada, 2006; Li, 2014). Hence, the effects of recasts may be more affected by proficiency than the other two explicit feedback types. As explicit correction provides correct target forms and metalinguistic information, it may neutralize individual differences in previous L2 knowledge or proficiency. Conversely, although metalinguistic prompts also provide metalinguistic information, the feedback requires learners to reformulate their ill-formed utterances independently, thereby posing a heavy processing load for low-proficiency learners. Thus, L2 proficiency may moderate the benefit of the feedback.

Methods

Participants

The participants were first-year students at a university in Japan. Their ages ranged from 19 to 20. They had learned English as a foreign language for over six years. None of them had lived in an English-speaking country for more than two weeks. The English classes they attended at the university were form-oriented. Their English proficiency was assumed to be around beginner to lower intermediate levels. Although they studied the past progressive form when they were junior or senior high school students, the pretest scores of these participants indicated that they did not possess much declarative or procedural knowledge of its structure. Five English classes were chosen for this study, each of which was assigned to the recast group ($n = 43$), the explicit correction group ($n = 38$), the metalinguistic prompt group ($n = 38$), the task-only group ($n = 36$), and the control group ($n = 42$). Analyses of variance (ANOVAs) of the pretest scores with the group as a between-subject variable found no significant group differences for the UGJT or EIT, $F(4, 192) = 1.06, p = .38, \eta^2_p = .02, F(4, 93.9) = 2.40, p = .06, \eta^2_p = .04$, respectively.

Target Structure

Various linguistic structures have been used in CF studies. In this study, the past progressive form was chosen as the target structure, mainly because the effects of CF on past progressive forms have remained relatively unexplored in CF research compared to others, such as past tense forms. Révész (2012) selected this structure for two reasons: First, past progressive forms are “realized via a free morpheme (*was/were*) and a syllabic bound morpheme (*-ing*); thus, it is physically salient” (p. 104); when the structure is salient, learners may notice the corrective intent of implicit types of feedback such as recasts (Li, 2014). Second, the structure “denotes grammatical tense and aspect, meaning that it has some communicative value” (Révész, 2012, p. 104). Thus, past progressive forms would be appropriate for studies that include a recast group and low-proficiency learners who tend to miss the corrective intent of feedback. Further, Shirahata (2015) reported that the structure is easier to acquire for Japanese learners than others such as third-person *-s*. Considering the relatively short treatment duration (one-hour) secured for this study, the past progressive form was deemed appropriate.

Treatment

The treatment task designed to elicit the targeted form was a picture description task adapted from Révész (2012). The task scenario was as follows. The participants were near a place where a crime had occurred between 7 a.m. and 12 p.m. A police officer (the teacher) was looking for the suspect and asked the participants to give their alibis. The participants were asked to tell the police officer where they had been when the crime occurred and describe the activities of other people present there. The police had some information about those people and checked whether the participants lied.

First, the participants saw two pictures that showed some people doing different things on their computer screen. The two pictures varied in terms of the activities of those people, the time, and the place. Relatively difficult verbs related to the people’s activities were listed in the pictures. Initially, the participants prepared to describe the pictures by themselves. Subsequently, they practiced it in pairs. Afterward, they were asked to memorize the activities of the people, the time, and the place in the pictures. Finally, the picture was removed from the screen, and the teacher asked the students to remember it and recount the activities of the people, the time, and the place in the pictures to the whole class. Only the CF groups received CF during the task. The task-only group performed the same task but did not receive any CF, and the control group did not perform the task. This task was performed twice in two weeks with different pictures (30 minutes × 2).

CF

The recast, the explicit correction, and the metalinguistic prompt groups consistently received recasts, explicit correction, and metalinguistic prompts, respectively, from the teacher during the tasks.

The recast group received relatively implicit recasts that were full and declarative without any emphasis on the erroneous part (Zhao & Ellis, 2022); for example, see the following:

Student: A woman painting a picture at seven in the park.

Teacher: Oh, a woman was painting a picture at seven in the park.

The explicit correction group received metalinguistic explanations and the correct form; for example:

Student: A man use a computer at seven in the park.

Teacher: No. You need to use the past progressive; *was/were* plus *-ing* form of verbs. So, a man was using a computer at seven in the park.

The metalinguistic prompt group was pushed to reformulate their output after the metalinguistic explanation. Thus, unlike the explicit correction group, the correct form was not provided by the teacher; for example:

Student: Two boys run at seven in the park.

Teacher: You need to use the past progressive; was/were plus -ing form of verbs. So...

Student: Two boys were running at seven in the park.

The metalinguistic explanation part was provided in the learners' first language (Japanese) for the explicit correction and metalinguistic prompt groups.

Learning Outcome Measures

Students' development of the declarative and procedural knowledge of the past progressive form was measured by the UGJT and the EIT, respectively. Zhao and Ellis (2022) stated that these tests are distinct measures of the two types of knowledge, which has also been validated by others (Bowles, 2011; Ellis, 2005, 2009; Zhang, 2014). The UGJT and the EIT were conducted at the pretest, the immediate posttest, and the delayed posttest.

In the UGJT, the participants were required to judge whether the given sentences were grammatically correct. If the sentences were incorrect, they were asked to revise the erroneous parts. There was no time constraint. The UGJT consisted of 16 target sentences and 8 distractor sentences. Half of the target sentences were grammatical, and the other half were ungrammatical. Although the sentences were provided with PowerPoint slides, the participants were asked to write their answers and corrections on the answer sheet. If the learner judged grammatical sentences correctly, they were given one point. However, for ungrammatical sentences, a point was awarded only if the learner judged them as ungrammatical and corrected the target structure. Moreover, if the ungrammatical sentence was changed to the present progressive form, 0.5 point was awarded. The reliabilities of the test were computed using Cronbach's alpha. They were $\alpha = .67$, $\alpha = .68$, and $\alpha = .72$, for the pretest, the immediate posttest, and the delayed posttest, respectively.

In the EIT, the participants were asked to listen to prerecorded sentences (e.g., "I was playing volleyball at two yesterday"), judge whether the sentences were true for themselves, and write their answers (*true, not true, not sure*) on their answer sheets. This process required the participants to focus on meaning (Erlam, 2006). Subsequently, they were required to reproduce the sentences. There was a six-second pause between each sentence. This time pressure, in particular, encouraged the use of procedural knowledge (Ellis, 2009). The EIT was composed of 12 target sentences and 8 distractor sentences. Half of these were grammatical, and the other half were not grammatical. The participants' reproduction of the sentences was audio-recorded. For the grammatical sentences, if the learner correctly repeated the target structure, one point was awarded. Regarding the ungrammatical sentences, if the learner corrected the ungrammatical part of the target structure, one point was awarded. If the ungrammatical part was changed to the present progressive form, 0.5 point was awarded. The participants' judgements about their personal life were not rated. The reliabilities of the test were $\alpha = .65$, $\alpha = .69$, and $\alpha = .78$, for the pretest, the immediate posttest, and the delayed posttest, respectively.

L2 Proficiency

As used by Li (2014), the term "proficiency" refers to general proficiency measured by standardized tests. In this study, the test of English for international communication (TOEIC) Bridge test was employed to measure the learners' L2 proficiency. The TOEIC Bridge test consists of reading and listening sections. Each section comprises fifty questions, and the learners are required to answer them in one hour. The possible test scores ranged from 20 to 180. As the test is intended for beginners and lower

intermediate learners (IIBC, 2019), it was appropriate for the participants in this study. They were obliged to take the test by their university and received their test scores about a week before the commencement of the study. ANOVA did not detect a significant difference among the groups' proficiency scores, $F(4, 192) = 1.63, p = .17, \eta^2_p = .03$

Procedure

All the tasks and tests were conducted in regular classes, and all the sessions were taught by one teacher (the author of the study). In the first week, all participants completed the pretests consisting of the EIT and the UGJT in that order. In the second week, the experimental groups (recast, explicit correction, metalinguistic prompt and task-only) performed the first task in class. In the third week, the experimental groups performed the second task and completed the immediate posttests, while the control group completed only the posttests. In the sixth week, all the participants completed the delayed posttests. Only the feedback groups (recast, explicit correction, and metalinguistic prompt) received CF during the tasks. All the tasks and the tests were conducted using computers.

Analysis

To examine the first research question on the relative effectiveness of recasts, explicit correction, and metalinguistic prompts on the L2 development of the past progressive, one-way ANOVAs were conducted on gains in scores (immediate or delayed posttest scores - pretest scores) of the UGJT and the EIT with group as a between-subject variable. Gain scores were used to factor out the effects of pretest scores on the results of statistical analyses as much as possible. When the ANOVA detected significant differences, Bonferroni-adjusted post-hoc analyses were carried out to locate where the differences lay. To compare the magnitudes of the effects of CF, effect sizes (Cohen's *d*) were also reported. Effect sizes larger than 0.2, 0.5 and 0.8 were considered small, medium, and large, respectively. Subsequently, to investigate the second research question on whether L2 proficiency differentially influences the effectiveness of the three CF types, simple linear regression analyses were conducted with the UGJT or the EIT gain scores as a dependent variable and L2 proficiency scores as an independent variable.

Results

The total numbers of CF episodes the recasts, explicit correction, and metalinguistic prompt groups received during the tasks were 11, 8, and 9, respectively, which indicated that the groups received similar numbers of CF during the tasks and the effects of the three feedback types were comparable.

UGJT

Table 1 presents the descriptive statistics of the pretest, immediate posttest, delayed posttests, Gain 1 scores (immediate post test scores - pretest scores), and Gain 2 scores (delayed posttest scores - pretest scores) of the UGJT. Generally, all the treatment groups showed some improvement from the pretest to the two posttests. To examine the relative effectiveness of CF on the UGJT, one-way ANOVAs were conducted on Gain 1 and Gain 2 scores with group as a between-subject variable. The results detected a significant group difference among the groups, $F(4, 192) = 6.33, p < .001, \eta^2_p = .11$ for Gain 1 scores, but did not for Gain 2 scores, $F(4, 192) = 2.18, p = .07, \eta^2_p = .04$. Subsequently, post-hoc comparisons determined which pairs of the groups were different for Gain 1 scores. Table 2 presents the summary of the post-hoc analyses. The results revealed that the explicit correction and metalinguistic prompt groups' scores were significantly different from those of the test control group with large effect sizes ($d = 1.09$ and 0.94 , respectively). Moreover, the explicit correction group performed better than the task-only group

with a large effect size ($d = 0.83$). Thus, the effects of the explicit types of CF that included metalinguistic information on the L2 development of the declarative knowledge of the past progressive form measured by the UGJT were advantageous over those of the implicit type.

In terms of the influence of L2 proficiency on the effectiveness of the CF types on the UGJT scores, simple regression analyses were run with Gain 1 or 2 scores as a dependent variable and L2 proficiency scores as an independent variable for the CF groups. Table 3 summarizes the results of the regression analyses, which indicated that L2 proficiency explained 10% of the variance in Gain 2 scores for the recast group. No other significant associations were observed between L2 proficiency scores and gain scores for the explicit correction and metalinguistic prompt groups. This result suggests that L2 proficiency may influence only the efficacy of the implicit types of CF on the L2 development of declarative knowledge, and that these effects may appear later.

TABLE 1

Descriptive Statistics of the Pretest, Posttest, and Gain Scores of the UGJT

	Pretest			Immediate posttest		Delayed posttest		Gain 1		Gain 2	
	<i>n</i>	<i>M</i> (%)	<i>SD</i>	<i>M</i> (%)	<i>SD</i>	<i>M</i> (%)	<i>SD</i>	<i>M</i> (%)	<i>SD</i>	<i>M</i> (%)	<i>SD</i>
Recast	43	52.7	11.8	58.9	13.1	59.7	15.9	6.3	14.3	7.0	13.3
Explicit correction	38	53.4	12.5	65.0	14.5	61.6	15.3	11.6	13.7	8.2	14.6
Prompts	38	53.4	12.0	61.9	15.1	61.4	15.9	8.6	11.1	8.1	13.7
Task-only	36	54.7	11.9	57.3	12.7	58.2	13.7	2.6	6.6	3.5	10.3
Test control	42	49.6	10.9	50.4	11.5	51.3	13.9	0.9	3.9	1.8	9.3

Note. UGJT, untimed grammaticality judgment test.

TABLE 2

Group Comparisons of Gain Scores of the UGJT

Group comparisons	Gain 1		Gain 2	
	<i>d</i>	<i>p</i>	<i>d</i>	<i>p</i>
recast vs. explicit correction	0.38	.24	0.09	1.00
recast vs. prompt	0.18	1.00	0.08	1.00
recast vs. task-only	0.32	1.00	0.29	1.00
recast vs. test control	0.51	.21	0.62	.55
explicit correction vs. prompt	0.24	1.00	0.01	1.00
explicit correction vs. task-only	0.83	.00*	0.37	1.00
explicit correction vs. test control	1.09	.00*	0.53	.22
prompt vs. task-only	0.65	.31	0.38	1.00
prompt vs. test control	0.94	.03*	0.42	.25
task-only vs. test control	0.32	1.00	0.17	1.00

Note. UGJT, untimed grammaticality judgment test.

* $p < .05$

TABLE 3

Summary of Simple Regression Analyses of the UGJT

Group	Gain 1			Gain 2		
	β	R^2	<i>p</i>	β	R^2	<i>p</i>
Recasts	.01	.00	.93	.68	.10	.03*
Explicit correction	-.13	.02	.43	.04	.00	.81
Prompts	.27	.07	.10	.23	.05	.17

Note. UGJT, untimed grammaticality judgment test.

* $p < .05$

EIT

Table 4 presents the descriptive statistics of the pretest, immediate posttest, delayed posttest, and gain scores of the EIT. In general, all the treatment groups demonstrated an improvement in their posttest scores over their pretest scores.

Regarding group differences, the results of one-way ANOVAs established that there were significant group differences in Gain 1 scores, $F(4, 192) = 8.94, p < .001, \eta^2_p = .16$ and Gain 2 scores, $F(4, 192) = 3.89, p = .005, \eta^2_p = .08$. A summary of the results of post-hoc analyses is provided in Table 5. The results demonstrated that the metalinguistic prompt group outperformed all the other groups with medium or large effect sizes (d s ranged from 0.55 to 1.30) in terms of Gain 1 scores. With regard to Gain 2 scores, the explicit correction and metalinguistic prompt groups showed more beneficial effects than the test control group with large ($d = 0.96$) and medium ($d = 0.71$) effect sizes, respectively. These results indicated that the explicit types of CF were more effective than the implicit type in developing procedural knowledge of the past progressive form. Further, output opportunities to modify the erroneous utterances provided by the metalinguistic prompts might be more important than input or correct target forms provided by recasts or explicit correction at least in the short-term.

In terms of the moderating role of L2 proficiency on the efficacy of the CF types, simple regression analyses were computed with the EIT gain scores as the dependent variable and the L2 proficiency scores as the independent variable for the CF groups. Table 6 summarizes the results of the regression analyses. They reveal that L2 proficiency was significantly associated with the benefits of metalinguistic prompts and explained 19% of the variance in the Gain 1 scores for the group, whereas L2 proficiency was significantly related to the efficacy of recasts and accounted for 11% of the variance in the Gain 2 scores for the group. These results suggest that L2 proficiency may differentially influence the effects of different CF types on developing the procedural knowledge of the past progressive forms at varied times.

TABLE 4

Descriptive Statistics of the Pretest, Posttest, and Gain Scores of the EIT

	<i>n</i>	Pretest		Immediate posttest		Delayed posttest		Gain 1		Gain 2	
		<i>M</i> (%)	<i>SD</i>	<i>M</i> (%)	<i>SD</i>	<i>M</i> (%)	<i>SD</i>	<i>M</i> (%)	<i>SD</i>	<i>M</i> (%)	<i>SD</i>
Recast	43	12.2	13.6	17.2	15.2	17.2	14.7	4.9	13.1	5.0	11.4
Explicit correction	38	7.0	7.9	12.0	10.0	15.4	13.2	4.9	11.2	8.3	10.8
Prompts	38	10.5	9.8	22.6	17.7	17.5	14.3	12.1	12.8	7.0	12.6
Task-only	36	7.8	7.6	9.5	8.0	10.8	10.6	1.7	5.8	3.0	9.8
Test control	42	7.3	5.8	6.9	7.7	7.5	7.8	-0.4	5.4	0.2	5.6

Note. EIT, elicited imitation test.

TABLE 5

Group Comparisons of Gain Scores of the EIT

Group comparisons	Gain 1		Gain 2	
	<i>d</i>	<i>p</i>	<i>d</i>	<i>p</i>
recast vs. explicit correction	0.00	1.00	0.3	1.00
recast vs. prompt	0.55	.01*	0.17	1.00
recast vs. task-only	0.31	1.00	0.19	1.00
recast vs. test control	0.53	.21	0.54	.39
explicit correction vs. prompt	0.6	.04*	0.11	1.00
explicit correction vs. task-only	0.36	1.00	0.51	.28
explicit correction vs. test control	0.62	.10	0.96	.01*
prompt vs. task-only	1.03	.00*	0.35	.97
prompt vs. test control	1.3	.00*	0.71	.04*
task-only vs. test control	0.38	1.00	0.36	1.00

Note. EIT, elicited imitation test.

* $p < .05$

TABLE 6
Summary of Simple Regression Analyses of the EIT

Group	Gain 1			Gain 2		
	β	R^2	p	β	R^2	p
Recasts	.25	.06	.10	.32	.11	.04*
Explicit correction	.30	.09	.07	.20	.04	.22
Prompts	.43	.19	.01*	.20	.04	.22

Note. EIT, elicited imitation test.

* $p < .05$

Discussion

Relative Effectiveness of the Different Types of CF on Developing of the Past Progressive Form

Before discussing the research questions, the effects of the task itself are examined. Irrespective of the outcome measures, the task-only group showed some improvement in the posttests. However, there were no statistically significant differences detected between this group and the test control group. Hence, the significant differences detected between the CF groups and task-only or test control groups were derived from the additional CF effects.

Research Question 1 addresses the relative effectiveness of recasts, explicit correction, and metalinguistic prompts on the L2 development of past progressive forms. In terms of the UGJT, the results confirmed that the explicit correction group outperformed the task-only and the test control group on the immediate posttest with large effect sizes. Furthermore, the metalinguistic group performed better than the test control group on the immediate posttest with a large effect size. Both explicit correction and metalinguistic prompts are categorized as explicit types of feedback. Thus, for the learners in this study, the explicit types of feedback were more beneficial in developing the declarative knowledge of past progressive forms than the implicit type. These results are in line with those of past studies (Ellis et al., 2006; Li, 2014; Sheen, 2007; Yilmaz, 2012) that reported the superiority of explicit feedback over implicit feedback. For instance, Sheen (2007) identified the preeminence of explicit correction over recasts in the learning of articles. Sheen referred to Schmidt's (1995) noticing hypothesis and argued that explicit correction is conducive to both levels of awareness, that is, noticing and understanding. She stated that, "it is, therefore, perhaps not so surprising that provision of the correct form together with metalinguistic feedback proved more effective than the recasts alone because it led to noticing and understanding of the underlying rule" (p. 318). The results that recasts were less effective than the explicit feedback types may also be explained in terms of redundancy. Ellis (2015) stated that a structure is noticed less if it is "communicatively redundant (i.e., they are not necessary for understanding of an utterance)" (p. 15). Although Révész (2012) affirmed that the past progressive form has some communicative value, it may not have been high for performing the tasks in this study. For instance, the learners in this study were able to convey the information on past events with adverbial phrases without any morphological changes. As the task required the learners to use an adverbial phrase such as "At seven-thirty yesterday," they might not have felt the necessity of using past progressive forms to convey information on past events in pictures to their interlocutors. Thus, the explicit types of feedback might have been more beneficial by directing more of their attention to the targeted forms. Moreover, as the study was conducted in a classroom, the corrective intent of recasts was less noticeable than in laboratory studies. Regarding the comparison between the two explicit feedback types, as the effect size of explicit correction was larger than that of metalinguistic prompts, providing metalinguistic information with correct target forms would be more beneficial than that with output opportunities to modify their non-target like utterances. These results are in line with those of Sato (2021). As the learners in the metalinguistic group were not given correct forms from the teacher, some of them might not have been

able to reach the correct target forms independently even if they were provided with metalinguistic information.

In terms of the EIT as a measure of procedural knowledge, the metalinguistic prompt group made significantly more gains than all the other groups at the immediate posttest. This result supports the hypothesis that explicit and output-prompting feedback is the best CF in classroom contexts (Sheen & Ellis, 2011) and is consistent with the findings of past studies (Lyster, 2004; Sato, 2021; Yang & Lyster, 2010). In terms of this advantageous effect of prompts over the other feedback types, Lyster and Sato (2013) argued that opportunities to retrieve the relevant declarative knowledge to a target structure and modify erroneous utterances using the knowledge are more beneficial “for L2 development than merely noticing target forms during interaction by drawing on Skill Acquisition Theory” (p. 83). The Skill Acquisition Theory postulates that declarative knowledge is proceduralized when the knowledge is applied in communicative activities (DeKeyser, 2015). While recasts and explicit correction might not have afforded such opportunities for proceduralization, metalinguistic prompts could have, thereby leading to significant differences between the groups. However, not only the metalinguistic prompt group but also the explicit correction group that were not provided with opportunities to modify their output made significantly more gains than the control group at the delayed posttest. Thus, it may be possible that when learners understand how to construct past progressive forms or develop the declarative knowledge of the structure, the metalinguistic/declarative knowledge may be partially and gradually proceduralized over time with a little practice. Regarding this aspect, DeKeyser (2015, p. 95) stated that, “proceduralization can be complete after just a few trials/instances.” After the explicit correction group received the feedback, several of the learners might have practiced applying the metalinguistic/declarative knowledge on their own while observing other students interacting with the teacher or performing the immediate posttest, thus leading to a partial proceduralization.

Effects of L2 Proficiency

Research Question 2 asked whether L2 proficiency differentially affects the effectiveness of the three different types of CF in developing the past progressive form. The results indicated that L2 proficiency influenced only the effects of recasts on the delayed posttest in terms of the UGJT. This result supports our prediction and is in line with that of previous studies (Ammar & Spada, 2006; Li, 2014). Ammar and Spada (2006) found that recasts were less effective than prompts only for low-proficiency learners. Li (2014) also found that recasts were less beneficial for L2 learning than metalinguistic correction (i.e., explicit correction in this study) only for the lower-proficiency group. These results established that L2 proficiency influenced the benefits of recasts but not those of explicit types of CF. Therefore, in terms of developing the declarative knowledge of the past progressive form, compared to the effects of explicit types of feedback, those of implicit types of feedback might be more moderated by L2 proficiency. Ammar and Spada (2006) explained the ineffectiveness of recasts from the standpoint of the implicitness of recasts and the cognitive processing load posed in L2 interaction. They argue that low-proficiency learners need more explicit signals, which lessens their cognitive processing load in L2 interactions. As recasts in this study were relatively implicit, low-proficiency learners in the recast group might not have been able to notice the corrective intent of the feedback due to the high cognitive processing load taxed by accomplishing the task in itself. Moreover, as they were not provided with metalinguistic information, their cognitive processing load would not have been reduced. Conversely, as explicit correction and metalinguistic feedback included metalinguistic information, it might have neutralized the effects of the difference in L2 proficiency on L2 development.

In terms of the EIT, L2 proficiency affected the effectiveness of metalinguistic prompts and recasts on the immediate and delayed posttests, respectively. Only the learners in the metalinguistic group were required to modify their erroneous output independently with the help of the metalinguistic information provided. Thus, unlike explicit correction that afforded correct forms, as the metalinguistic group needed to apply the rule of the target structure and reformulate their ill-formed output independently, it may have

taxed the learners with a heavier cognitive processing load than the other feedback types. Although high-proficiency learners might have been able to use the metalinguistic information effectively and promptly in the subsequent reformulation of their output, low-proficiency learners might not have done so, as their cognitive capacity was easily filled up merely in processing/understanding the metalinguistic feedback. In terms of recasts, L2 proficiency affected the delayed posttest scores, which was in line with the results of the UGJT. This result revealed that recasts might have leveled out the effects of L2 proficiency during the immediate posttest. As recasts contained correct target forms, the learners in the group might have been able to store the correct forms as unanalyzed chunks or exemplars in their short-term memory and withdraw them easily, at least during the immediate posttest, thus leveling out the difference in L2 proficiency. However, as time passed, only the high-proficiency learners who noticed the corrective intent of the feedback and integrated the forms into existing knowledge in long-term memory may have developed the knowledge of the target structure. Thus, the effects of L2 proficiency on the feedback might have appeared only during the delayed posttest.

As explicit correction afforded both metalinguistic information and correct target forms, L2 proficiency might not have played a significant role in the development of L2 irrespective of the outcome measures. In other words, all the learners in the explicit correction group may have the potential to benefit from feedback regardless of their proficiency. It may be one of the reasons why explicit correction might have been superior to recasts. Thus, it would be appropriate for teachers to use the feedback in classrooms where learners' proficiency levels are low.

Conclusion

This study explored the relative effectiveness of recasts, explicit correction, and metalinguistic prompts on the development of the past progressive form and examined whether L2 proficiency differentially influenced the effectiveness of the feedback types. The results showed that the explicit types of feedback were more effective than the implicit type regardless of the outcome measures. In particular, explicit correction and metalinguistic prompts were useful to develop declarative and procedural knowledge, respectively. In terms of the influence of L2 proficiency on the efficacy of the feedback types, L2 proficiency played a role in developing the declarative knowledge for the recast group, while it was significantly associated with developing procedural knowledge for the recast and metalinguistic prompt groups. Explicit correction might have been more effective than recasts because explicit correction that afforded both metalinguistic information and correct forms might have neutralized the influence of L2 proficiency on the effectiveness of the feedback.

This study had several limitations. First, the short treatment duration did not allow learners to develop the knowledge of the target structure substantially. Therefore, if a longer treatment time had been secured in this study, more significant CF effects might have been observed. Second, the numbers of each CF provided during the task were not completely the same across the groups. Although it is not realistic to equate the numbers of CF in each group in classroom-based studies such as this one, if the volume of feedback can be controlled properly, the results would be more generalizable. Moreover, though learners' proficiency was measured by the TOEIC Bridge test, it did not include speaking and writing measures. Thus, if another L2 proficiency measure had been used, different results might have been obtained.

As a pedagogical implication, teachers should employ explicit correction in classes, especially when various levels of L2 learners are mixed in one classroom. However, if the learners attain high L2 proficiency, divergent types of feedback would be effective. Thus, it is important to take learners' proficiency into account when choosing and providing appropriate feedback.

The Author

Yusuke Sato is an associate professor in the Dept. of Business and Informatics of Nagano University in Japan.

Department of Business and Informatics
Nagano University
Shimonogo 658-1, Ueda-shi
Nagano-ken, 386-1298, Japan
Email: yusuke-sato@nagano.ac.jp

References

- Ammar, A., & Spada, N. (2006). One size fits all? Recasts, prompts, and L2 learning. *Studies in Second Language Acquisition*, 28, 543-574. <https://doi.org/10.1017/S0272263106060268>
- Bowles, M. A. (2011). Measuring implicit and explicit linguistic knowledge. *Studies in Second Language Acquisition*, 33(2), 247-271. <https://doi.org/10.1017/S0272263110000756>
- DeKeyser, R. (2015). Skill acquisition theory. In B. VanPatten & J. Williams (Eds.), *Theories in second language acquisition: An introduction* (pp. 94-112). Routledge.
- DeKeyser, R. (2017). Knowledge and skill in ISLA. In S. Loewen & M. Sato (Eds.), *The Routledge handbook of instructed second language acquisition* (pp. 15-32). Routledge.
- Ellis, R. (2005). Measuring implicit and explicit knowledge of a second language. *Studies in Second Language Acquisition*, 27, 141-172. <https://doi.org/10.1017/S0272263105050096>
- Ellis, R. (2009). Measuring implicit and explicit knowledge of a second language. In R. Ellis, S. Loewen, C. Elder, R. Erlam, J. Philp, & H. Reinders (Eds.), *Implicit and explicit knowledge in second language learning, testing and teaching* (pp. 31-64). Multilingual Matters.
- Ellis, R. (2015). *Understanding second language acquisition* (2nd ed.). Oxford 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. <https://doi.org/10.1017/S0272263106060141>
- Erlam, R. (2006). Elicited imitation as a measure of L2 implicit knowledge: An empirical validation study. *Applied Linguistics*, 27, 464-491. <https://doi.org/10.1093/applin/aml001>
- IIBC. (2019). *Test Results*. IIBC. <https://www.iibc-global.org/english/toeic/test/bridge.html>
- Iizuka, T., & Nakatsukasa, K. (2020). Implicit and explicit corrective feedback and feedback exposure conditions. *Instructed Second Language Acquisition*, 4(1), 3-48. <https://doi.org/10.1558/isla.38113>
- Li, H., & Iwashita, N. (2019). The role of recasts and negotiated prompts in an FL learning context in China with non-English major university students. *Language Teaching Research*. <https://doi.org/10.1177/1362168819839727BR>
- Li, S. (2010). The effectiveness of corrective feedback in SLA: A meta-analysis. *Language Learning*, 60, 309-365. <https://doi.org/10.1111/j.1467-9922.2010.00561.x>
- Li, S. (2014). The interface between feedback type, L2 proficiency, and the nature of the linguistic target. *Language Teaching Research*, 18(3), 373-396. <https://doi.org/10.1177/1362168813510384>
- Long, M. H. (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). Academic Press.
- Lyster, R. (2004). Different effects of prompts and effects in form-focused instruction. *Studies in Second Language Acquisition*, 26, 399-432. <https://doi.org/10.1017/S0272263104263021>
- Lyster, R., & Ranta, L. (1997). Corrective feedback and learner uptake. *Studies in Second Language Acquisition*, 19, 37-66. <https://doi.org/10.1017/S0272263197001034>
- Lyster, R., & Saito, K. (2010). Oral feedback in classroom SLA: A meta-analysis. *Studies in Second Language Acquisition*, 32, 265-302. <https://doi.org/10.1017/S0272263109990520>

- Lyster, R., & Sato, M. (2013). Skill acquisition theory and the role of practice in L2 development. In M. García Mayo, J. Gutierrez-Mangado, & M. Martínez Adrián (Eds.), *Contemporary approaches to second language acquisition* (pp. 71-92). John Benjamins.
- Mackey, A., & Philp, J. (1998). Conversational interaction and second language development: Recasts, responses, and red herrings? *Modern Language Journal*, 82, 338-356. <https://doi.org/10.1111/j.1540-4781.1998.tb01211.x>
- Mackey, A., & Goo, J. (2007). Interaction research in SLA: A meta-analysis and research synthesis. In A. Mackey (Ed.), *Conversational interaction in second language acquisition: A collection of empirical studies* (pp. 407-452). Oxford University Press.
- Pawlak, M. (2021). Corrective feedback, developmental readiness, and language proficiency. In H. Nassaji & E. Kartchava (Eds.), *The Cambridge handbook of corrective feedback in second language learning and teaching* (pp. 733-753). Cambridge University Press.
- Révész, A. (2012). Working memory and the observed effectiveness of recasts on different L2 outcome measures. *Language Learning*, 62, 93-132. <https://doi.org/10.1111/j.1467-9922.2011.00690.x>
- Sagarra, N. (2007). From CALL to face-to-face interaction: The effect of computer delivered recasts and working memory on L2 development. In A. Mackey (Ed.), *Conversational interaction in second language acquisition: A collection of empirical studies* (pp. 229-248). Oxford University Press.
- Sato, Y. (2021). Effects of recasts, explicit correction and metalinguistic prompts on L2 knowledge of the English past tense. *The Journal of Asia TEFL*, 18(2), 390-405. <http://dx.doi.org/10.18823/asiatefl.2021.18.2.1.390>
- Schmidt, R. W. (1995). Consciousness and foreign language learning: A tutorial on the role of attention and awareness in learning. In R. W. Schmidt (Ed.), *Attention and awareness in foreign language learning* (pp. 1-63). University of Hawai'i, National Foreign Language Resource Center.
- Schmidt, R. W. (2001). Attention. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 3-32). Cambridge University Press.
- Sheen, Y. (2007). The effects of corrective feedback, language aptitude, and learner attitudes on the acquisition of English articles. In A. Mackey (Ed.), *Conversational interaction in second language acquisition: A collection of empirical studies* (pp. 301-322). Oxford University Press.
- Sheen, Y., & Ellis, R. (2011). Corrective feedback in language teaching. In E. Hinkel (Ed.), *Handbook of research in second language teaching and learning* (Vol. 2, pp. 593-610). Routledge.
- Shirahata, T. (2015). *Eigoshidoo ni okeru koukatekina ayamari teisei: Daini gengo syuutoku kenkyuu no chiken kara. [Effective grammatical corrections in English teaching: From the viewpoint of second language acquisition research]*. Taishukan Shoten.
- Yang, Y., & Lyster, R. (2010). Effects of form-focused practice and feedback on Chinese EFL learners' acquisition of regular and irregular past tense forms. *Studies in Second Language Acquisition*, 32(2), 235-263. <https://doi.org/10.1017/S0272263109990519>
- Yılmaz, Y. (2012). The relative effects of explicit correction and recasts on two target structures via two communication modes. *Language Learning*, 62, 1134-1169. <https://doi.org/10.1111/j.1467-9922.2012.00726.x>
- Zhang, R. (2014). Measuring university-level L2 learners' implicit and explicit linguistic knowledge. *Studies in Second Language Acquisition*, 37(3), 1-30. <https://doi.org/10.1017/S0272263114000370>
- Zhao, Y., & Ellis, R. (2022). 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*. 26(3), 361-381. <https://doi.org/10.1177/1362168820903343>

(Received November 22, 2021; Revised May 20, 2022; Accepted June 18, 2022)