

EFFECTS OF FORM-FOCUSED PRACTICE AND FEEDBACK ON CHINESE EFL LEARNERS' ACQUISITION OF REGULAR AND IRREGULAR PAST TENSE FORMS

Yingli Yang

University of International Business and Economics

Roy Lyster

McGill University

Conducted in English-as-a-foreign-language (EFL) classrooms at the university level in China, this quasi-experimental study compared the effects of three different corrective feedback treatments on 72 Chinese learners' use of regular and irregular English past tense. Three classes were randomly assigned to a prompt group, a recast group, or a control group and then participated in form-focused production

This study was funded by a doctoral fellowship from the Fonds Québécois pour la Recherche sur la Société et la Culture (120263) and a McGill internal fellowship for dissertation research, both awarded to the first author, and partially by a grant from the Social Sciences and Humanities Research Council of Canada (410-2006-2212) awarded to the second author. Parts of this study were presented at the annual meeting of the Canadian Association of Applied Linguistics in Vancouver, BC, on June 6, 2008, and at the McGill University Education Graduate Student Society Annual Conference in Montreal, QC, on March 14, 2008. We are grateful to the participating teachers and students, to José Correa for his assistance with the statistical analyses, and to the following research assistants for contributions to various phases of this research: Susan Ballinger, Zhengjun Lin, and Kazuya Saito. We gratefully acknowledge Younghee Sheen and two anonymous *SSLA* reviewers for their helpful comments on earlier versions of this article.

Address correspondence to: Yingli Yang, University of International Business and Economics, Room 1312, Chengxin Building, No. 10, Huixin Dongjie, Chaoyang District, Beijing, China, 100029; e-mail: yyingliuibe@gmail.com; or Roy Lyster, Department of Integrated Studies in Education, McGill University, 3700 McTavish St., Montreal, QC, H3A 1Y2, Canada; e-mail: roy.lyster@mcgill.ca.

activities that elicited the target forms. In the two feedback groups, teachers consistently provided one type of feedback (i.e., either recasts or prompts) in response to learners' errors during the activities, whereas in the control group, the teacher provided feedback only on content. Pretests, immediate posttests, and delayed posttests administered 2 weeks after the treatment assessed participants' acquisition of regular and irregular past tense forms in both oral and written production. Comparisons of group means across testing sessions using a repeated-measures ANOVA consistently revealed large effects for time. Post hoc within-group analyses of the eight immediate- and delayed-posttest measures revealed significant gains by the prompt group on all eight measures, the recast group on four, and the control group on three. The effects of prompts were larger than those of recasts for increasing accuracy in the use of regular past tense forms, whereas prompts and recasts had similar effects on improving accuracy in the use of irregular past tense forms.

Research in contexts of English as a foreign language (EFL) has shown that the exclusive use of traditional grammar translation approaches is problematic and results in learners who are able to achieve high scores on discrete-point grammar tests yet unable to communicate fluently and accurately in communicative contexts (Hu, 2003). In contrast, research in immersion and other communicatively oriented classrooms has demonstrated that even though learners achieve high levels of communicative ability, simply being exposed to rich second language (L2) input falls short of developing high levels of grammatical accuracy and error-free production (Harley & Swain, 1984; Lightbown & Spada, 1990). A solution that has proven effective in improving learners' grammatical accuracy in immersion and content-based L2 programs is the integration of form-focused instruction (e.g., Day & Shapson, 2001; Lyster, 2004b). Similarly, such an integrative pedagogy that enables learners to practice language forms within communicative contexts may also be necessary in the highly form-oriented Chinese context.

Form-focused instruction refers to "any pedagogical effort which is used to draw the learners' attention to language form either implicitly or explicitly" (Spada, 1997, p. 73). Form-focused instruction differs from the traditional grammar translation method because target forms are usually introduced during communicative interaction or embedded in meaning-based tasks and are thus expected to be retrieved more easily in similar communicative contexts. In contrast, the grammar points presented in a traditional decontextualized manner may be remembered in similar contexts, such as a discrete-point grammar test, but difficult to retrieve in a communicative context (e.g., Lightbown, 2008).

Corrective feedback (CF), a reactive type of form-focused instruction, is considered effective in promoting noticing and is thus conducive to L2 learning (Mackey & Philp, 1998; Philp, 2003; Sheen, 2007; Trofimovich, Ammar, & Gatbonton, 2007). An increasing amount of research has been conducted, in both laboratory and classroom contexts, regarding the type of evidence CF provides (Egi, 2007a; Leeman, 2003) as well as the effectiveness of different types of CF (Ammar & Spada, 2006; Ellis, Loewen, & Erlam, 2006; Lyster, 2004a; Lyster & Izquierdo, 2009; Sheen).

Most studies that compare different CF techniques have targeted rule-based grammatical features (e.g., Ammar & Spada, 2006; Ellis et al., 2006; Lyster, 2004a; Sheen, 2007). Until now, the efficacy of different types of CF on various types of grammatical features as well as the cognitive mechanisms that these different techniques invoke remain to be examined (Ellis, 2007). The present study investigates how different types of CF together with form-focused practice activities may benefit Chinese EFL learners' acquisition of regular and irregular English past tense, in an attempt to determine whether recasts and prompts have differential effects on the acquisition of rule-based versus exemplar-based grammatical forms.

THE ROLE OF RECASTS AND PROMPTS IN SLA

Researchers tend to differentiate feedback in terms of how explicit or implicit it is. Implicit types of feedback often take the form of recasts (Long, 1996; Long & Robinson, 1998). However, this dichotomous classification of feedback can be problematic (Egi, 2007a). Research shows that recasts can also be quite explicit (Ellis & Sheen, 2006; Sheen, 2006), depending on the context (see Sheen, 2004) as well as the characteristics of the recasts (e.g., linguistic targets, length, and number of changes). Similarly, explicit types of feedback can also take a variety of forms. For some researchers, explicit feedback involves the indication that an error has been made (referred to as *explicit rejection* in Carroll, 2001, and Carroll & Swain, 1993); for others, it also includes metalinguistic information (Ellis et al., 2006), the correct form (Lyster & Ranta, 1997), or both (Sheen, 2007).

Although of theoretical interest, comparing the effects of different feedback techniques only in accordance with their implicit or explicit nature thus remains problematic. An alternative and more categorical way of classifying feedback types is to distinguish between input-providing CF and output-pushing CF (Ellis, 2006): Input-providing CF supplies the correct reformulation through recasts and explicit correction; output-pushing CF withholds the correct reformulation and instead encourages learners to self-repair through prompts such as clarification requests, repetition of learner error, metalinguistic clues, and elicitation (Lyster,

2002, 2007; Lyster & Mori, 2006, 2008; Ranta & Lyster, 2007). Input-providing and output-pushing feedback engages learners in different levels of cognitive processing, which arguably involves cognitive comparison in working memory in the case of input-providing CF and retrieval from long-term memory in the case of output-pushing CF.

Recasts and prompts differ not only in terms of whether the target forms are given but also in the types of evidence provided. Nicholas, Lightbown, and Spada (2001) argued that recasts afford learners with positive evidence, but whether negative evidence is also provided is less clear. Other researchers (e.g., Egi, 2007a; Ellis & Sheen, 2006) believe that whether recasts provide positive evidence, negative evidence, or both largely depends on learners' perceptions, which, in turn, determine the effectiveness of recasts. It has been argued that by providing positive evidence in classroom input, recasts may help learners establish new knowledge. Prompts, in contrast, aim to provide negative evidence because they signal that the learners' utterance is problematic. The self-repair process is claimed to help learners to reanalyze what has already been learned (at some level) and to restructure their interlanguage (Lyster, 2002). According to de Bot (1996), learners benefit more from being pushed to "make the right connection on one's own" (p. 549) than from hearing the correct structures in the input. Furthermore, prompts may help learners to gain greater control over already acquired forms and to access these forms more quickly.

Despite the ongoing debate that pertains to the classification of CF and the types of evidence that CF provides, there has been a growing interest in the role of different types of CF in SLA in the past decade, which has spawned a substantial number of studies in a variety of classroom and laboratory contexts. Classroom studies generally show more benefits for prompts than for recasts. With younger learners, Lyster (2004a) investigated the effects of form-focused instruction and CF on students' acquisition of grammatical gender in French. He found that form-focused instruction was more effective when combined with prompts than with recasts. Also with young learners, Ammar and Spada (2006) investigated the potential benefits of recasts and prompts on the acquisition of possessive determiners by French-speaking English-as-a-second-language (ESL) learners. Whereas both CF groups showed superior performance compared to the control group, the group that received prompts significantly outperformed the recast group on written and oral posttests. Prompts were especially effective for learners who had pretest scores below 50%, whereas learners with pretest scores above 50% benefited similarly from both recasts and prompts.

With adult learners, Ellis et al. (2006) investigated the differential effects of recasts and metalinguistic feedback (i.e., a prompt consisting of repetition of the error followed by a clue such as *you need past tense*) on the acquisition of regular past tense in English. Results showed that

metalinguistic feedback was overall more effective than recasts, a finding more often demonstrated in the delayed posttests than the immediate posttests. Loewen and Nabei (2007) compared the effects of recasts, clarification requests, and metalinguistic feedback provided during meaning-focused tasks on English question formation in a Japanese EFL context. In contrast with previous studies (e.g., Lyster, 2004a; Ellis et al.), there was no significant difference among the feedback groups, which, as the researchers acknowledged, might be due to the brevity of the treatment session and the varying amount of feedback provided to different treatment groups during the treatment session. Additionally, the sample sizes in the treatment groups were small, and there was no delayed posttest, which might more rigorously show effects of feedback (Ellis et al.).

Although classroom quasi-experimental studies have generally demonstrated the overall beneficial effects of prompts over recasts, results obtained from laboratory studies have not yielded similar results. For example, McDonough (2007) carried out a laboratory study in a Thai EFL context that compared recasts with clarification requests. The results suggested that both clarification requests and recasts facilitated the emergence of simple past activity verbs. Similarly, a laboratory study by Lyster and Izquierdo (2009) that investigated adult L2 learners' acquisition of French grammatical gender reported similar effects for prompts and recasts. The researchers argued that these two types of feedback provide different opportunities for L2 learning: Learners who receive prompts benefit from the repeated exposure to negative evidence as well as the opportunity to modify their output, whereas recasts provide positive evidence as well as opportunities for learners to infer negative evidence due to the enhanced saliency of recasts in dyadic interactions.

CF EFFECTIVENESS IN RELATION TO TARGET STRUCTURES

Several studies have suggested that recasts of phonological and lexical errors are more noticeable than recasts of errors in morphosyntax (Carpenter, Jeon, MacGregor, & Mackey, 2006; Ellis, Basturkmen, & Loewen, 2001; Han, 2008; Lyster, 1998; Mackey, Gass, & McDonough, 2000; Sheen, 2006). These empirical studies suggest that CF might be relatively more effective for L2 phonological and lexical development than for other domains such as L2 morphosyntax because learners tend to perceive the corrective force of teachers' CF on pronunciation and vocabulary errors.

A study by Egi (2007b), in which learners of Japanese participated in dyadic communicative tasks with native speakers and received recasts, provided some evidence that learners processed recasts differently according to whether the linguistic target was morphosyntactic or lexical. Using stimulated recall measures and tailor-made posttests to investigate

which components of recasts (i.e., positive or negative evidence) benefit learners' morphosyntactic and lexical development, Egi concluded that learners' noticing of positive evidence may be more likely to result in immediate interlanguage changes in lexical learning than in morphosyntactic learning. In their meta-analysis of 25 interaction studies (7 lexis studies and 18 grammar studies), Mackey and Goo (2007) confirmed that the effects of interaction are significantly larger on lexical development than on grammatical development, especially on immediate posttests ($d = 1.32$ for lexical items, $d = 0.59$ for grammatical structures).

Most laboratory and classroom experimental studies of CF effectiveness have targeted grammatical structures—with variable effectiveness—and these grammatical targets vary across studies in terms of difficulty and inherent properties. However, some questions remain unanswered: (a) which grammatical structures are most amenable to CF and (b) whether recasts and prompts have differential effects on different kinds of grammatical structures. Ellis (2007) seemed to be the only study to date that compared the differential effects of CF on two different grammatical targets: regular past tense *-ed* and comparative *-er*, which were hypothesized to differ in terms of grammatical difficulty, the regular past tense being the easier of the two structures. Three classes of students in a private language school participated in the study. The students in the treatment groups took part in 1-hr communicative tasks during which the researcher provided CF, either in the form of metalinguistic feedback, which served as a prompt, as in *you need a comparative adjective*, or recasts. The group that received metalinguistic explanations showed more improvement on comparative than on past tense forms, whereas the recast group did not show any significant gains over the control group on any of the measures. Ellis (2007) concluded by proposing that “what is needed in future research is to determine *how* linguistic factors may determine *when* different types of feedback will work for acquisition” (p. 360).

REGULAR AND IRREGULAR PAST TENSE FORMS IN ENGLISH

Past tense forms were chosen as the target of this quasi-experimental study for two reasons. First, although the past tense is introduced early on in textbooks, learners demonstrate difficulties in gaining full control of this structure, even at intermediate or advanced levels (Ellis et al., 2006). Mandarin Chinese, the native language of all participants in this study, does not mark past tense with morphological change (Cai, 2007); past tense marking thus remains especially difficult for Chinese EFL learners, owing arguably to expectancy, which is one of the factors that influences the noticing of grammatical features (Schmidt, 2001).

Second, past tense forms are either regular or irregular; this distinction allowed us to compare not only the effects of recasts and prompts but also their differential effects on different linguistic features: one rule based (regular past tense) and the other item based (irregular past tense). Regular English past tense forms are a rule-based feature because there is a clear general rule: Add *-ed* to the base form of a regular verb. In contrast, there is no similarly clear rule for irregular past tense forms, which are thus considered to be exemplar based or item based (Ellis, 2005). Due to the predictable and rule-based nature of regular past tense verbs and the complex and unpredictable characteristics of irregular past tense verbs, the representation and acquisitional processes of the two kinds of verbs also presumably vary to a great extent (Ellis). The comparison of learners' acquisition of regular versus irregular past tense forms allows for the investigation of a grammatical feature with different realizations in form but identical semantic and functional value, thus precluding any variability that may result from differences in the communicative value of the target forms.

The representation and processing of regular versus irregular morphology, however, continues to be a topic of unresolved debate. On the one hand, analogy-based network models, such as the connectionist model (Rumelhart & McClelland, 1986), posit that all inflectional morphology, whether regular or irregular, is based on frequency-driven generalizations from exposure to input. On the other hand, researchers such as Pinker (Pinker & Prince, 1994; Pinker & Ullman, 2002) propose that regular and irregular past tense forms are processed through two distinct channels. Specifically, the processing of regular past tense requires computation and therefore involves the procedural system, whereas the representation and processing of irregular past tense requires declarative memory. Similarly, Skehan (1998) described a dual-mode system that comprises two interrelated representational systems that coexist and interact in language processing: an analytic rule-based system and a memory-driven exemplar-based system. If rule-based and item-based structures coexist in the language system but involve different processing mechanisms, then it follows that different types of instructional intervention will have variable effects on changes in rule-based and item-based representations. For example, DeKeyser (1995) found that participants learned simple abstract morphosyntactic rules in a miniature linguistics system significantly better under explicit learning conditions than under implicit conditions, but they learned similarity patterns better in the implicit-inductive conditions than in the explicit-deductive conditions. Robinson and Ha (1993) also suggested that both structural complexity and developmental readiness determine whether learners rely on item retrieval or rule searching in L2 processing.

RESEARCH QUESTIONS

Building on the theoretical argument that recasts and prompts provide different types of linguistic evidence and engage learners in different types of processing (Lyster, 2004a; Lyster & Izquierdo, 2009; Lyster & Mori, 2006) as well as on results from empirical investigations that have begun to show that the effects of feedback may be mediated by linguistic structures (Egi, 2007b; Ellis, 2007; Mackey, 2006), the present study compares the relative effects of recasts and prompts on EFL learners' acquisition of regular and irregular past tense forms. To investigate these effects, the following research questions are addressed:

1. Do the groups that performed form-focused production activities while receiving CF show an overall superiority in learning regular and irregular past tense over the control group, which performed the same communicative classroom activities but without receiving CF on past tense errors?
2. Do recasts have differential effects on the acquisition of regular versus irregular English past tense forms?
3. Do prompts have differential effects on the acquisition of regular versus irregular English past tense forms?

METHOD

Participants

The present study was conducted in three intensive reading classes, which, according to the teachers, relied heavily on the explanation of grammar points and translation exercises as well as reading comprehension and vocabulary activities. Each of these three intact classes was assigned to a treatment condition (i.e., recast, prompt, control). Each class had between 22 and 26 students, with some fluctuation in numbers due to different conditions for oral and written measures as well as some attrition across testing times.

Participants included 72 undergraduate EFL students (64 females and 8 males) at a university in northern China. All in their second year of university and majoring in English language and literature, their ages ranged from 18 to 24, with an average age of 20. The students reported an average of 9.25 years of experience in learning English and were uniformly from a Chinese ethnic background, with Mandarin Chinese as their native language.

The control group teacher was the regular classroom teacher, a male with a master's degree in English literature and 1 year of teaching experience. The teachers of the CF groups were both female and were filling in for the regular classroom teachers (due to unforeseen circumstances).

The teacher assigned to the recast condition had a master's degree in English linguistics, had been teaching for 3 years, and had participated in the pilot study. The teacher assigned to the prompt condition was the first author, a doctoral candidate with 2 years of prior teaching experience. The two teachers of the CF groups were assigned to the prompt or the recast group based on their own preference for feedback conditions. Before the actual experiment, participating teachers were provided with the schedule of the study, general procedure, and treatment materials. The teachers were also provided with a booklet that described the two CF types with examples.

Feedback Conditions

Recasts were operationalized as the teacher's reformulation, either partial or full (Sheen, 2007), of the students' ill-formed utterances that contained past tense errors. To allow for learner responses following recasts to flow naturally, the teacher was not explicitly requested either to discourage repair, as in (1), or to encourage repair, as in (2).

- (1) Student: *Well. Cinderella was such an, such a simple girl that she never knew what is waiting for her.*
Teacher: *What was waiting for her. Next picture.*
- (2) Student: *And they have a happy life after and many years they have a lot of kids around the...*
Teacher: *They had...*
Student: *They had a lot of kids around the palace.*

Corrective feedback that withheld the target form and pushed learners to self-repair was given as prompts that included four feedback types, illustrated in the examples in (3)–(6) taken from the data of the present study.

- (3) Metalinguistic clue
Student: *I went to the train station and pick up my aunt.*
Teacher: *Use past tense consistently.*
Student: *I went to the train station and picked up my aunt.*
- (4) Repetition
Student: *Mrs. Jones travel a lot last year.*
Teacher: *Mrs. Jones travel a lot last year?*
Student: *Mrs. Jones traveled a lot last year.*
- (5) Clarification request
Student: *Why does he fly to Korea last year?*

Teacher: *Pardon?*

Student: *Why did he fly to Korea last year?*

(6) Elicitation

Student: *Once upon a time, there lives a poor girl named Cinderella.*

Teacher: *Once upon a time, there...*

Student: *There lived a girl.*

During the implementation by all three teachers of a set of form-focused oral practice activities, the teachers in the prompt and the recast groups provided the designated type of feedback in response to past tense errors, whereas the teacher in the control group ignored past tense errors and instead reacted to the meaning of student responses, as in (7) and (8).

(7) Student: *George tried a pair of scissors and it flew up and put his eye out.*

Teacher: *Put his eye out, already put his eye out?*

Student: *Nearly.*

Teacher: *Nearly. Right?*

(8) Student: *George's head.*

Teacher: *What, head?*

Student: *Hat.*

Teacher: *umm. It was George's hat. Anyway, it should be the hat that saved his life.*

In (7), the teacher ignores the past tense error and instead asks for clarification about the extent of damage caused by the scissors. In (8), the teacher repeats the student's utterance *head* to elicit the more accurate item *hat*. These teacher responses were classified as feedback on content because they addressed possible confusion or misunderstanding in meaning.

Treatment Instruments

The classroom treatment consisted of four form-focused production activities developed for the purpose of this study and implemented in sequence by the instructors for approximately 2 hr over a period of 2 weeks. Prior to the actual treatment phase, all four tasks had been piloted in another class with the same proficiency level at the same university. Additionally, the students' textbook and the teachers were consulted to tailor the material to the students' appropriate level in terms of vocabulary. The four activities included two dictogloss tasks, one question-and-answer activity, and one picture-cued narrative task. While the students were interacting with their teachers during these

activities, the teachers in the CF groups provided either recasts or prompts in response to errors in past tense, whereas the teacher in the control group ignored past tense errors and instead reacted to content.

Dictogloss Activity. Two dictogloss tasks (Wajnryb, 1990; see Swain & Lapkin, 2001) were used for the purpose of raising learners' awareness of past tense forms and to elicit their use in oral production. One was adapted from the classic fairy tale "Cinderella" (Perrault, 1697/2004) and the other from a humorous story, "Three Men in a Boat" (Jerome, 1889). The teacher read the narrative twice at a normal speed. In the first reading, the students were not allowed to take notes, whereas during the second reading, the students wrote down what they heard as much as possible. In pairs, students then compared their notes and reconstructed the text before being called on to narrate part of the complete story, during which time the treatment group teachers provided CF on past tense errors. In the control group, the teacher waited until the students finished presenting their texts to summarize some common errors. Finally, teachers in all three groups provided the original texts with the past tense verbs highlighted in bold for students to compare with their own texts.

Question-and-Answer Activity. To provide students with individual question-and-answer production practice with a focus on past tense forms, two sets of cards (one for the students and the other for the teacher) that included 30 questions and corresponding information for the answers were prepared. On each of the teacher's cards, there was a question about an action in the past (e.g., *Did Mrs. Jones travel last year?*). Below each question on the teacher's card, a specific noun or adverbial phrase (or clause) was presented (e.g., *a lot*). On the corresponding student card, only the specific phrase *a lot* was given. After listening to the teacher's question, the students were expected to answer affirmatively and to incorporate the information on their cards to complete the answer (e.g., *Yes, she traveled a lot last year*).

Picture-Cued Narrative Activity. This type of activity has been used frequently in previous studies on past tense (Ellis, 2007; Ellis et al., 2006; Han, 2002). In the present study, pictures illustrating the story of "Cinderella" were chosen because this fairy tale had previously appeared as a dictogloss and thus was expected to allow students to communicate without being overloaded with new information. The teacher first divided the class into 10 small groups of two or three students according to their seating positions. Ten numbered pictures, each with several characters depicting a scene of the story, were distributed among the groups. The students in each group were allowed to prepare for 5 min, and then together the 10 groups told the story in sequence. As students presented their part of the story, teachers in the treatment groups provided one

type of CF following errors in past tense, whereas in the control group, the teacher only provided feedback on vocabulary errors.

Treatment Procedures

Students' performance on the accurate use of the past tense was measured three times: before, immediately after, and 2 weeks after the treatment. For each testing session, an oral production measure and a written production measure were administered. The oral production measure in the present study was designed to assess the learners' implicit knowledge through spontaneous use of English past tense (Ellis, 2005; Erlam, 2006). The written production measure also elicited the use of past tense in a productive way but differed from the oral narrative in that the learners had more time to reflect on their use of language.

Oral Production. Adapted from an online grammar exercise, the oral production measure required learners to retell a story based on a series of word cues. The test consisted of a short passage that described a crazy beach party and included 9 types (9 tokens) of regular and 17 types (18 tokens) of irregular past tense forms in English. For 3 min, the students silently read the story in the L2, before returning the text to the researcher. To assist them in retelling the story, students were given a series of word cues, including all content words as well as verbs in their base form and adverbial phrases that indicated past time referent (e.g., *in the year 2000, fly...*). These adverbials created an obligatory context for retelling the story in the past tense: For example, the student could say, *In the year 2000, I flew to Korea for the first time.* There was no planning time after the word cues were presented, which limited the learners' ability to draw on explicit grammatical knowledge (Ellis, 2005).

To control for test-retest effects, two versions (A and B) of the oral test were used alternately over the three testing sessions following the sequence ABA for all participants. Each version had exactly the same target regular and irregular verbs, but the sequence of actions and people who performed the actions alternated in the two versions so that participants could not rely exclusively on their memory to perform the task. The oral data during the three testing sessions were audio-recorded, transcribed, and then, using obligatory context analysis, coded for accuracy by the first author, a Mandarin speaker with nativelike fluency of English. A second rater then coded 12% of the oral data, taken randomly from the three testing sessions. The interrater reliability for the oral production data was considered high at 91.6%.

Written Production. The written production measure was designed in the form of a written narrative, which aimed to test the students' knowledge of

the past tense forms in open-ended written production. Provided with a topic and 12 verbs (5 regular and 7 irregular), participants were required to compose a story in 15–20 min. The order of the 12 verbs was randomly altered at each testing session and a different topic to describe was provided: (a) a crazy day, (b) the day the Titanic sank, and (c) Cinderella. Instructions required the use of transition words (e.g., *after*, *before*, *finally*) and forbade the use of infinitives, participles, or negation; thus, for example, the students were not allowed to use *a party was thrown* or *the prince did not throw a party*. The written data were coded for accuracy by the first author and then a random selection of 12% of the written data was coded by a second rater. The interrater reliability for the written production data again proved high at 94.9%.

Exit Questionnaire. A short exit questionnaire with a combination of closed-item and open-ended questions (Mackey & Gass, 2005) was administered immediately following the immediate posttest session (see the Appendix). This questionnaire was adapted from Sheen's (2007) exit questionnaire, which was designed to examine whether the students had become aware of the focus of the CF treatments and tests. Two open-ended questions were added to explore the learners' general attitude toward CF and the production tasks used in the present study. The questionnaire was written in English, but the students were allowed to use Chinese to answer.

RESULTS

The audio recordings of the instructional treatment sessions were transcribed and analyzed to identify the number of errors in the use of past tense, CF moves, and immediate repair across different groups. As shown in Table 1, all three groups produced a relatively similar number of errors during the treatment sessions. Table 1 further illustrates that the teachers consistently provided the appropriate CF type in their respective treatment group with only minimal use of the wrong feedback type. Of the 27 prompts, 8 were elicitation moves, 8 were metalinguistic clues, 6 were repetition, and 5 were clarification requests. There were 25 instances of repair following prompts: Only one clarification request and one repetition did not elicit immediate repair. There was only one instance of repair in the recast group, and the control group did not produce any repair. The low rate of repair in the recast group is noteworthy because a post hoc analysis of the audio recordings revealed that 61% of the recasts provided to this group were followed by pragmatically appropriate opportunities for uptake.

A mixed-design two-way ANOVA with repeated measures on transformed data was employed to assess differences in accuracy scores

Table 1. Distribution of error, feedback moves, and repair across groups

| Group | Errors | Feedback moves | | Repairs |
|---------|--------|----------------|---------|---------|
| | | Prompts | Recasts | |
| Prompt | 31 | 27 | 1 | 26 |
| Recast | 32 | 1 | 23 | 1 |
| Control | 45 | 2 | 1 | 0 |

among the prompt group, the recast group, and the control group across three testing times. Because initial examination of residual plots indicated problems with normality or heterogeneity of variances, the data were transformed using the arcsine square-root transformation to normalize the distribution of the percentage scores (Sokal & Rohlf, 1995). Separate ANOVAs were conducted for oral and written production tests using the mixed-model procedure (PROC MIXED) of Statistical Analysis System (SAS), version 9.1.3 (SAS Institute, Inc., 2002–2003). Because the groups were unbalanced, a Kenward-Roger correction was applied to the degrees of freedom for better control of Type I error. Post hoc multiple comparisons were performed by calculating all-pair contrasts using the SAS macro %SimTests with the unconstrained step-down method, a closed testing procedure to account for logical dependencies between the hypotheses. Moreover, *p*-values were adjusted for multiplicity (Westfall, Tobias, Rom, Wolfinger, & Hochberg, 1999). All tests were two-sided and the significance level was set at .05.

Effect sizes were calculated to measure the magnitude of effects of the different treatments with respect to within-group contrasts (i.e., time difference) and between-group contrasts (i.e., group difference) in contexts in which significance was detected. Taking into account the nature of instructed SLA studies, Cohen’s *d* index was chosen following recommendations by Norris and Ortega (2000, 2006) and effects were considered as small ($d \leq 0.20$), medium ($0.20 < d < 0.80$), or large ($d \geq 0.80$).

Oral Test Results

Irregular Past Tense. The means and standard deviations of the accuracy scores in the use of irregular past tense forms in oral production for each group across three testing times are displayed in Table 2, and the group means are presented graphically in Figure 1. A repeated-measures ANOVA revealed that there was a significant time effect,

Table 2. Group means of irregular past tense percentage scores in oral production

| Group | Pretest | | | Immediate posttest | | | Delayed posttest | | |
|---------|----------|-----------|----------|--------------------|-----------|----------|------------------|-----------|----------|
| | <i>M</i> | <i>SD</i> | <i>N</i> | <i>M</i> | <i>SD</i> | <i>N</i> | <i>M</i> | <i>SD</i> | <i>N</i> |
| Prompt | 63.64 | 23.99 | 22 | 78.95 | 13.30 | 19 | 76.14 | 11.91 | 17 |
| Recast | 59.78 | 20.74 | 25 | 76.85 | 21.07 | 18 | 66.67 | 23.08 | 20 |
| Control | 55.56 | 20.85 | 25 | 72.69 | 15.01 | 24 | 64.44 | 16.75 | 20 |

$F(2, 62) = 20.16, p < .0001$, but neither a significant group effect, $F(2, 69) = 1.76, p = .18$, nor a significant interaction effect between group and time, $F(4, 72) = 0.80, p = .53$. Post hoc analyses revealed that all three groups significantly increased their scores from the pretest to the posttest, with equally large effects (prompt: $d = 0.78$; recast: $d = 0.81$; control, $d = 0.94$). Additionally, the prompt group ($p < .05$) and the control group, ($p < .005$), but not the recast group, demonstrated significant gains in accuracy scores from pretest to delayed posttest, with medium effects for both groups (prompt: $d = 0.66$; control: $d = 0.46$).

Regular Past Tense. Table 3 displays group means and standard deviations of all three groups’ accurate use of regular past tense across three testing times in oral production, whereas Figure 2 graphically displays the group means over time. A repeated-measures ANOVA revealed a significant time effect, $F(2, 62) = 9.88, p < .0002$, and a significant group effect, $F(2, 66) = 3.35, p < .05$, but no Group \times Time interaction effect, $F(2, 72) = 0.42, p = .80$. Post hoc analyses revealed that only the prompt

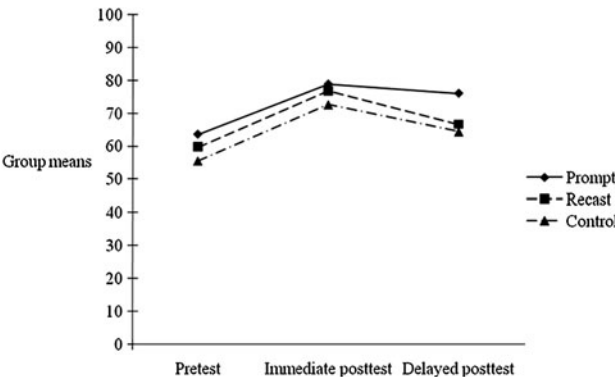


Figure 1. Group means of irregular past tense percentage scores in oral production.

Table 3. Group means of regular past tense percentage scores in oral production

| Group | Pretest | | | Immediate posttest | | | Delayed posttest | | |
|---------|----------|-----------|----------|--------------------|-----------|----------|------------------|-----------|----------|
| | <i>M</i> | <i>SD</i> | <i>N</i> | <i>M</i> | <i>SD</i> | <i>N</i> | <i>M</i> | <i>SD</i> | <i>N</i> |
| Prompt | 62.62 | 29.61 | 22 | 76.61 | 18.10 | 19 | 79.09 | 16.14 | 17 |
| Recast | 61.33 | 24.03 | 25 | 72.22 | 17.57 | 18 | 70.56 | 22.30 | 20 |
| Control | 54.67 | 22.89 | 25 | 64.82 | 18.44 | 24 | 64.44 | 20.58 | 20 |

group made significant gains, with medium effects between pretest and immediate posttest ($p < .05$; $d = 0.57$) and between pretest and delayed posttest ($p < .05$; $d = 0.69$).

Written Test Results

Irregular Past Tense. The means and standard deviations of learners’ accurate use of irregular past tense forms in the written test across three testing times appear in Table 4, and Figure 3 graphically displays the group means over time. A repeated-measures ANOVA revealed a significant effect for time, $F(2, 67) = 26.08$, $p < .0001$, as well as a significant Group \times Time interaction effect, $F(4, 79) = 5.35$, $p < .01$, but no significant group effect, $F(2, 68) = 1.48$, $p = .23$. Post hoc analyses confirmed that the prompt group significantly improved from the pretest to the posttest

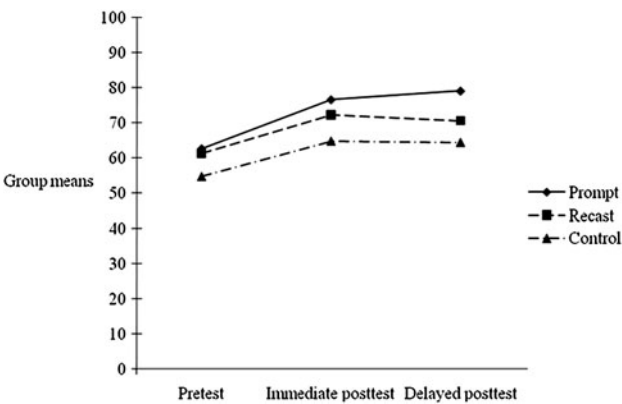


Figure 2. Group means of regular past tense percentage scores in oral production.

Table 4. Group means of irregular past tense percentage scores in written production

| Group | Pretest | | | Immediate posttest | | | Delayed posttest | | |
|---------|----------|-----------|----------|--------------------|-----------|----------|------------------|-----------|----------|
| | <i>M</i> | <i>SD</i> | <i>N</i> | <i>M</i> | <i>SD</i> | <i>N</i> | <i>M</i> | <i>SD</i> | <i>N</i> |
| Prompt | 30.09 | 23.48 | 22 | 77.27 | 27.11 | 22 | 75.45 | 20.41 | 22 |
| Recast | 42.61 | 25.08 | 23 | 67.83 | 25.40 | 23 | 65.45 | 26.32 | 22 |
| Control | 53.85 | 23.85 | 26 | 60.80 | 26.13 | 25 | 59.20 | 21.97 | 25 |

($p < .0001$), with a large effect ($d = 1.51$), and maintained this improvement in scores at the time of the delayed posttest ($p < .0001$), with a large effect ($d = 1.65$). Similar results were found for the recast group, which also demonstrated significant gains from pretest to immediate posttest ($p < .005$), with a large effect ($d = 0.99$), as well as from the pretest to the delayed posttest ($p < .0005$), also with a large effect size ($d = 0.88$). The control group made no significant gains over time. Post hoc analyses also revealed that the prompt group significantly outperformed the control group ($p < .05$) at the time of the delayed posttest in accurate use of irregular past tense forms in written production with a medium effect size ($d = 0.76$).

Regular Past Tense. Table 5 displays the group means and standard deviations for the accuracy scores obtained by the three groups over time in their use of regular past tense in written production; Figure 4 graphically presents the group means over time. The ANOVA results

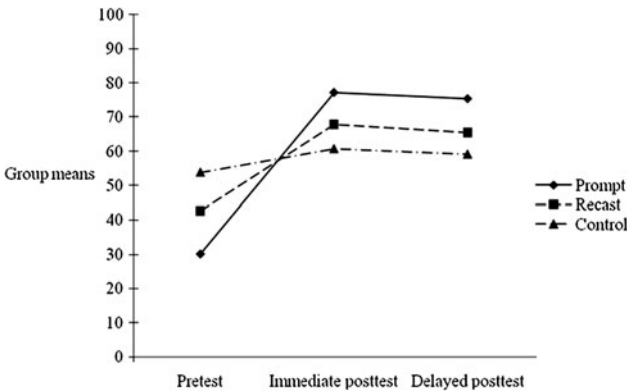


Figure 3. Group means of irregular past tense percentage scores in written production.

Table 5. Group means of regular past tense percentage scores in written production

| Group | Pretest | | | Immediate posttest | | | Delayed posttest | | |
|---------|----------|-----------|----------|--------------------|-----------|----------|------------------|-----------|----------|
| | <i>M</i> | <i>SD</i> | <i>N</i> | <i>M</i> | <i>SD</i> | <i>N</i> | <i>M</i> | <i>SD</i> | <i>N</i> |
| Prompt | 65.58 | 21.89 | 22 | 94.16 | 8.43 | 22 | 84.42 | 15.84 | 22 |
| Recast | 73.91 | 19.09 | 23 | 91.30 | 11.18 | 23 | 84.42 | 13.88 | 22 |
| Control | 67.58 | 26.36 | 26 | 85.14 | 14.57 | 25 | 74.88 | 17.14 | 25 |

again revealed a significant main effect for time, $F(2, 67) = 27.59, p < .0001$, but no significant group effect, $F(2, 69) = 2.13, p = .13$, or any significant interaction effect between group and time, $F(4, 80) = 0.98, p = .42$. Post hoc analyses confirmed that all three groups significantly improved their scores from pretest to immediate posttest, with large effect sizes (recast: $d = 1.11$; prompt: $d = 1.92$; control: $d = 0.82$), and that only the prompt group demonstrated significant gains between pretest scores and delayed posttest scores ($p < .05$), with a large effect size ($d = 0.98$).

Questionnaire Results

A total of 94 participants completed the exit questionnaire (see the Appendix) following the immediate posttest. In response to the first, multiple-choice question (*Now that you have completed the tasks and the tests, what do you think they were all about?*), 83 participants considered

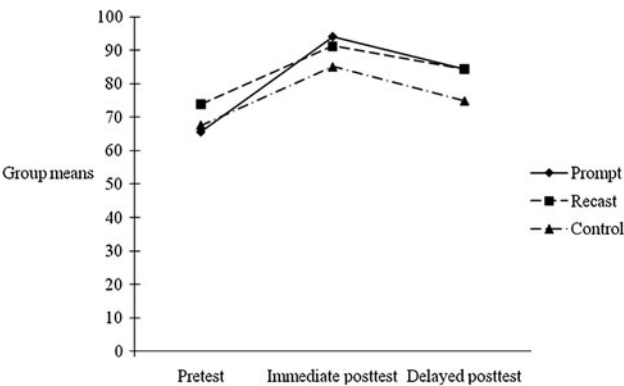


Figure 4. Group means of regular past tense percentage scores in written production.

the tasks and tests to be about listening and speaking, 10 chose writing, 10 chose grammar, and 7 selected vocabulary as the focus. The second question asked students to write a short paragraph indicating what they had learned from the experience. The most frequently mentioned themes were skill improvement in dictation ($n = 33$), listening ($n = 32$), speaking ($n = 26$), and vocabulary ($n = 26$). Only 17 participants reported learning grammar. Six suggested a focus on tense and only two explicitly indicated past tense as the focus of their learning.

The third question, which asked participants if they liked the instructional activities, elicited 60 affirmative responses, 24 negative responses, and 10 undecided. In response to the fourth question, which asked participants about their attitude toward feedback, 75 participants indicated that they liked to be corrected while speaking in class, whereas 19 participants indicated that they did not.

Summary

Table 6 provides a summary of the significant within-group time contrasts (i.e., differences within each group in performance across testing times) along with the magnitude of effect sizes. Of eight possible scores (two scores in oral production and two in written production contrasted at two different times), the prompt group achieved significance on all eight measures; the recast group achieved significance on four measures, and the control group achieved significance on three measures. Overall, the prompt group exhibited relatively large effect sizes in statistically significant contexts compared to the recast and control groups. Additionally, the prompt group significantly outperformed the control group in accurate use of irregular past tense forms at the delayed posttest.

DISCUSSION

The prompt group excelled by making significant gains on all measures, regardless of time (whether immediate or delayed), mode (whether oral or written), and structure (whether regular or irregular). Additionally, only the prompt group showed any significant between-group differences in its favor, significantly outperforming the control group in the oral production of irregular past tense at the delayed posttest. In oral production, the recast group showed significant short-term within-group gains in the use of irregular past tense but no significant gains in the use of regular past tense. In written production, the recast group showed significant short- and long-term gains in the use of irregular past tense forms and short-term gains in the use of regular past tense.

Table 6. Significant within-group time contrasts and magnitude of effect sizes

| Test | Control group | Recast group | Prompt group |
|------------------------------|-------------------------------|------------------------------|--|
| Oral production irregular | | | |
| Immediate posttests | Medium effects ($d = 0.78$) | Large effects ($d = 0.81$) | Large effects ($d = 0.94$) |
| Delayed posttests | Medium effects ($d = 0.46$) | ns | Medium effects ($d = 0.66$) ^a |
| Oral production regular | | | |
| Immediate posttests | ns | ns | Medium effects ($d = 0.57$) |
| Delayed posttests | ns | ns | Medium effects ($d = 0.69$) |
| Written production irregular | | | |
| Immediate posttests | ns | Large effects ($d = 0.99$) | Large effects ($d = 1.51$) |
| Delayed posttests | ns | Large effects ($d = 0.88$) | Large effects ($d = 1.65$) |
| Written production regular | | | |
| Immediate posttests | Large effects ($d = 0.82$) | Large effects ($d = 1.11$) | Large effects ($d = 1.92$) |
| Delayed posttests | ns | ns | Large effects ($d = 0.98$) |

Note. The abbreviation “ns” corresponds to not significant.
^a The prompt group also significantly outperformed the control group on this measure.

The control group only made significant short- and long-term gains in oral production of irregular past tense forms and short-term gains in written production of regular forms. The overall larger effect sizes in written production compared to oral production, for all groups, may reflect the fact that fewer items were tested in writing ($n = 12$) than in oral production ($n = 27$).

Prompts are predicted to help learners gain greater control over already acquired forms and access them in faster ways, whereas recasts might be more effective for providing positive exemplars of relatively new linguistic forms. Given their prior experience in Chinese EFL classrooms, the participants in the present study had received ample exposure to regular and irregular past tense forms in English as well as the rules that govern their formation but were arguably at a point at which they could benefit from opportunities for production practice and feedback in the context of classroom interaction. The superior performance of the prompt group over the recast group may be due to the fact that participants in the prompt group were pushed to retrieve target forms that already existed in their long-term memory, and the repeated instances of such retrieval strengthened connections between knowledge stored in memory and actual language production. The results that reveal advantages for prompts over recasts and for prompts versus no feedback are congruent with findings of a number of previous quasi-experimental studies in a variety of instructional settings (Ammar & Spada, 2006; Ellis, 2007; Ellis et al., 2006; Lyster, 2004b). The findings of the present study suggest that the differential effects of CF treatments were further mediated by (a) the extent of immediate self-repair, (b) varying degrees of CF saliency, and (c) the nature of the structural targets.

Self-Repair

Prompts and recasts differ in the extent to which learners modify their output immediately after CF. In the present study, because learners consistently modified their output after prompts but not after recasts, production opportunities following CF could, to some extent, explain the greater benefits yielded by prompts in comparison to recasts and to no feedback. Modified output is considered to contribute to L2 development, as suggested by Levelt's (1989) speech production model, because it allows learners to either generate new output or reprocess their original output, both of which trigger additional grammatical encoding (Izumi, 2003). Similarly, according to Swain's (1993, 1995) output hypothesis, learners need to be pushed to make use of their own resources and to extend their productive repertoire.

The effectiveness of prompts has indeed been associated with the self-repair process they trigger (Loewen & Philp, 2006; Lyster & Izquierdo, 2009), arguably because by self-repairing in response to negative evidence, learners are pushed to engage in some degree of reanalysis that increases awareness of the rule or item in question (de Bot, 1996; Lyster, 2004b). In the case of recasts, however, many studies have shown that their variable effectiveness is associated more with specific input features of positive exemplars than with modified output (Loewen & Philp; Lyster & Izquierdo; Mackey & Philp, 1998; McDonough, 2007; McDonough & Mackey, 2006), arguably because repair following a recast may not involve any additional processing other than mechanical imitation. Because the absence of repair following recasts in the present study cannot alone explain the greater number of within-group gains of the prompt group relative to the recast group, the input features generally associated with recast effectiveness will be considered next.

CF Saliency

Saliency is considered to be a primary factor in determining the effectiveness of recasts. The saliency of recasts can be enhanced by pragmatic variables relating to the research context, with recasts provided consistently by a researcher to a learner in a laboratory setting considered as more salient than recasts provided by a teacher in a classroom setting (e.g., Lyster & Izquierdo, 2009; Nicholas et al., 2001). The saliency of recasts can also be enhanced by input features relating to shortness in length, minimal changes, and intonational stress (Egi, 2007a; Ellis & Sheen, 2006; Loewen & Philp, 2006; Nicholas et al.; Sheen, 2006).

Audio recordings of the teacher-student interaction in the present study revealed that recasts were generally delivered by the teacher rather implicitly during oral production activities. Whereas recasts delivered during the question-and-answer activity often involved only one error, recasts provided during the dictogloss tasks often involved multiple changes with no stress or specific emphasis on the error. In (9), the learner made multiple errors in her original utterance by omitting a verb and failing to use the past tense for the two verbs. The teacher corrected the whole utterance by adding the missing verb, supplying the missing adverbial phrase of place *at the palace*, and correcting the errors on the use of past tense. All of these alterations occurred in a single move and may not have drawn the learners' attention exclusively to the target forms.

- (9) Student: *When Cinderella at the ball, the prince set eyes on Cinderella and walk to her, ask her to dance.*

Teacher: *When Cinderella **entered the ballroom** at the palace, the prince set eyes on Cinderella. Walking over to her, he bowed deeply and **asked** her to dance.*

As pointed out in some studies, the intended target features in long recasts with multiple errors are especially difficult for learners to perceive (Loewen & Philp, 2006; Philp, 2003; Sheen, 2006).

Additionally, the teacher of the recast group occasionally repeated students' error-free utterances, as in (10), in a way that discursively resembled the use of recasts.

- (10) Student: *We turned out everything in the basket.* (Correct utterance)

Teacher: *We turned out everything in the basket.* (Repetition)

Recasts in the present study, therefore, may have shared some traits with those that appear in communicative and immersion classrooms, where "recasts of ill-formed utterances and repetitions of well-formed utterances together appear to confirm or disconfirm the meaning of a learner's message" (Lyster, 2004a, p. 404).

In contrast, the corrective purpose of most prompts is generally more salient and, thus, their corrective function is easier for learners to notice in classroom settings. By definition, prompts indicate that something is wrong in learner utterances while withholding correct forms, thus providing negative evidence and inciting learners to modify their output. In some prompts, the learners were provided with metalinguistic clues (e.g., *What tense should we use here?*). Metalinguistic clues help learners to identify the nature and locus of the errors, especially errors for which they already have metalinguistic knowledge (Ammar & Spada, 2006). Yet, in spite of the greater saliency of prompts, surprisingly few participants, in response to items on the exit questionnaire, indicated that they remembered having their attention drawn specifically to grammar or past tense during the oral production activities, which suggests that prompts can be integrated into such classroom activities just as smoothly as recasts, but with greater effects.

Structural Targets

Degrees of saliency pertain not only to different CF types but also to different structural targets. Regular past tense forms are generally associated with low saliency (owing at times to the addition of a voiceless *-ed*), low communicative value, and high regularity (DeKeyser, 1998; Ellis, 2005), whereas irregular past tense forms are associated with high

saliency and high frequency but low regularity (Salaberry, 2000). In Mackey's (2006) classroom study that investigated the relationship between feedback and noticing, question formation was the most noticeable of the three different grammatical structures examined, whereas regular past tense forms were the least noticeable. Because recasts in the present study were delivered in an implicit manner during production activities, with neither intonational stress nor isolation of the error, recasts of irregular past tense forms were likely more noticeable than recasts of regular forms, due to the greater saliency of the irregular forms. Consequently, the recast group made more significant gains in irregular past tense forms (i.e., short- and long-term gains in written production and short-term gains in oral production) relative to regular forms (i.e., only short-term gains in written production and none in oral production). Prompts, however, succeeded in drawing attention to irregular and regular past tense forms alike and significantly increased learners' accuracy scores for both structures with durable effects over time.

Regular past tense forms are more amenable to prompts than recasts, because of the negative evidence generally afforded by prompts, but not by recasts that incorporate a nonsalient morpheme. In contrast, irregular past tense forms are amenable to both recasts and prompts, again because of the negative evidence afforded by prompts and by recasts that incorporate salient target features. It could be that learners' control over rule-based forms improves more from negative evidence and opportunities for self-repair than from positive evidence alone, whereas improved control over item-based forms results from either (a) positive evidence or (b) negative evidence together with self-repair. Drawing on Skehan's (1998) model of a dual-mode system, it may be argued that because the generative and compact rule-based system is otherwise difficult to access during online processing, learners benefit from the consolidating practice effects of a teacher's prompt to retrieve the regular past tense rule during online communication and to apply this rule in actual production. In contrast, because items stored in the less structured exemplar-based system can usually be retrieved quickly and with fewer processing constraints and no internal computation, learners might improve their control over irregular past tense forms already stored in the exemplar-based system simply by hearing them in the input.

CONCLUSION

The prompt group made significant gains on all oral and written measures, both in the short and long term, and significantly outperformed the control group on one measure. The overall superior performance of the

prompt group relative to the recast and control groups can be attributed to the self-repair that prompts consistently elicited as well as to their greater saliency during oral production activities. These results support findings from previous studies that demonstrate that prompts are generally more effective than recasts in classroom settings (e.g., Ammar & Spada, 2006; Ellis et al., 2006; Lyster, 2004a; see also the meta-analysis by Lyster & Saito, this issue).

Given the design of the current study and its focus on both irregular and regular past tense, the results presented here add to the emerging body of research that investigates the differential effects of feedback on different linguistic targets. Specifically, the effects of prompts were larger than those of recasts for increasing accuracy in the use of regular past tense forms, whereas prompts and recasts had similar effects on improving accuracy in the use irregular past tense forms. Given the nature of the dual-mode representational system (Skehan, 1998), it seems that during online communication, prompts more than recasts trigger access to the rule-based system, whereas recasts and prompts alike trigger access to the exemplar-based system. That recasts and prompts differentially affect rule-based and exemplar-based targets in this way is worthy of further investigation.

In conclusion, some limitations to this study should be acknowledged. First, each treatment group was limited to only one intact class, thus creating a possible teacher effect above and beyond CF type. An increased number of classes with comparable participants and teachers would yield more reliable and more robust results in future studies. Second, a practice effect might have been shown in the written test, because it included only a limited number of target items. In future studies, increasing the number of test items in both oral and written tests would certainly yield more robust results. Third, only recasts and prompts were compared without exploring the specific effects of different types of recasts (e.g., implicit vs. explicit) or the differential effects of specific prompts that range broadly from simple clarification requests to meta-linguistic clues. Future research should compare the differential effects of the various moves that comprise prompts, relative to each other and to different types of recasts, on a range of linguistic features and on classroom learners distinguished by individual differences such as proficiency level (Ammar & Spada, 2006), language analytic ability and aptitude (Sheen, 2007), and working memory (Trofimovich et al., 2007). These are all worthwhile areas for further investigation.

REFERENCES

- Ammar, A., & Spada, N. (2006). One size fits all? Recasts, prompts, and L2 learning. *Studies in Second Language Acquisition*, 28, 543–574.
- Cai, J. (2007). *The effects of multiple linguistic factors on the simple past use in English inter-language*. Beijing: Foreign Language Teaching and Research Press.

- Carpenter, H., Jeon, K. S., MacGregor, D., & Mackey, A. (2006). Learners' interpretations of recasts. *Studies in Second Language Acquisition*, 28, 209–236.
- Carroll, S. (2001). *Input and evidence: The raw material of second language acquisition*. Amsterdam: Benjamins.
- Carroll, S., & Swain, M. (1993). Explicit and implicit negative feedback: An empirical study of the learning of linguistic generalizations. *Studies in Second Language Acquisition*, 15, 357–366.
- Day, E., & Shapson, S. (2001). Integrating formal and functional approaches to language teaching in French immersion: An experimental study. *Language Learning*, 51, 47–80.
- de Bot, K. (1996). The psycholinguistics of the output hypothesis. *Language Learning*, 46, 529–555.
- DeKeyser, R. (1995). Learning second language grammar rules: An experiment with a miniature linguistic system. *Studies in Second Language Acquisition*, 17, 379–410.
- DeKeyser, R. (1998). Beyond focus on form: Cognitive perspectives on learning and practicing second language grammar. In C. J. Doughty & J. Williams (Eds.), *Focus on form in classroom second language acquisition* (pp. 42–63). New York: Cambridge University Press.
- Egi, T. (2007a). Interpreting recasts as linguistic evidence: The roles of linguistic target, length, and degree of change. *Studies in Second Language Acquisition*, 29, 511–537.
- Egi, T. (2007b). Recasts, learners' interpretations, and L2 development. In A. Mackey (Ed.), *Conversational interaction in second language acquisition: A collection of empirical studies* (pp. 249–267). Oxford: Oxford University Press.
- Ellis, R. (2005). Measuring implicit and explicit knowledge of second language: A psychometric study. *Studies in Second Language Acquisition*, 27, 141–172.
- Ellis, R. (2006). Researching the effects of form-focussed instruction on L2 acquisition. *AILA Review*, 19, 18–41.
- Ellis, R. (2007). The differential effects of corrective feedback on two grammatical structures. In A. Mackey (Ed.), *Conversational interaction in second language acquisition: A collection of empirical studies* (pp. 339–360). Oxford: Oxford University Press.
- Ellis, R., Basturkmen, H., & Loewen, S. (2001). Learner uptake in communicative ESL lessons. *Language Learning*, 51, 281–318.
- Ellis, R., Loewen, S., & Erlam, R. (2006). Implicit and explicit corrective feedback and the acquisition of L2 grammar. *Studies in Second Language Acquisition*, 28, 339–368.
- Ellis, R., & Sheen, Y. (2006). Re-examining the role of recasts in L2 acquisition. *Studies in Second Language Acquisition*, 28, 575–600.
- Erlam, R. (2006). Elicited imitation as a measure of L2 implicit knowledge: An empirical validation study. *Applied Linguistics*, 27, 464–491.
- Han, Z. (2002). A study of the impact of recasts on tense consistency in L2 output. *TESOL Quarterly*, 36, 542–572.
- Han, Z. (2008). On the role of meaning in focus on form. In Z. Han (Ed.), *Understanding second language process* (pp. 45–79). Clevedon, UK: Multilingual Matters.
- Harley, B., & Swain, M. (1984). The interlanguage of immersion students and its implications for second language teaching. In A. Davies, C. Cripser, & A. Howatt (Eds.), *Interlanguage* (pp. 291–311). Edinburgh: Edinburgh University Press.
- Hu, G. (2003). English language teaching in China: Regional differences and contributing factors. *Journal of Multilingual and Multicultural Development*, 24, 290–318.
- Izumi, S. (2003). Comprehension and production processes in second language learning: In search of the psycholinguistic rationale of the output hypothesis. *Applied Linguistics*, 24, 168–196.
- Jerome, J. K. (1889). *Three men in a boat (to say nothing of the dog)*. Bristol, UK: Arrowsmith.
- Leeman, J. (2003). Recasts and second language development: Beyond negative evidence. *Studies in Second Language Acquisition*, 25, 37–63.
- Levelt, W. (1989). *Speaking: From intention to articulation*. Cambridge, MA: MIT Press.
- Lightbown, P. M. (2008). Transfer appropriate processing as a model for class second language acquisition. In Z. Han (Ed.), *Understanding second language process* (pp. 27–44). Clevedon, UK: Multilingual Matters.

- Lightbown, P. M., & Spada, N. (1990). Focus-on-form and corrective feedback in communicative language teaching: Effects on second language learning. *Studies in Second Language Acquisition*, 12, 429–448.
- 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: Oxford University Press.
- Loewen, S., & Philp, J. (2006). Recasts in the adult L2 classroom: Characteristics, explicitness and effectiveness. *Modern Language Journal*, 90, 536–556.
- 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: Vol. 2. Second language acquisition* (pp. 413–468). San Diego, CA: Academic Press.
- Long, M. H., & Robinson, P. (1998). Focus on form: Theory, research, and practice. In C. J. Doughty & J. Williams (Eds.), *Focus on form in classroom second language acquisition* (pp. 15–41). New York: Cambridge University Press.
- Lyster, R. (1998). Recasts, repetition, and ambiguity in L2 classroom discourse. *Studies in Second Language Acquisition*, 20, 51–81.
- Lyster, R. (2002). Negotiation in immersion teacher-student interaction. *International Journal of Educational Research*, 37, 237–253.
- Lyster, R. (2004a). Differential effects of prompts and recasts in form-focused instruction. *Studies in Second Language Acquisition*, 26, 399–432.
- Lyster, R. (2004b). Research on form-focused instruction in immersion classrooms: Implications for theory and practice. *Journal of French Language Studies*, 14, 321–341.
- Lyster, R. (2007). *Learning and teaching languages through content: A counterbalanced approach*. Amsterdam: Benjamins.
- Lyster, R., & Izquierdo, J. (2009). Prompts versus recasts in dyadic interaction. *Language Learning*, 59, 453–498.
- Lyster, R., & Mori, H. (2006). Interactional feedback and instructional counterbalance. *Studies in Second Language Acquisition*, 28, 269–300.
- Lyster, R., & Mori, H. (2008). Instructional counterbalance in immersion pedagogy. In R. Fortune & D. Tedick (Eds.), *Pathways to bilingualism and multilingualism: Evolving perspectives on immersion education* (pp. 133–151). Clevedon, UK: Multilingual Matters.
- Lyster, R., & Ranta, L. (1997). Corrective feedback and learner uptake: Negotiation of form in communicative classrooms. *Studies in Second Language Acquisition*, 19, 37–66.
- Mackey, A. (2006). Feedback, noticing, and second language development: An empirical study of L2 classroom interaction. *Applied Linguistics*, 27, 405–430.
- Mackey, A., & Gass, S. M. (2005). *Second language research: Methodology and design*. Mahwah, NJ: Erlbaum.
- Mackey, A., Gass, S. M., & McDonough, K. (2000). How do learners perceive interactional feedback? *Studies in Second Language Acquisition*, 22, 471–497.
- 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: Oxford University Press.
- Mackey, A., & Philp, J. (1998). Conversational interaction and second language development: Recasts, responses, and red herrings? *Modern Language Journal*, 82, 338–356.
- McDonough, K. (2007). Interactional feedback and the emergence of simple past activity verbs in L2 English. In A. Mackey (Ed.), *Conversational interaction in second language acquisition: A collection of empirical studies* (pp. 323–338). Oxford: Oxford University Press.
- McDonough, K., & Mackey, A. (2006). Responses to recasts: Repetitions, primed production, and linguistic development. *Language Learning*, 56, 693–720.
- Nicholas, H., Lightbown, P. M., & Spada, N. (2001). Recasts as feedback to language learners. *Language Learning*, 51, 719–758.
- Norris, J., & Ortega, L. (2000). Effectiveness of L2 instruction: A research synthesis and quantitative meta-analysis. *Language Learning*, 50, 417–528.
- Norris, J., & Ortega, L. (2006). The value and practice of research synthesis for language learning and teaching. In J. Norris & L. Ortega (Eds.), *Synthesizing research on language learning and teaching* (pp. 3–52). Amsterdam: Benjamins.

- Perrault, C. (2004). Cinderella. In A. E. Johnson (Trans.), *Perrault's fairy tales* (pp. 56–65). Hartfordshire, UK: Wordsworth. (Original work published 1697)
- Philp, J. (2003). Constraints on “noticing the gap”: Non-native speakers’ noticing of recasts in NS-NNS interaction. *Studies in Second Language Acquisition*, 25, 99–126.
- Pinker, S., & Prince, A. (1994). Regular and irregular morphology and the psychological status of rules of grammar. In S. Lima, R. Corrigan, & G. Iverson (Eds.), *The reality of linguistic rules* (pp. 353–388). Amsterdam: Benjamins.
- Pinker, S., & Ullman, M. (2002). The past and future of the past tense. *Trends in Cognitive Science*, 6, 456–463.
- Ranta, L., & Lyster, R. (2007). A cognitive approach to improving immersion students’ oral language abilities: The awareness-practice-feedback sequence. In R. DeKeyser (Ed.), *Practice in a second language: Perspectives from applied linguistics and cognitive psychology* (pp. 141–160). New York: Cambridge University Press.
- Robinson, P., & Ha, M. (1993). Instance theory and second language rule learning under explicit conditions. *Studies in Second Language Acquisition*, 13, 413–438.
- Rumelhart, D. E., & McClelland, J. L. (1986). On learning the past tenses of English verbs. In J. L. McClelland, D. E. Rumelhart, & the P.D.P. Research Group (Eds.), *Parallel distributed processing: Explorations in the microstructures of cognition: Vol. 2. Psychological and biological models* (pp. 216–271). Cambridge, MA: MIT Press.
- Salaberry, R. (2000). *The development of past tense morphology in L2 Spanish*. Amsterdam: Benjamins.
- Schmidt, R. (2001). Attention. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 3–32). New York: Cambridge University Press.
- Sheen, Y. (2004). Corrective feedback and learner uptake in communicative classrooms across instructional settings. *Language Teaching Research*, 8, 263–300.
- Sheen, Y. (2006). Exploring the relationship between characteristics of recasts and learner uptake. *Language Teaching Research*, 10, 361–392.
- 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: Oxford University Press.
- Skehan, P. (1998). *A cognitive approach to language learning*. Oxford: Oxford University Press.
- Sokal, R. R., & Rohlf, F. J. (1995). *Biometry: The principles and practice of statistics in biological research* (3rd ed.). New York: Freeman.
- Spada, N. (1997). Form-focussed instruction and second language acquisition: A review of classroom and laboratory research. *Language Teaching*, 30, 73–87.
- Swain, M. (1993). The output hypothesis: Just speaking and writing aren’t enough. *Canadian Modern Language Review*, 50, 158–164.
- Swain, M. (1995). Three functions of output in second language learning. In G. Cook & B. Seidlhofer (Eds.), *Principles and practice in applied linguistics: Studies in honour of H. G. Widdowson* (pp. 125–144). Oxford: Oxford University Press.
- Swain, M., & Lapkin, S. (2001). Focus on form through collaborative dialogue: Exploring task effects. In M. Bygate, P. Skehan, & M. Swain (Eds.), *Researching pedagogic tasks: Second language learning, teaching, and testing* (pp. 99–118). London: Longman.
- Trofimovich, P., Ammar, A., & Gatbonton, E. (2007). How effective are recasts? The role of attention, memory, and analytic ability. In A. Mackey (Ed.), *Conversational interaction in second language acquisition: A collection of empirical studies* (pp. 171–195). Oxford: Oxford University Press.
- Wajnryb, R. (1990). *Grammar dictation*. Oxford: Oxford University Press.
- Westfall, P. H., Tobias, R. D., Rom, D., Wolfinger, R. D., & Hochberg, Y. (1999). *Multiple comparisons and multiple tests using the SAS system*. Cary, NC: SAS Institute.

APPENDIX

EXIT QUESTIONNAIRE

1. Now that you have completed the tasks and the tests, what do you think they were all about?
 - a) They were practicing and testing writing.
 - b) They were practicing and testing my grammar.
 - c) They were practicing and testing listening and speaking.
 - d) They were practicing and testing my vocabulary.
2. Please write a short paragraph saying what you think you learned from these communicative tasks and tests.
3. Do you like the tasks you performed in class? Why do you like or dislike them?
4. Do you like to be corrected while you are speaking in class? Why or why not?