



Effects of Recasts, Explicit Correction and Metalinguistic Prompts on L2 Knowledge of the English Past Tense

Yusuke Sato

International University of Health and Welfare, Japan

This study examined the effects of three different types of corrective feedback (CF) on the acquisition of declarative/procedural knowledge of the English past tense in a foreign language (FL) classroom setting. The participants comprised Japanese university students, and five classes were assigned to recasts, explicit correction, metalinguistic prompt, task-only, and test control groups. The first three groups performed tasks designed to elicit past-tense forms and received CF. The task-only group performed the tasks without any CF, while the test control group completed pretests and posttests only. Participants' declarative and procedural knowledge of the past tense was measured at the pretest, immediate posttest, and delayed posttest. The results generally showed that recasts and explicit correction were effective for the development of declarative knowledge of the structures, while metalinguistic prompts were beneficial for the development of procedural knowledge. Thus, metalinguistic prompts would be one of the best types of feedback to provide in an FL classroom setting to help learners develop procedural knowledge of past-tense forms. However, to effectively develop their declarative knowledge of past-tense forms, recasts and explicit correction should also be provided.

Keywords: corrective feedback, past tense, declarative knowledge, procedural knowledge

Introduction

Teachers and researchers specializing in second language acquisition (SLA) have taken great interest in the effectiveness of corrective feedback (CF). Numerous studies have reported the positive effects of CF on second language (L2) learning (Li, 2010; Lyster & Saito, 2010; Mackey & Goo, 2007). However, the effectiveness of different types of CF can vary depending on the learning context. In Mackey and Goo's meta-analysis (2007), which included laboratory studies, recasts had a larger effect on L2 learning than either negotiation or metalinguistic feedback. Conversely, a meta-analysis conducted by Lyster and Saito (2010) found that prompts, including metalinguistic feedback, were more beneficial than recasts in classroom settings. Sheen and Ellis (2011) have argued that "in general, the types of CF that have the greatest impact on L2 development in a classroom context are those that are explicit and output-prompting rather than implicit and input-providing" (p. 607). However, the roles of different CF types in L2 development in that context remain unclear (Goo & Mackey, 2013; Li & Iwashita, 2021; Lyster & Ranta, 2013; Lyster et al., 2013) as the effectiveness of CF differ according to classroom contexts (L2 or FL contexts; Lyster & Mori, 2006; Sheen, 2004).

Furthermore, several researchers insist on the importance of distinguishing two types of knowledge in CF studies (e.g., Ellis et al., 2006; Iizuka & Nakatsukasa, 2020; Rouhi & Javan-Amani, 2011; Zhao &



Ellis, 2020): declarative and procedural knowledge. However, several past studies have not distinguished these two types of knowledge, which may partly contribute to their inconsistent results. As Sheen and Ellis (2011) claimed, is it the case that in a classroom context, explicit and output-prompting feedback provides the most benefit to learners attempting to acquire both declarative and procedural knowledge?

This study investigates the effects of three different types of CF (recast, explicit correction, and metalinguistic prompt) on the development of declarative and procedural knowledge of English past-tense forms (regular and irregular past tense) in an FL classroom context. Only a few studies have investigated the effects of three different types of CF on L2 development in a single study (e.g., Loewen & Nabei, 2007). Comparing the three types of feedback provides insight into what type of evidence or processing is conducive to L2 development of declarative or procedural knowledge (Lyster & Saito, 2010).

Literature Review

Types of CF

CF refers to “responses to the errors that learners commit in their production or comprehension of a second language” (Li, 2018, p. 34). Among the various types of CF, Lyster and Saito (2010) stated comparing recasts, explicit correction, and prompts is feasible and clarifies the types of evidence and processing involved in L2 development. Although prompts comprise various CF types such as clarification requests, repetitions, elicitations, or metalinguistic feedback, in this study, only the effects of metalinguistic feedback are investigated. Recasts are “the teacher’s reformulation of all or part of a student’s utterance, minus the error” (Lyster & Ranta, 1997, p. 46). Explicit correction refers to “the explicit provision of correct form” (Lyster & Ranta, 1997, p. 46) with/without metalinguistic information. Metalinguistic feedback refers to “comments, information, or questions related to the well-formedness of the student’s utterance, without explicitly providing the correct form” (Lyster & Ranta, 1997, p. 47). These types of CF were broadly classified into implicit and explicit CF depending on whether its corrective force is covered. According to Lyster and Saito (2010), while recasts are implicit types of CF, explicit correction and metalinguistic feedback are explicit types. In addition, types of CF can be divided into input-providing and output-prompting. Input-providing feedback provides “learners with the correct target form,” while output-prompting feedback pushes “learners to self-correct their own errors” (Ellis & Shintani, 2013, p. 265). In this dichotomy, recasts and explicit correction constitute input-providing feedback, while metalinguistic feedback constitutes output-prompting feedback. Regarding types of evidence, Lyster and Saito (2010) stated that recasts basically provide positive evidence but possibly negative evidence; explicit correction affords both positive and negative evidence; and metalinguistic feedback/prompts provide only negative evidence. In summary, the three types of CF are different in that they are implicit or explicit, input-providing or output-prompting, and provide positive or negative evidence. Thus, comparisons of the three types of CF will lead to a better understanding of the role of CF in L2 development.

Theoretical Issues

The noticing hypothesis (Schmidt, 1995, 2001) is often referred to in discussing the effectiveness of recasts and explicit correction. According to Schmidt (1995, 2001), for learners to develop L2 knowledge, noticing is necessary. If learners are afforded the correct target forms through recasts or explicit correction, their attention may be directed toward the target forms, and they may notice the difference between their interlanguage and the target language, thus leading to their learning of the forms. Although this learning might be related solely to awareness at the level of noticing for recasts, explicit correction provided with metalinguistic information may bring the learning further to awareness at the level of understanding (Sheen, 2007).

Conversely, with respect to the effectiveness of metalinguistic prompts, skill acquisition theory is frequently referenced. Skill acquisition theory (DeKeyser, 2015) posits that L2 knowledge develops from initial declarative knowledge to procedural knowledge. Procedural knowledge is then automatized through repeated practice. Metalinguistic prompts allow learners to modify their output or practice using the target form instead of giving them the correct target form. Thus, compared to recasts and explicit correction, metalinguistic prompts may afford more opportunities for learners to turn declarative knowledge into procedural knowledge (Lyster & Saito, 2013).

CF in Classroom Contexts

Numerous past studies have compared recasts and prompts. Most of the studies conducted in L2 classroom contexts have consistently reported that prompts are superior to recasts (Ammar & Spada, 2006; Ellis et al., 2006; Lyster, 2004). However, in the studies by Ammar and Spada (2006) and Lyster (2004), the prompts included various types of feedback (including both explicit and implicit feedback). Therefore, the role that the degree of explicitness plays in L2 development was unclear. Conversely, Ellis et al. (2006) compared implicit (recasts) and explicit (metalinguistic prompt) feedback and found that the explicit metalinguistic prompt group performed better than the implicit recast group. However, comparing only the two feedback types did not clarify which was more associated with L2 development, explicitness (metalinguistic information), or opportunities for learners to reformulate their output. One solution to overcome this limitation is the addition of an explicit correction group that is given metalinguistic information and input (correct target form).

Compared to the relatively consistent results in L2 classroom contexts, the results in FL classroom contexts are mixed (Iizuka & Nakatsukasa, 2020; Li & Iwashita, 2021; Yang & Lyster, 2010). The study by Yang and Lyster (2010) showed that prompts were generally more effective than recasts, which was in line with L2 classroom studies. However, Li and Iwashita (2021) stated that the classroom context of Yang and Lyster's study was not a common FL context as the learners comprised English major students and were immersed with extensive L2 input. Li and Iwashita examined the effects of recasts and negotiated prompts that offered opportunities for output only (clarification requests, repetitions, and elicitations) in a form-oriented classroom for non-English major students and found that recasts were more effective (large effects) than prompts (medium effects). Moreover, although Iizuka and Nakatsukasa (2020) compared the effects of recasts and metalinguistic prompts on L2 development for learners of Japanese, unlike the results of past L2 classroom studies, they found that recasts and metalinguistic prompts were equally effective. These mixed results cast doubt on Sheen and Ellis's (2011) statement that explicit and output-prompting feedback is more effective than implicit and input-providing feedback in classroom settings, necessitating further studies in FL classroom contexts.

In terms of the contrast between the two types of input-providing feedback, recasts, and explicit correction, Sheen (2007) showed that explicit correction was more effective than recasts in L2 classroom contexts. Thus, the results suggest that explicit feedback may be more effective than implicit feedback in the context. The results of Lyster and Saito's (2010) meta-analysis also showed that the effect size of explicit correction is larger than that of recast, although no statistically significant difference was detected.

Regarding the comparison between explicit correction and prompts, surprisingly, few studies have explored this. Thus, comparing the effects of the two feedback types would extend the research.

Although researchers have investigated which type of CF is the most effective in classroom settings, the results suggest that CF effects differ according to classroom contexts. Li and Iwashita (2021) stated that there are fewer CF studies conducted in a form-oriented FL classroom context than in a meaning-oriented L2 classroom context, necessitating more CF studies in such a context.

Effects of CF on Declarative or Procedural Knowledge

In CF studies, several researchers have separately measured L2 development of declarative/explicit and procedural/implicit knowledge (Ellis et al., 2006; Iizuka & Nakatsukasa, 2020; Loewen & Nabei, 2007; Zhao & Ellis, 2020). Declarative knowledge refers to “knowledge of facts (semantic memory) and events (episodic memory), usually consciously accessible and often verbalizable, but not necessarily” (DeKeyser, 2017, p. 17), while procedural knowledge refers to “knowledge that can only be performed, such as how to swim, do mental arithmetic, or speak fluently” (DeKeyser, 2017, p. 17). Several studies have shown that CF is effective for both implicit/procedural knowledge and explicit/declarative knowledge (Ellis et al., 2006; Iizuka & Nakatsukasa, 2020; Zhao & Ellis, 2020). However, the study by Loewen and Nabei (2007) showed that CF was only effective for implicit/procedural knowledge. These results suggest that different CF types may be involved in developing the two types of knowledge differently. Thus, measuring the two types of knowledge separately would shed further light on the roles of different CF types.

The Current Study

In line with Sheen and Ellis (2011), the literature suggests that explicit and output-prompting feedback is the most effective type of feedback in L2 classroom contexts. However, the results of studies in FL classroom contexts are mixed. Furthermore, the effectiveness of different types of CF can vary depending on the type of knowledge (declarative or procedural) measured and linguistic structures involved. This study provided learners with three different types of CF in an FL classroom context and measured the L2 development of both declarative and procedural knowledge. The three types of CF were recasts, explicit correction, and one type of prompt (metalinguistic feedback). The target structure was the English past tense. To date, only in a few cases has a single study compared the three types of CF. Such a comparison would reveal the evidence types and processing required for L2 development (Lyster & Saito, 2010) more clearly than studies that compare only two CF types. Accordingly, the following research questions were formulated:

Research Question 1: To what extent do recasts, explicit correction, and metalinguistic prompts help students in an FL classroom context effectively develop L2 declarative knowledge of English regular/irregular past-tense forms, and which has a larger effect?

Research Question 2: To what extent do recasts, explicit correction, and metalinguistic prompts help students in an FL classroom context effectively develop L2 procedural knowledge of English regular/irregular past-tense forms, and which has a larger effect?

It was generally hypothesized that all the CF types used in this study would help students develop L2 declarative and procedural knowledge (Ellis et al., 2006; Iizuka & Nakatsukasa, 2020; Zhao & Ellis, 2020). However, as this study was conducted in a classroom context, metalinguistic prompts may be more effective than recasts or explicit corrections because they constituted explicit and output-prompting feedback (Sheen & Ellis, 2011).

Methods

Participants

This study was conducted in an FL classroom context in Japan. The participants comprised 201 first-year Japanese university students who had studied English as an FL over the course of six years. These participants were chosen mainly because they still had difficulty in using the past-tense forms correctly

despite possessing some declarative knowledge of such forms. In other words, they had not fully proceduralized the declarative knowledge at the time of the study. Hence, the effects of different types of feedback would be observed not only on declarative knowledge but also on procedural knowledge. Moreover, as they learned English as an FL and took primarily form-oriented classes, they seemed to be accuracy-oriented. Thus, investigating the effects of different feedback types on L2 development for these learners would clarify the effects of the learning context on the benefits of the feedback types (Li & Iwashita, 2021). Five intact classes were assigned to the recast, explicit correction, metalinguistic prompt, task-only, and test control groups. The first four groups performed tasks designed to elicit past-tense forms during class. However, only the recast, explicit correction, and metalinguistic prompt groups were given CF during the tasks. The test control group simply took a pretest and two posttests. The number of participants in each group ranged from 37 to 45. Data from participants who missed the pretest or posttests were excluded from the analysis. There were no significant group differences in the pretest scores of the untimed grammaticality judgment test (UGJT) as a measure of declarative knowledge, $F(4, 85.1) = 2.09, p = .09$ (regular past tense), and $F(4, 76.9) = 1.17, p = .33$ (irregular past tense). In addition, no significant group differences were found among the groups for the pretest scores of the elicited imitation test (EIT) as a measure of procedural knowledge, $F(4, 91.9) = 1.54, p = .20$ (regular past-tense forms), and $F(4, 92.2) = 1.10, p = .36$ (irregular past-tense forms), indicating that the groups were comparable.

Target Structures

The target structure used in this study is the English past tense; although it is introduced at a relatively early stage of L2 learning (Yang & Lyster, 2010), even intermediate and advanced learners find it difficult to master this form (Ellis et al., 2006). Research has shown that Japanese university students often make mistakes in the past-tense form (Shirahata, 2015) even after learning English for several years. The rule for regular past-tense forms is simple: *-ed* is added to the end of the base form of a regular verb. However, there is no clear rule for the irregular past-tense form. Thus, regular past-tense forms involve rules-based learning, whereas irregular past-tense forms involve exemplar-based learning (Yang & Lyster, 2010). In this regard, Ullman (2015) insists that regular and irregular past-tense forms are learned in different memory systems: regular and irregular past-tense forms involve procedural and declarative memory systems, respectively. Moreover, the two forms have differing degrees of saliency (Yang & Lyster, 2010). That is, irregular past-tense forms are more salient than regular past-tense forms (Yang & Lyster, 2010). As the saliency of structures may be related to noticing, regular and irregular past-tense forms may differ in the efficacy of implicit types of feedback such as recasts. The use of these two different types of structures allows researchers to investigate whether the effects of CF are mediated by linguistic structures (Yang & Lyster, 2010). Consequently, investigating the effects of CF on the development of declarative and procedural knowledge of past-tense forms provides useful information for both teachers and researchers.

Instructional Material

The tasks involved picture-cued narrative tasks, which are frequently used to examine the effects of CF on the learning of past-tense forms (Ellis, 2007; Ellis et al., 2006; Yang & Lyster, 2010). This study chose the story of *Momotaro*, one of Japan's most famous folk tales, as all participants were familiar with the story. The learners were given pictures of *Momotaro* and told to describe scenes from the story. Some difficult words were listed on the pictures for participants to use during the task. In addition, several sentences were written next to the pictures to reduce the effort required to describe everything they contained. The learners were asked to prepare to describe the scenes, first by themselves and then in pairs. Subsequently, the teacher asked some of the students to tell the whole class about a scene in one sentence. When they made mistakes involving the past-tense form, the teacher offered CF to students in the three

CF groups but not to those in the task-only group. The story of *Momotaro* was split in half; the first and second halves of the story were used during the first and second weeks, respectively. The three feedback and task-only groups spent approximately an hour, in total, completing the two tasks over two weeks. Although the treatment time secured for this study was relatively short (an hour), past studies (Ellis et al., 2006; Li & Iwashita, 2021) showed that learners were able to develop their knowledge of past-tense forms even with such limited instructional time. Although it may not have been sufficient for the learners to fully develop their implicit knowledge, they would have been able to develop their explicit and declarative knowledge of the structures and at least partially proceduralize the knowledge. Thus, the treatment time was considered acceptable for the purpose of the study.

Feedback Conditions

The same L2 teacher (the researcher) provided CF across the various feedback groups. The recast group received relatively implicit recasts (Zhao & Ellis, 2020) when they made mistakes involving the past-tense form. The recasts provided were declarative and full types. For example,

Learner: The woman goes to the river.
Teacher: Oh, the woman went to the river.

The explicit correction group received the correct target form after metalinguistic information. For example,

Learner: He gives one dumpling to the pheasant.
Teacher: No. You need to use the past tense. So, he gave one dumpling to the pheasant.

The metalinguistic prompt group was encouraged to produce the modified output following metalinguistic information. For example,

Learner: She picks it up.
Teacher: You need to use the past tense. So...
Learner: She picked it up.

If a learner could not produce the modified output on their own, even after the metalinguistic cue, the teacher asked another student to produce the output on their behalf.

Tests

The UGJT and EIT were used to measure the participants' declarative and procedural knowledge. These tests were used by several researchers (e.g., Zhao & Ellis, 2020) to measure the two types of knowledge. The tests were administered using a computer in a computer room.

During the UGJT, participants were given grammatical and ungrammatical sentences and asked whether each sentence was grammatically correct, with no time constraint (e.g., "Miki played badminton yesterday"). If a sentence was incorrect, they were asked to correct the error and wrote their answers on their answer sheets. The UGJT consisted of 16 sentences targeting past-tense forms and eight distractor sentences. Half of the sentences (12 sentences) were grammatical, while the other half (12 sentences) were ungrammatical. Of the target sentences, nine targeted regular past-tense forms, and seven targeted irregular past-tense forms. Students received one point for recognizing grammatical sentences correctly. They also received one point for judging ungrammatical sentences correctly and correcting the errors. In terms of reliability, the Cronbach's alphas were .68, .72, and .77 for the pretest, immediate posttest, and delayed posttest, respectively.

During the EIT, the participants listened to a sentence (e.g., “Last week, I took a bus to come to school.”) and were asked whether that sentence accurately described their personal lives to focus their attention on meaning rather than forms. They wrote their responses (true, false, and not sure) on answer sheets. They were then asked to repeat the sentence, and their responses were audio-recorded. A time limit (6 seconds between each sentence) was established to elicit the participants’ procedural knowledge. The EIT consisted of 12 target sentences and 6 distractor sentences. Half of the sentences (nine sentences) were grammatical and the other half (nine sentences) were ungrammatical. Of the target sentences, six each elicited regular and irregular past-tense forms. The participants received one point for correctly repeating the past-tense forms in grammatical items; they also received one point for correcting ungrammatical forms using past-tense forms in ungrammatical items. Errors regarding non-target structures were not considered. In terms of reliability, the Cronbach’s alphas were .68, .70, and .67 for the pretest, immediate posttest, and delayed posttest, respectively.

Three versions of the UGJT and EIT were developed for the pretest and the two posttests. The test items in the tests were identical, but the items were presented in a different order in each test. Two practice items were included in each test.

Procedure

The study was conducted during the participants’ regular English classes. In Week 1, all groups were briefed about the study and took the UGJT and EIT pretests. All the pretests and posttests were conducted in computer rooms with an instructor. In the UGJT, the participants were provided with Microsoft PowerPoint slides, which included grammatical and ungrammatical sentences. They were asked to judge the grammaticality of the sentences and correct the errors on their answer sheets if a sentence was incorrect, without time constraint. In the EIT, the participants were asked to listen to prerecorded sentences. Subsequently, they were asked to answer whether the sentences were true for their personal life on their answer sheets and repeat the sentences within six seconds. Their repetition was audio-recorded. In Weeks 2 and 3, the three feedback groups (recasts, explicit correction, metalinguistic prompts) and the task-only group performed the tasks (about 30 minutes per week). These tasks elicited the use of past-tense forms from the learners. If they produced erroneous forms, the feedback groups were given CF during the tasks. After the tasks, the feedback groups and task-only group completed the immediate posttest of the UGJT and EIT. The control group also took these tests in Week 3. In Week 6, all groups completed the delayed posttest of the UGJT and EIT.

Analysis

Data from students who scored more than 90% on their pretests were not included in the analysis to measure learning. All analyses were computed separately for regular and irregular past-tense forms. A mixed-design repeated measure analysis of variance (ANOVA) was computed, with group as the between-group variable and scores on the pretests and immediate and delayed posttests as the within-group variable to examine whether the differences in group scores were statistically significant over time. Subsequently, one-way ANOVAs were conducted on gain scores - immediate or delayed posttest scores minus pretest scores - with group as the between-group variable to examine if there were significant differences among the groups. Although pretest scores did not detect significant group differences for both declarative and procedural knowledge tests, gain scores were used for the group comparisons to eliminate the differences in pretest scores. Post hoc analyses were conducted to detect where the differences lay when an ANOVA detected a significant difference. Effect sizes in the form of Cohen’s *d* were also reported in post hoc comparisons to show the magnitude of the difference. The effect sizes of 0.2, 0.5, and 0.8 were recognized as small, medium, or large, respectively.

Results

Number of Corrective Feedback Episodes

Table 1 presents the number of CF episodes that the learners received during the tasks. The three CF groups received a fairly similar number of CFs for their incorrect utterances.

TABLE 1
Number of Corrective Feedback Episodes During the Tasks

Groups	Regular past tense	Irregular past tense
Recasts	5	5
Explicit correction	4	6
Prompts	4	5

Regular Past-Tense Forms

UGJT

Table 2 shows the means (*M*) and standard deviations (*SD*) of the UGJT pretest, posttests, and gain scores for regular past-tense forms. The means of the pretest scores ranged from 68.2% to 75.9%, indicating that all groups possessed a certain amount of declarative knowledge of regular past-tense forms before the treatment. The results of the repeated measures ANOVA revealed a significant main effect of time, $F(2, 344) = 53.36, p < .001$, group, $F(4, 172) = 4.11, p = .003$, and the significant interaction between time and group, $F(8, 344) = 5.10, p < .001$, indicating that the groups performed differently from each other over time.

Of primary concern in this study is whether there were group differences in gain scores. One-way ANOVAs conducted on Gain 1 scores (immediate posttest scores minus pretest scores) and Gain 2 scores (delayed posttest scores minus pretest scores) revealed significant group differences, $F(4, 85.2) = 8.30, p < .001$, and $F(4, 84.8) = 4.86, p = .001$, respectively. The results of post hoc comparisons are shown in Table 3. The recast and explicit correction groups outperformed the task-only group with the same medium effect size ($d = 0.71$) and the test control group with large effect sizes ($d = 0.99$ and $d = 1.09$, respectively) at the immediate posttest. The metalinguistic prompt group performed better than the test control group with a medium effect size ($d = 0.69$) but not better than the task-only group at the immediate posttest. Conversely, none of the CF groups were significantly different from the task-only group at the delayed posttest, and only the recast group performed better than the test control group with a large effect size ($d = 1.09$). Comparisons of effect sizes indicated that recast and explicit correction (input-providing feedback) were more effective than metalinguistic prompts (output-prompting feedback) in developing declarative knowledge of regular past-tense forms. Moreover, the effects of recasts were sustained during the delayed test.

TABLE 2

Means and Standard Deviations of the Pretest, Posttests, and Gain Scores of the Untimed Grammaticality Judgment Test (UGJT) for Regular Past-Tense Forms

	<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	37	68.2	15.3	81.4	12.6	82.0	11.5	13.2	16.1	13.8	12.9
Explicit correction	35	75.9	10.6	87.0	9.9	83.5	14.9	11.1	11.4	7.6	12.9
Prompts	34	70.6	18.5	78.1	17.2	76.5	17.5	7.5	11.5	5.9	12.6
Task-only	33	73.4	12.7	77.4	10.9	80.5	14.2	4.0	8.3	7.1	8.7
Test control	38	69.0	17.4	70.2	16.2	70.8	14.0	1.2	6.2	1.8	8.8

TABLE 3

Group Comparisons of UGJT Gain Scores (Regular Past Tense)

Group comparisons	Immediate posttest		Delayed posttest	
	<i>d</i>	<i>p</i>	<i>d</i>	<i>p</i>
recast vs. explicit correction	0.15	.97	0.48	.26
recast vs. prompt	0.40	.43	0.62	.08
recast vs. task-only	0.71	.03	0.60	.09
recast vs. test control	0.99	.00	1.09	.00
explicit correction vs. prompt	0.31	.69	0.14	.98
explicit correction vs. task-only	0.71	.04	0.05	1.00
explicit correction vs. test control	1.09	.00	0.54	.17
prompt vs. task-only	0.34	.62	0.11	.99
prompt vs. test control	0.69	.05	0.38	.51
task-only vs. test control	0.40	.48	0.61	.09

EIT

Table 4 shows the means and standard deviations of the EIT scores for regular past-tense forms. The means of the EIT scores ranged from 17.5% to 28.0%, suggesting that the participants had not fully proceduralized the declarative knowledge before the treatment. The results of the repeated measures ANOVA revealed a significant main effect of time, $F(2, 370) = 22.52, p < .001$, and group, $F(4, 185) = 3.75, p = .006$. In addition, there was a significant interaction between time and group, $F(8, 370) = 4.30, p < .001$, indicating that the groups' performance varied over time.

One-way ANOVAs conducted on Gain 1 and Gain 2 scores revealed significant group differences, $F(4, 91.2) = 7.43, p < .001$ and $F(4, 92.0) = 6.46, p < .001$, respectively. A summary of post hoc comparisons is presented in Table 5. The recast and metalinguistic prompt groups outperformed the test control group with medium effect sizes ($d = 0.77$ and $d = 0.58$, respectively) at the immediate posttest. At the delayed posttest, the metalinguistic prompt group outperformed all the other treatment groups and the test control group with medium ($d = 0.67$ – 0.71) and large effect sizes ($d = 0.94$), respectively. Thus, although recasts were effective in the short-term, metalinguistic prompts were most beneficial in developing procedural knowledge of regular past-tense forms over time.

TABLE 4

Means and Standard Deviations of the Pretest, Posttests, and Gain Scores of the Elicited Imitation Test (EIT) for Regular Past-Tense Forms

	<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	39	28.0	22.0	42.7	23.8	31.2	21.7	14.7	18.9	3.2	19.8
Explicit correction	39	24.8	24.4	29.1	27.5	29.9	22.4	4.3	19.4	5.1	16.7
Prompts	38	24.1	21.8	35.1	24.7	39.9	21.1	11.0	17.4	15.8	15.0
Task-only	36	23.6	23.4	32.4	24.2	27.3	19.2	8.8	17.1	3.7	19.2
Test control	38	17.5	16.9	19.3	15.8	18.9	15.6	1.8	14.4	1.3	15.7

TABLE 5

Group Comparisons of EIT Gain Scores (Regular Past Tense)

Group comparisons	Immediate posttest		Delayed posttest	
	<i>d</i>	<i>p</i>	<i>d</i>	<i>p</i>
recast vs. explicit correction	0.55	.12	0.11	.99
recast vs. prompt	0.21	.89	0.71	.02
recast vs. task-only	0.33	.61	0.03	1.00
recast vs. test control	0.77	.01	0.11	.99
explicit correction vs. prompt	0.36	.51	0.67	.03
explicit correction vs. task-only	0.25	.82	0.08	.99
explicit correction vs. test control	0.15	.97	0.24	.84
prompt vs. task-only	0.12	.98	0.70	.03
prompt vs. test control	0.58	.01	0.94	.00
task-only vs. test control	0.45	.32	0.14	.98

Irregular Past-Tense Forms

UGJT

The means and standard deviations of the UGJT scores for irregular past-tense forms are presented in Table 6. The means of the pretest scores ranged from 68.0% to 71.9%, indicating that all groups possessed a certain amount of declarative knowledge of irregular past-tense forms before the treatment. The results of the repeated measures ANOVA revealed a significant main effect of time, $F(2, 316) = 36.70$, $p < .001$, and group, $F(4, 158) = 5.0$, $p < .001$. Moreover, there was a significant interaction between time and group, $F(8, 316) = 2.83$, $p = .005$, indicating that the groups performed differently over time.

One-way ANOVAs conducted on Gain 1 and Gain 2 scores revealed significant group differences, $F(4, 77.2) = 8.47$, $p < .001$, and $F(4, 77.9) = 5.09$, $p = .001$, respectively. The results of post hoc analyses are shown in Table 7. The performance of the recast and explicit correction groups was superior to that of the test control group with medium ($d = 0.73$) and large ($d = 1.08$) effect sizes, respectively, at the immediate posttest. In addition, the performance of the explicit correction group was better than that of the task-only group with a medium effect size ($d = 0.75$). At the delayed posttest, no significant difference was found between groups, suggesting that input-providing feedback would work better than output-prompting feedback for developing declarative knowledge of irregular past-tense forms, but the effectiveness of the feedback might be short-lived.

TABLE 6

Means and Standard Deviations of the Pretest, Posttests, and Gain Scores of the UGJT for Irregular Past-Tense Forms

	<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	32	71.9	14.3	82.6	11.3	80.8	13.4	10.7	15.8	8.9	14.9
Explicit correction	33	71.4	11.8	84.8	11.8	81.0	12.7	13.4	12.8	9.5	13.2
Prompts	30	68.6	16.9	76.2	19.6	77.1	16.2	7.6	17.1	8.6	15.7
Task-only	30	74.3	13.7	79.0	12.3	83.8	11.7	4.8	10.2	9.5	11.5
Test control	38	68.0	11.7	69.2	12.7	70.7	13.7	1.1	10.2	2.6	10.9

TABLE 7

Group Comparisons of UGJT Gain Scores (Irregular Past Tense)

Group comparisons	Immediate posttest		Delayed posttest	
	<i>d</i>	<i>p</i>	<i>d</i>	<i>p</i>
recast vs. explicit correction	0.19	.94	0.04	1.00
recast vs. prompt	0.19	.94	0.02	1.00
recast vs. task-only	0.44	.40	0.05	1.00
recast vs. test control	0.73	.04	0.49	.29
explicit correction vs. prompt	0.39	.56	0.07	1.00
explicit correction vs. task-only	0.75	.03	0.00	1.00
explicit correction vs. test control	1.08	.00	0.57	.14
prompt vs. task-only	0.20	.93	0.07	1.00
prompt vs. test control	0.48	.36	0.45	.41
task-only vs. test control	0.36	.59	0.62	.10

EIT

The means and standard deviations of the EIT scores for irregular past-tense forms are presented in Table 8. The mean pretest scores ranged from 11.4% to 18.1%, suggesting that the declarative knowledge of all groups was only partially proceduralized before the treatment. The results of the repeated measures ANOVA revealed a significant main effect of time, $F(2, 370) = 38.51, p < .001$, and group, $F(4, 185) = 3.18, p = .02$, and a significant interaction between time and group, $F(8, 370) = 4.21, p < .001$. Thus, the groups' performance varied over time.

The one-way ANOVAs conducted on the Gain 1 and Gain 2 scores revealed significant group differences, $F(4, 92.1) = 4.82, p = .001$, and $F(4, 91.8) = 4.66, p = .002$, respectively. The results of the post hoc analyses are summarized in Table 9. It shows that only the recast group outperformed both the task-only and test control groups at the immediate posttest, and the effect sizes were medium ($d = 0.77$) and large ($d = 0.86$), respectively. Conversely, at the delayed posttest, only the metalinguistic prompt group was significantly different from the task-only and test control groups with medium ($d = 0.62$) and large ($d = 0.88$) effect sizes, respectively. This suggests that the effects of output-prompting feedback on the development of procedural knowledge of irregular past tense may be more persistent than those of input-providing feedback.

TABLE 8

Means and Standard Deviations of the Pretest, Posttests, and Gain Scores of the EIT for Irregular Past-Tense Forms

	<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	39	17.1	17.7	35.9	27.4	28.6	24.8	18.8	22.4	11.5	19.2
Explicit correction	39	11.5	18.8	21.4	20.9	20.9	21.2	9.8	17.8	9.4	16.6
Prompts	38	12.7	16.6	26.3	24.1	31.1	29.6	13.6	20.1	19.3	23.7
Task-only	36	18.1	20.1	22.7	20.4	24.1	22.3	4.6	13.6	6.0	15.5
Test control	38	11.4	16.5	14.0	18.4	13.2	18.6	2.6	14.8	1.8	13.3

TABLE 9

Group Comparisons of the EIT Gain Scores (Irregular Past Tense)

Group comparisons	Immediate posttest		Delayed posttest	
	<i>d</i>	<i>p</i>	<i>d</i>	<i>p</i>
recast vs. explicit correction	0.45	.30	0.12	.98
recast vs. prompt	0.25	.82	0.32	.51
recast vs. task-only	0.77	.01	0.31	.64
recast vs. test control	0.86	.00	0.59	.08
explicit correction vs. prompt	0.19	.91	0.45	.27
explicit correction vs. task-only	0.34	.61	0.21	.89
explicit correction vs. test control	0.45	.31	0.51	.18
prompt vs. task-only	0.52	.17	0.62	.04
prompt vs. test control	0.63	.06	0.88	.00
task-only vs. test control	0.14	.97	0.30	.71

Discussion

Before discussing the effects of CF on L2 development, task effects without CF are examined. In relation to the task-only group, in line with previous studies (Ammar & Spada, 2006; Lyster, 2004; Yang & Lyster, 2010), this study showed that performing tasks without CF may have partly contributed to the development of the declarative and procedural knowledge of the target structures as the posttest score of the group was higher than the pretest score. However, its effects were limited as statistically significant differences were not detected between the task-only and control groups. Thus, significant differences found between the CF groups and task-only or test control groups derived from additional CF effects.

Effects of CF on Declarative Knowledge

Research question 1 addresses the impact of recasts, explicit correction, and metalinguistic prompts on the acquisition of L2 declarative knowledge of English regular/irregular past-tense forms in an FL classroom setting.

The results of this study showed that every CF technique improved the development of declarative knowledge of past-tense forms; however, the effectiveness varied according to the CF types or linguistic structures. The recast and explicit correction groups outperformed the task-only and/or test control groups on the declarative knowledge measure, irrespective of linguistic structures at the immediate posttest. Although a significant difference was also detected between the metalinguistic prompt group and the test control group at the immediate posttest, it was observed only for regular past-tense forms. Furthermore, the comparison of the CF groups indicated that explicit correction and recasts had larger effects than metalinguistic prompts on the immediate posttest. What is common to the former two types of CF is that

they are categorized as input-providing feedback. As both the implicit and explicit input-providing feedback types were effective, it is suggested that evidence types, rather than explicitness, may play a key role in developing declarative knowledge. As the learners in the recast and explicit correction groups received correct forms from the teacher, they may have easily noticed differences between their interlanguage and the target language (Schmidt, 1995, 2001), leading them to develop more declarative knowledge than those in the metalinguistic prompt group who only received negative evidence. In terms of the comparison between recasts and explicit correction, explicit correction showed larger effects than recasts for irregular past-tense forms. This result was consistent with Sheen (2007), who reported advantageous effects of explicit correction over recasts. This larger effect for explicit correction may have stemmed from the difference in awareness levels (Schmidt, 1995). Sheen (2007) claimed that recasts lead to awareness at the level of noticing, while explicit correction with metalinguistic information leads to awareness at the level of understanding. Therefore, the explicit correction group may have been able to deepen their metalinguistic understanding of irregular past-tense forms, resulting in more development of the declarative knowledge. Conversely, in terms of regular past-tense forms, only the effects of recasts were endured at the delayed posttest. The endured effects of recasts are in line with the results of Li's (2010) meta-analysis, which indicated that implicit feedback was more effective than explicit feedback in the long run. However, these results seem inconsistent with those of Ellis et al. (2006), who targeted regular past-tense forms and did not detect significant group differences between recast and control groups in the UGJT. Nevertheless, this non-significant difference was likely caused by "the fact that learners possessed ceiling levels of explicit knowledge at the beginning of the study" (Ellis et al., 2006, p. 362).

CF Effects on Procedural Knowledge

In terms of procedural knowledge, only the effects of metalinguistic prompts on learning were durable until the delayed posttest irrespective of linguistic structures. Further, the metalinguistic prompt group outperformed recasts and explicit correction groups at the delayed posttest in terms of regular past-tense forms. These results support Sheen and Ellis's (2011) statement that explicit and output-prompting feedback is the best feedback in classrooms. The comparison of the explicit correction (explicit and input-providing feedback) and metalinguistic prompt (explicit and output-prompting feedback) groups suggested that opportunities to reformulate output were more important for proceduralization than input. To turn declarative knowledge into procedural knowledge, learners must practice applying declarative knowledge in actual language use (DeKeyser, 2015). Although all treatment groups had such opportunities during the task, only the metalinguistic prompt group had additional output opportunities, which may have helped to further proceduralize their declarative knowledge (Lyster, 2004; Lyster & Sato, 2013; Yang & Lyster, 2010). The results of the study are congruent with most of the previous studies in L2 classroom contexts (Ammar & Spada, 2006; Lyster, 2004; Yang & Lyster, 2010) that reported the advantageous effects of prompts over recasts in L2. However, the results are partly inconsistent with Li and Iwashita's (2021) study, which showed that recasts were more effective than prompts in FL classroom contexts. A possible explanation for this difference is that different types of prompts were employed between Li and Iwashita's study and this study. Li and Iwashita used negotiated prompts that only required output, while this study employed metalinguistic prompts that provided metalinguistic information and opportunities for reformulation. Thus, the learners in this study may have been able to modify their output independently with the aid of metalinguistic information. However, those in the study of Li and Iwashita may not have been able to achieve this by themselves because some of the learners in their study may have lacked declarative knowledge of past tense forms to apply. Thus, the output opportunities may not have been appropriate practice opportunities for proceduralization for some of the learners in Li and Iwashita's study, even if they were provided with such opportunities. Although only the recast group outperformed the task-only or control groups at the immediate posttest, the effects were not durable. This result is partly consistent with previous studies that showed that recasts are more

effective than prompts or that recasts and prompts were equally effective (Iizuka & Nakatsukasa, 2020; Li & Iwashita, 2021). However, the results of this study suggest that additional output opportunities provided by prompts are more effective for proceduralization than input provided by recasts in the long run.

In summary, recasts and explicit corrections were generally more beneficial for deepening the metalinguistic understanding of the past tense, while metalinguistic prompts were more beneficial for proceduralizing the target structures. According to Sheen and Ellis (2011), prompts are the best form of CF for use in classroom settings. This was true for proceduralizing the target structure even in the FL classroom context of this study. However, skill acquisition theory (DeKeyser, 2015) posits that initial declarative knowledge is important during proceduralization. Thus, if learners lack a solid understanding of the target structure, and the class aims to deepen their understanding of the rule, recasts or explicit correction may be better options than prompts.

Conclusion

This study investigated the effects of different types of CF on the acquisition of L2 declarative or procedural knowledge of past-tense forms in an FL classroom context. The results suggest that the explicit and output-prompting technique is the best type of feedback to use even in an FL classroom context for developing procedural knowledge of the structures. However, input-providing feedback, such as recasts or explicit correction, should also be an option for teachers working with learners who lack adequate declarative knowledge of the target structure in the context.

As with other research, this study has several limitations. The first limitation is the study's short treatment duration. Increasing the treatment time could further clarify the differential effects of various types of CF on L2 learning. Further, the effects of CF on different linguistic structures need to be examined using linguistic structures other than the past tense. Moreover, the amount of CF provided during tasks was not completely controlled (Li, 2018) in this study. Although it may not be realistic for classroom-based studies to control fully for such differences, future results could be more generalizable if this were done. In addition, as there are no pure measures of declarative or procedural knowledge, multiple measures for each type of knowledge need to be included to generalize the results.

The Author

Yusuke Sato is a lecturer in the Dept. of Language Education of International University of Health and Welfare in Japan. His research interests cover corrective feedback and language aptitude.

Department of Language Education
International University of Health and Welfare
2600-1, Kitakanemaru, Otawara, Tochigi, 324-8501, Japan
Tel: +81-287-24-3076
Email: yskst@iuhw.ac.jp

References

- Ammar, A., & Spada, N. (2006). One size fits all? Recasts, prompts, and L2 learning. *Studies in Second Language Acquisition*, 28, 543-74. doi.org/10.1017/S0272263106060268
- 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. (2007). The differential effects of corrective feedback on two grammatical structures. In A. Mackey (Ed.), *Conversational interaction in second language acquisition* (pp. 339-360). 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>
- Ellis, R., & Shintani, N. (2013). *Exploring language pedagogy through second language acquisition research*. Routledge.
- Goo, J., & Mackey, A. (2013). The case against the case against recasts. *Studies in Second Language Acquisition*, 35, 127-65. doi.org/10.1017/S0272263112000708
- 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. (Lydia), & Iwashita, N. (2021). The role of recasts and negotiated prompts in an FL learning context in China with non-English major university students. *Language Teaching Research*, 25(2), 209-233. <https://doi.org/10.1177/1362168819839727>
- 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. (2018). Data collection in the research on the effectiveness of corrective feedback: A synthetic and critical review. In A. Gudmestad & A. Edmonds (Eds.), *Critical reflections on data in second language acquisition* (pp. 33-62). John Benjamins Publishing Company.
- Loewen, S., & Nabei, T. (2007). Measuring the effects of oral corrective feedback on L2 knowledge. In A. Mackey (Ed.), *Conversational interaction in second language acquisition: A collection of empirical studies* (pp. 361-377). Oxford University 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., & Mori, H. (2006). Interactional feedback and instructional counterbalance. *Studies in Second Language Acquisition*, 28, 269-300. <https://doi.org/10.1017/S0272263106060128>
- 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., Saito, K., & Sato, M. (2013). Oral corrective feedback in second language classrooms. *Language Teaching*, 46, 1-40.
- 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., & 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.
- Rouhi, A., & Javan-Amami, Z. (2011). Declarative knowledge: Does it mediate the effect of recasts and prompts in an EFL classroom setting? *The Journal of Asia TEFL*, 8(2), 1-25.
- 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. (2004). Corrective feedback and learner uptake in communicative classrooms across instructional settings. *Language Teaching Research*, 8, 263-300. <https://doi.org/10.1191/1362168804lr1460a>

- 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.
- Ullman, M. T. (2015). The declarative/procedural model: A neurobiologically-motivated theory of first and second language. In B. VanPatten, & J. Williams (Eds.), *Theories in second language acquisition: An introduction* (2nd ed., pp. 135-158). Routledge.
- 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>
- Zhao, Y., & Ellis, R. (2020). The relative effects of implicit and explicit corrective feedback on the acquisition of 3rd person -s by Chinese university students: A classroom-based study. *Language Teaching Research*. Advance online publication. <https://doi.org/10.1177/1362168820903343>

(Received February 23, 2021; Revised May 20, 2021; Accepted June 18, 2021)