

The role of recasts and negotiated prompts in an FL learning context in China with non-English major university students

Language Teaching Research

2021, Vol. 25(2) 209–233

© The Author(s) 2019

Article reuse guidelines:

sagepub.com/journals-permissions

DOI: 10.1177/1362168819839727

journals.sagepub.com/home/ltr**Huifang (Lydia) Li**

University of New England, Armidale, New South Wales, Australia

Noriko Iwashita 

University of Queensland, Brisbane, Queensland, Australia

Abstract

This quasi-experimental classroom study examines the effects of recasts and negotiated prompts on oral interactions in a foreign language (FL) context where the instructional approach primarily focuses on grammar and accuracy. Ninety adult native speakers of Mandarin Chinese from two intact classes were randomly assigned to one of two experimental groups (recast or negotiated prompt) or a control group, and were asked to complete pre-/post-tests on their use of question and past tense forms in English. Between the pre-test and the post-tests, the learners of the two experimental groups had three treatment lessons for feedback outside their usual English classes, the interaction data of which are reported. ANOVA analyses revealed that recasts were highly effective for accuracy development of *wh*-questions, interrogative questions and irregular past tense verbs. Negotiated prompts showed moderate effectiveness at most times for accuracy development of interrogative questions and both irregular and regular past tense verbs. These findings provide further insights into the role of corrective feedback (CF) in L2 development.

Keywords

corrective feedback, recasts, negotiated prompts, questions, past tense, second language acquisition

Corresponding author:

Huifang (Lydia) Li, University of New England, LaTT, Armidale, 2350, NSW, Australia.

Email: hli24@une.edu.au

I Introduction

There has been considerable interest in investigating the role of two types of corrective feedback (CF), namely, recasts and prompts, in second language acquisition (SLA) research. Many studies have reported a positive role of these two CF types in second language (L2) development by comparing their effects on different structures within laboratory and classroom settings (Lyster & Saito, 2010; Lyster, Saito, & Sato, 2013; Russell & Spada, 2006). While both CF types are similarly effective for L2 learning in the laboratory setting (Li, 2010; Lyster & Izquierdo, 2009; Nicholas, Lightbown, & Spada, 2001), their role in L2 learning in the classroom setting remains unclear (Goo & Mackey, 2013; Lyster & Ranta, 2013; Lyster et al., 2013). This lack of clarity might result from the fact that a combination of variables in a classroom context (e.g. learner age) interacts with the CF types and will have an impact on their effectiveness. As a result, we suggest, as do others (Lyster & Ranta, 2013; Lyster et al., 2013), that further classroom research is needed to obtain a better understanding of how the two CF types work for L2 learning in the classroom setting, and to offer implications for L2 pedagogies.

This article reports on a quasi-experimental classroom study that investigated the effectiveness of recasts and prompts in the context of English as an foreign language (FL). Specifically, this context involved non-English major students at a Chinese university. To our knowledge, little research has investigated how this group of learners utilizes CF for accuracy development in their English classes. Given that non-English major students constitute a large majority of the annual five million enrollees in Chinese higher education (Chen & Klenowski, 2009; Liu, Xu, & Ran, 2015), it is of significance to examine their L2 learning of English, including their use of CF when they are engaged in language tasks. Moreover, built on a few experimental classroom studies that targeted two structures (e.g. R. Ellis, 2007; Mifka-Profozic, 2013; van de Gauche, Braaksma, Rijlaarsdam, & Bimmel, 2015; Yang & Lyster, 2010), this study further tested the differential effects of recasts and prompts on two types of questions (wh-questions and interrogative questions) and two types of past tense verbs (irregular and regular past tense verbs). Finally, in contrast with previous experimental classroom research (e.g. Ammar & Spada, 2006; R. Ellis, 2007; Gooch, Saito, & Lyster, 2016; Lyster, 2004; Mifka-Profozic, 2013; van de Guchte et al., 2015; Yang & Lyster, 2010), this study used a new combination of prompts to determine their effectiveness in relation to recasts. This combination includes three prompts that provide opportunities for output only: clarification requests, repetitions, and elicitations. We refer to the three prompts as negotiated prompts. Through these variations in method, this study aims to cast further light on how recasts and prompts work for L2 learning in the classroom setting.

I Theoretical background

Recasts are the most widely studied CF in oral interaction research. They constitute partial or full reformulation of L2 learners' non-target-like utterances with target-like models while keeping the learners' central meaning. Theoretical support for the contribution of recasts to SLA draws largely on Long's (1996, 2007) updated interaction hypothesis,

which refers to Schmidt (1990, 1993, 1995) to highlight the role of noticing in L2 learning. Long believes that the utility of recasts lies in the feedback's semantic transparency in freeing up learner attention to form, contingent juxtaposition of incorrect and correct utterances to induce a cognitive comparison of form, and unobtrusiveness in conversation flow to encourage form–function mapping. Some other scholars (e.g. Doughty, 2001; Nicholas et al., 2001) add that recasts can also be a scaffolding technique when the targeted forms are beyond learners' interlanguage capacity to produce.

Prompts are often compared to recasts in SLA research. Whereas recasts provide models of language and are referred to as input-providing CF, prompts encourage L2 learners to modify their non-target-like utterances and are referred to as output-promoting CF (R. Ellis, 2006). According to Lyster and his colleagues (Lyster, 2004; Lyster et al., 2013), prompts contain clarification requests, repetitions, elicitations, and metalinguistic clues. Theoretical support for the contribution of prompts to SLA can be found in the output hypothesis (Swain, 1985, 1995, 2005) and skill acquisition theory (Anderson, 1993, 2000). According to the output hypothesis, modified or pushed output following prompts helps learners to notice their deviant L2 use, to search in their rule-based knowledge system to test their hypotheses, and to talk about language forms. These functions of modified or pushed output induce a processing of language form at the conscious level, which Lyster (2004) argues will make prompts more informative for resolving problematic interlanguage forms. With reference to skill acquisition theory, prompts are postulated to provide 'guided practice in the context of communicative interaction' (Lyster et al., 2013, p. 11), which contributes to the gradual transition from effortful production of the target language to more automatic use.

Drawing on the theories explained above, many studies have examined the effectiveness of recasts and prompts. The existing observational and experimental classroom studies seem to point to recasts being overall less effective than prompts. However, as many factors can affect CF effectiveness, the results are not conclusive (e.g. R. Ellis, 2007; Goo & Mackey, 2013; Lyster & Ranta, 2013; Lyster et al., 2013; Russell & Spada, 2006). In the following sections, relevant studies will be reviewed, with a focus on the factors that have motivated the current study.

2 Classroom context

While the effects of recasts and prompts differ between laboratories and classrooms (Li, 2010; Lyster & Izquierdo, 2009), the effects of the two CF types may also differ across classroom contexts. In a study that compared the rate of uptake following CF across two Canadian classrooms, English L2 classrooms in New Zealand, and English FL classrooms in Korea, Sheen (2004) found that the rate of uptake following recasts differed substantially. Comparably, Lyster and Mori (2006) reported that a much higher proportion of uptake followed recasts in Japanese immersion classes in the USA, where a focus-on-form was emphasized with repetition and reading aloud exercises, than in French immersion classes in Canada, where the instruction was mainly focused on meaning or content. Based on the differential rates of uptake in different classrooms, researchers (Han & Kim, 2008; Loewen & Philp, 2006; Nicholas et al., 2001; Sheen, 2004, 2006) speculate that recasts should work as well as prompts in a classroom context that is more

oriented to language form because this orientation reduces the ambiguity of recasts. Conversely, when the instructional context is mainly focused on meaning or content, recasts can be ambiguous as confirming meaning rather than disapproving form (Lyster, 1998, 2015).

Although observational classroom studies are invaluable in showing the role of recasts and prompts according to classroom contexts, these studies can be criticized for using uptake as a measure of learning. Uptake is not equivalent to evidence of learning (Goo & Mackey, 2013; Lyster et al., 2013). Therefore, experimental classroom studies (e.g. Ammar & Spada, 2006; R. Ellis, 2007; Gooch et al., 2016; Loewen & Nabei, 2007; Lyster, 2004; Mifka-Profozic, 2013; van de Guchte et al., 2015; Yang & Lyster, 2010) have used pre-/post-tests to measure the effects of recasts in comparison to a single prompt or a combined category of prompts on different structures (see Table 1). For example, in an early but influential study, Lyster (2004) examined the effects of recasts vs. four prompts in form-focused instruction on the learning of French gender agreement. Results indicate that subjects who received recasts embedded in form-focused instruction were superior in five of the eight measures when compared with the control group, but subjects who received prompts with form-focused instruction significantly out-performed the control group in all measures.

Collectively, the quasi-experimental classroom studies have tended to show less effectiveness of recasts than prompts. However, the studies offer a limited insight into how CF effectiveness differs according to classroom contexts. Experimental classroom studies are limited in number, less than 10 to our knowledge, and some studies only tested CF effectiveness without reporting the classroom context (Gooch et al., 2016; Loewen & Nabei, 2007). Among the studies that reported the classroom context, the greater effectiveness of prompts was achieved mainly in classroom contexts where the L2 was used as a tool for communication, such as an intensive language program or bilingual content-based program (Ammar & Spada, 2006; R. Ellis, 2007; Lyster, 2004; Yang & Lyster, 2010). There seems to be limited research in the classroom context that emphasizes grammar and accuracy, which is typically associated with the formal language classroom in the FL setting (Collentine & Freed, 2004). Only two studies, van de Guchte et al. (2015) and Mifka-Profozic (2013), were conducted in this latter context. Although Yang and Lyster (2010) claimed their study was conducted in a context focusing on language form, the learners in their study were English-major second year students at a Chinese university. Those students receive various language-related classes in the instructional language of English at university (Fang, 2009; Syllabus for the English Major in Higher Education, 2000), and their learning environment provides an extensive amount of input but may lack the opportunity for output.

Compared to Yang and Lyster's (2010) English-major second year students, non-English major university students in China undertake weekly classes of three to four hours in their first two years of study. As Cortazzi and Jin (1996a, 1996b, 2001) describe, the lessons usually start with an explanation of new words or phrases (i.e. parts of speech, collocations), continue with the teacher's explanation of the text sentence-by-sentence, and end with written exercises that mostly focus on language points. Also noted by Jin and Cortazzi (1998) and observed by us, due to waves of top-down reforms to promote communicative or task-based language teaching in China, language tasks are now infrequently

Table 1. Classroom experimental studies: Study design and contextual variables.

Context	Class size	Treatment (duration, type)	Prompt type	Target structure(s) and mode of testing	Frequency of corrective feedback (CF)
Lyster (2004)	bilingual French program; 10–11-year-olds, 3 schools, Canada	20–25 8–10 hours over 5 weeks; FFI + recasts or prompts, FFI only and control	elicitation, repetition, CR, MC	grammatical gender; written and oral production	not known
Ammar & Spada (2006)	intensive English class; Canada; $n = 64$, 3 teachers	not specified 30–35 minutes \times 12 over 4 weeks; recast, prompt and control	elicitation, repetition, MC	third person possessive; oral; listening	not known
R. Ellis (2007)	general English class; private language school; adults; New Zealand	10–12 1 hour over 2 days; recast, MC, and control	MC	past tense verbs; comparative	recasts 42 (past), 22 (comparative); MC 25 (past), 19 (comparative)
Loewen & Nabei (2007)	general English class at university, adults, Japan	4 in each group 30 minutes (recast, CR, MC, and no feedback)	CR and MC	question forms; GJT, oral production tasks	recasts 18.7; CR 18; MC 5.6
Yang & Lyster (2010)	general English class at university, English major students, adults, China	22–25 2 hours over 2 weeks; recast, prompt and control group	MC, CR, elicitation, repetition	regular and irregular past tense verbs; oral and written production	prompt 27; recasts 20
Mifka-Profozic (2013)	French, Croatia, 2 secondary schools Year 11–12	18 each 20–25 minutes per treatment over 3 weeks; recast, CR, and control	CR	passé composé, imperfait; oral and written production	CR 40; recasts 68
van de Guchte et al. (2015)	14-year-old (9th grade), German, $n = 64$; secondary school in the Netherlands	20–23 per class 1 hour per week over 3 weeks; recast, prompt and control	MC followed by elicitation	dative and comparative; written and oral production	not known
Gooch et al. (2016)	English learners at a Korean university	6–9 4 hours over 2 weeks FFI + recasts or prompts, FFI only	CR followed by elicitation	/ɹ/ pronunciation	recasts (68) prompts (97)

Note. CR = clarification request. FFI = form-focused instruction. GJT = grammaticality judgment test. MC = metalinguistic clues.

used in the English classes of non-English major students. In this classroom context, where there is a strong focus on grammar and accuracy, the provision of recasts during the performance of language tasks may facilitate learning equally as well as the provision of prompts. This prediction, however, would need to be supported by empirical data.

3 Target structures

Research has found that CF works for one structure in one condition but may not work for another structure. This structure-dependent effect of CF is reflected in the findings of four experimental classroom studies that targeted two different structures (R. Ellis, 2007; Mifka-Profozic, 2013; van de Guchte et al., 2015; Yang & Lyster, 2010). With adult Asian learners from three small language classes that formed the recast, metalinguistic clue, or control groups, R. Ellis (2007) investigated the effects of recasts vs. metalinguistic clues on the comparative and regular past tense in English. He rationalized his selection of the two structures in terms of learning difficulty, with the comparative being the more difficult. R. Ellis found metalinguistic clues were more effective in the comparative than in the regular past tense, with nil effects of recasts on both structures. Also with adult Asian learners, Yang and Lyster (2010) compared the effects of recasts and prompts on regular and irregular past tense verbs based on the two forms' distinction in rule-based vs. exemplar-based learning. It was found that there were larger effects for the prompt group than for the recast group in the learners' use of regular past tense verbs, and the two groups had comparable effects with irregular past tense verbs. Therefore, Yang and Lyster concluded that prompts are more effective than recasts in triggering access to the rule-based system, whereas both CF types are alike in triggering access to the exemplar-based system.

Another two experimental classroom studies involved secondary school learners. Van de Guchte et al. (2015) examined the interaction of recasts and prompts with two structures that differed in relatedness to the students' mother tongue (L1). Results showed that while prompts were effective for Dutch students' acquisition of both comparative and dative structures in German, recasts yielded positive findings only in the relatively easier, L1-related comparative structure. On the other hand, Mifka-Profozic (2013) found recasts were more beneficial than clarification requests for English learners' development of *passé composé* and imperfect structures in French. Specifically, the learners in the recast group significantly improved their accurate use of both structures in all the tests, whereas those in the clarification request group showed gains over time only for the imperfect structure. Accordingly, Mifka-Profozic suggests that recasts work better than clarification requests for the acquisition of morphologically complex structures, though both CF types can facilitate the acquisition of morphologically simple structures.

It can be seen that the number of experimental classroom studies that have investigated the efficacy of recasts and prompts on the acquisition of different structures is limited. The question raised by Yang and Lyster (2010) about how recasts and prompts work differently for different structures has been inconclusively answered. A number of studies have chosen English questions (e.g. Mackey & Philp, 1998; McDonough, 2005) or past tense verbs (e.g. R. Ellis, 2007; Yang & Lyster, 2010) to examine the effects of

CF on L2 development. However, no study has compared the effects of CF on questions and the past tense in the same study. Robinson (2001) discusses whether the learning of syntax may differ from the learning of morphology or lexicon regarding the level of attention required. A study that investigates the effects of recasts and prompts on both syntactic questions and morphological past tense will develop further understanding of how the two CF types work for L2 development.

4 Feedback operationalization

Another factor that leads to the lack of clarity in the role of recasts and prompts in previous experimental classroom studies is the varied administration of the two CF types. As shown in Table 1, prompts come in a variety of categories, involving only metalinguistic clues in R. Ellis (2007), a clarification request in Mifka-Profozic (2013), a metalinguistic clue followed by an elicitation in van de Guchte et al. (2015), a clarification request followed by an elicitation in Gooch et al. (2016), and metalinguistic clues with two or three other prompts in the rest of the studies (e.g. Ammar & Spada, 2006; Lyster, 2004; Yang & Lyster, 2010). The operationalization of metalinguistic clues in the studies also varies. In Lyster (2004), metalinguistic clues did not refer to grammar rules about the non-target-like parts of learner utterances; however, the other studies that used this CF did refer to grammar rules. When metalinguistic information is provided via metalinguistic clues, it is not clear whether the effects of prompts are due to metalinguistic input, opportunities for output or both (Goo & Mackey, 2013). Therefore, to clearly determine the role of prompts, future research on CF would need to compare the effects of recasts with prompts that provide opportunities for output only or with metalinguistic clues.

The same variation of administration occurs with recasts. While three studies (R. Ellis, 2007; Gooch et al., 2016; Mifka-Profozic, 2013) were specific in operationalizing recasts, other experimental classroom studies were not. In the other studies, recasts could come in a variety of forms, and the relative efficacy of these forms may be significantly different (Loewen & Philp, 2006; Sheen, 2006). This variation in the operationalization of recasts further complicates the results of experimental classroom research. Much remains unknown; for instance, what kind of recasts work more or less effectively than what kind of prompts, and to what extent do prompts contribute to learning due to opportunities for output? Researchers (Lyster & Ranta, 2013; Lyster et al., 2013) then argue that CF in the form of treatment should be documented so that the information obtained can help to interpret any differing results.

To summarize, within the prolific research that has examined the effects of recasts and prompts on L2 development, there are still issues warranting further investigation. Such issues include the relative effectiveness of recasts and prompts in terms of different classroom contexts and target structures, and whether the effects of prompts are due to output, metalinguistic input, or both. We believe that the claim made by Russell and Spada (2006) a decade ago is still applicable to the current research on CF effectiveness:

Without a sufficient accumulation of studies on any one of these variables and without researchers' attention to the constellation of moderating variables that could make a difference regarding CF effectiveness, we will not be able to establish clear patterns across studies. (p. 32)

II Research questions

In order to enhance further understanding of CF effectiveness in L2 development, this study investigated the effects of recasts vs. negotiated prompts on the non-English major university students' acquisition of English question and past tense forms in oral interaction. Four research questions were raised:

- Research question 1: To what extent will recasts and negotiated prompts facilitate accurate use of *wh*-questions, and which type of feedback shows a larger effect?
- Research question 2: To what extent will recasts and negotiated prompts facilitate accurate use of interrogative questions, and which type of feedback shows a larger effect?
- Research question 3: To what extent will recasts and negotiated prompts facilitate accurate use of irregular past tense verbs, and which type of feedback shows a larger effect?
- Research question 4: To what extent will recasts and negotiated prompts facilitate accurate use of regular past tense verbs, and which type of feedback shows a larger effect?

III Method

1 Student participants and their L2 learning environment

The first-year students from two intact English classes in the physics department at a provincial university in China were invited to participate in the study during the first author's visit to their classes. Ninety students volunteered to take part in the study, of which 72 were females and 18 were males. They all had Mandarin as their mother tongue and were between 17 and 18 years old. They had studied English for six years at secondary school. Their English classes in secondary school were large and delivered with traditional teacher-centered pedagogical practices. At the university, their learning situation did not seem to change much, remaining teacher-centered and grammar-focused but involving the occasional use of language tasks.

2 The teacher participant and training

The teacher participant was a native speaker of English from New Zealand and had a Master of Arts degree. He had taught oral English in the School of Foreign Languages at the university for a number of years before the data collection and was willing to assist throughout the study.

The teacher was instructed to use the required form of correction (recasts with the recast group and negotiated prompts with the prompt group) to correct the learners' non-target-like utterances of the target structures. For the provision of recasts, he could negotiate for meaning with a learner if the learner's utterance was incomprehensible to him. For the provision of negotiated prompts, he could provide a negotiated prompt of his choosing to address a learner's non-target-like utterance. If the learner could not provide

the correct utterance after being prompted several times, he was instructed to ask another learner to help with the correction. The teacher practiced providing the CF before the experiment commenced.

3 Design

The experimental design of the study contained a pre-test, three treatment lessons, and two post-tests, all of which were completed within four weeks. Participant assignment followed the stratified random sampling method. That is, the 90 learners were classified into three levels according to their scores on the National Matriculation English Test, and those at the same level were assigned into the recast, negotiated prompt, and control groups in a random manner. They were briefed about the study procedure on Day 1 and took the pre-test on Day 2. On Days 3–9, the experimental group learners participated in three treatment sessions outside their usual English classes; these sessions were video-recorded for the purpose of documenting the treatment. Immediate and delayed post-tests were administered on Days 10 and 31, respectively. While the control group learners also had their usual English classes, they took the three tests only and did not have any treatment lessons.

4 Target structures

The target structures in this study involved questions and the past tense in English. While the complexity of *wh*-questions and interrogative questions varies, questions and regular past tense verbs involve rule-based learning, and irregular past tense verbs involve exemplar-based learning. Given the variation, this study investigated the effects of recasts and negotiated prompts on two types of questions (i.e. *wh*-questions and *yes/no* questions) and two types of past tense verbs (i.e. irregular and regular past tense verbs).

The 90 learners were able to use the structures in their spontaneous communication with the explicit knowledge they had acquired in secondary schooling. A summary of the descriptive statistics showing the learners' attempts of the four structures in the three tests is provided in Online Appendix 1. However, their production of the structures included many non-target-like features. They randomly varied between the past and present tenses when a consistent use of the past tense was required, which is a common phenomenon in L2 production (Han, 2002). Similarly, variation existed in their production of varied question types regarding word order and accurate use of the moved elements when correct word order was followed (i.e. *wh*-words, auxiliaries/modals/copulas in relation to subjects or main verbs and subject cases). It is common to see a co-existence of the following questions in a learner's production: 'What *he is* drinking?', 'What *are* other people *doing*', 'Where is the umbrella?', 'Where *are* his right arm?', 'Is he *sit* the left of the bench or the right?' and 'Is the boy drinking water?' These patterns of variation in L2 learners' production of questions were also observed in other studies (e.g. Spada & Lightbown, 1993; Smith & Truscott, 2005).

5 Materials and administration for testing and treatment

The pre-/post-tests and the treatment lessons comprised various communicative activities to elicit the learners' oral production of the target structures. The pre-/post-tests were made up of three versions of oral production tasks (A, B, and C). Each version contained a warm-up, spot-the-difference, story-discovery, and storytelling tasks; the second and third tasks were to elicit the learners' use of questions and the fourth task was to determine their consistent use of past tense in narrative storytelling. To avoid any test order effects, the order of the three versions was counterbalanced, with a version being used by one third of the learners in each group and then rotated.

Pre-, immediate post-, and delayed post-test sessions were all conducted in three language labs, with an instructor in each lab to control the mainframe computer and to supervise the oral test. The first task was the warm-up, which took five minutes, during which the learners answered questions about their recent activities. The spot-the-difference task also took five minutes, during which the learners were instructed to ask questions to discover 20 differences between the picture on their individual monitor and a hidden picture. During the third task, story-discovery, four pictures were shown one-by-one on the monitor, and the learners were given one minute per picture to ask questions to discover the story. The last task was the storytelling. The four pictures used in the third task, along with another two pictures, were arranged sequentially on the monitor. The students then had one minute for preparation and subsequently had to tell a story, beginning with 'Last Sunday'. Most of them finished their story within five minutes.

Each of the three treatment lessons comprised 2×40 -minute sessions. In the first session, the learners completed a question task, and in the second, they were assigned a task on past tense verbs. The tasks included interviewing vs. sharing experiences, spot-the-differences vs. chain storytelling, and guessing objects vs. storytelling. The question tasks were carried out either by the teacher interacting with a group of students or by one group of students interacting with another group in front of the whole class. For the tasks on past tense verbs, the students were sometimes divided in small groups and were given a few minutes for preparation. Then, the tasks were carried out by one student speaking in front of the whole class to share/tell a story or to add a sentence building up the story. It should be noted that there was no grammar instruction on the target structures in the treatment lessons or in the learners' usual English classes, despite their usual English classes focusing on language forms.

The selection of task types used for the test sessions and the treatment lessons was based mainly on several previous studies on question and past tense forms (e.g. Han, 2002; Mackey, 1999; Mackey & Philp, 1998; Philp, 2003). All the tasks were piloted prior to the data collection and were found to elicit successful oral production of the target structures (see sample tasks in the Online Appendices 2 and 3).

6 Analyses and coding

The present study operationalized acquisition as the learners' accurate use of the target structures. The accuracy analyses for questions followed Spada and Lightbown (1993).

The accuracy rate of wh-questions a learner produced in a test was calculated by applying the following formula, and the same calculation was applied to calculating the accuracy rate of interrogative questions. A target-like question was defined as being not only correct in word order, but also correct in the use of the wh-question word, the auxiliary/modal/copula verb relative to the subject/the main verb, and the subject case. This method of analysis examined whether the CF treatment could help the learners to become more proficient in their use of question types.

$$\frac{n \text{ target-like Wh- questions}}{n \text{ total production of Wh- questions}} \times 100 = \text{Percentage frequency}$$

To code questions, certain question utterances were excluded: incomplete and formulaic questions (e.g. 'What's the weather like?'), statement questions with rising intonation (e.g. 'The boy is crying?'), and wh-subject questions (e.g. 'Who is near the door?'). For the remaining utterances of questions, each was labeled as a target-like wh-question or target-like interrogative question when it complied with the above-mentioned definition of being target-like.

The accuracy analyses for regular and irregular past tense verbs adopted target-like use analysis (TLU) (Pica, 1983, p. 33). For instance, to calculate the accuracy rate of irregular past tense verbs that a learner used in a test, we tallied the numbers for correct suppliance of irregular verbs in the obligatory contexts where the past tense was required (i.e. n correct suppliance in obligatory contexts), incorrect suppliance of irregular verbs in the obligatory contexts, and the over-suppliance of irregular verbs in the contexts where the past tense was not required (i.e. n suppliance in non-obligatory contexts). Then we added the first two numbers as n obligatory contexts and applied the three n numbers in the formula below. A similar calculation was undertaken separately for regular past tense verbs.

$$\frac{n \text{ correct suppliance in obligatory contexts}}{n \text{ obligatory contexts} + n \text{ suppliance in non-obligatory contexts}} \times 100 = \text{Percentage accuracy}$$

Three considerations were taken into account when coding past tense verbs. First, the verbs that could be ambiguous in the learners' oral production were excluded from the analyses. Because the learners in the study had poor pronunciation, it was difficult to differentiate between the present and past tenses of certain verbs in their oral production (such as the present tense form 'run' and the past tense form 'ran'). Second, when self-correction or repetition occurred, the learner's first production of the verb was counted as one obligatory context and rated. Finally, tokens of past tense were tallied; if a learner used the verb 'think' twice at different places in their story, then there were two obligatory occasions.

A trained research assistant double-coded 15% of the test data, with the inter-rater reliability being 97.7%. The inter-rater reliability was then calculated separately with 15% of the test data taken from each experimental group. The inter-rater reliability was

Table 2. Distribution of feedback episodes, feedback moves and learner responses.

Groups	Feedback episodes	Feedback moves		Learner responses
		Assigned feedback	Other feedback	
<i>Recasts:</i>				
Wh-	11	11	6	0
Interrogative	5	5	0	0
Irregular	11	11	0	4
Regular	5	5	0	1
<i>Negotiated prompts:</i>				
Wh-	11	22	2	29
Interrogative	5	9	1	9
Irregular	11	21	0	29
Regular	7	6	2	6

95.5% for the recast group, 99.5% for the negotiated prompt group and 98.2% for the control group.

IV Results

This section reports the preliminary and main analyses in examination of the research questions.

1 Preliminary analyses

a Interaction data. The interaction data that is presented below was from one treatment lesson randomly selected by drawing lots. The tasks involved in the selected lesson were guessing objects for questions and storytelling for past tense verbs. In the transcription, feedback episodes, feedback moves, and learner responses to feedback were identified. A feedback episode starts with a learner's non-target-like utterance of the target structures, which is then addressed by the assigned feedback type, and ends with a change back to topic focus or another linguistic form (Ellis, Basturkmen, & Loewen, 2001). Feedback moves indicate the type and number of feedback incidences provided in the identified episodes.

As summarized in Table 2, three findings are worthy of comment regarding the provision of recasts and negotiated prompts as treatment. First, the two experimental groups received a fairly even number of feedback episodes for the learners' non-target-like utterances of both questions and the past tense. For both groups, the majority of the CF episodes were given to wh-questions and irregular past tense.

Second, the recast group was provided with recasts and the negotiated prompt group received the vast majority of negotiated prompts, two thirds of which were elicitations. This provision of CF indicated the teacher's consistency in providing feedback according to the experimental conditions. Unsurprisingly, feedback types other than the assigned

feedback type occurred; that is, six clarification requests for wh-questions occurred in the recast group, and four recasts (two for wh-questions and two for regular past tense verbs) and one incidence of explicit feedback occurred in the negotiated prompt group. As mentioned earlier, the treatment was carried out in the L2 classroom, so pedagogical considerations were necessary, such as allowing meaning clarification when the teacher could not understand a student in the recast group and providing a resolution to the problematic utterance in the prompt group when no student could help with the correction.

A final note addresses the manner of delivery of the two CF types. Most of the recast episodes contained a single CF move and were provided in full reformulation of a learner's non-target-like utterance without stress or chance for repair, as shown in example 1. However, among the 16 recast episodes for the past tense, five learner responses were identified and they were all simple acknowledgments with the word 'yeah', as shown in example 2. Negotiated prompt feedback can also be short, like recasts (example 3). This short episode of negotiated prompt feedback occurred only a few times. However, most of the negotiated prompt episodes contained multiple negotiated prompt moves and learner responses, and sometimes a negotiated prompt triggered multiple responses, with one response or two from the student who was interacting with the teacher and one from his/her peers, as shown in example 4 and example 5 (notes: S = speaking student; T = teacher; Ss = peers).

Example 1: Recast for question

- S: Is your first object is a famous man?
T: Is your first object a famous person?
S: not a person.

Example 2: Recast for past-tense form

- S: All of a sudden, a man behind him who held a stick hit his head.
T: All of a sudden, a man behind him with a stick hit his head.
S: Yeah.

Example 3: Negotiated prompt for question

- S: Is it glasses?
T: Is it glasses? (Repetition)
S: Are they glasses?

Example 4

- S: Where we can see it?
T: Sorry. (Clarification request)
S: Where we can see it?
T: Where. (Elicitation)
Ss: Where can we.
T: Where. (Elicitation)
S: Where can we see it?

Example 5: Negotiated prompt for past tense form

- S: They all very happy.
T: They (Elicitation)

Table 3. Descriptive statistics of pre-, immediate post- and delayed post-tests for wh-questions.

Groups ($n = 30$ each)	Pre-test		Immediate post-test		Delayed post-test	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Recasts	.53	.15	.64	.15	.65	.14
Negotiated prompts	.52	.19	.59	.22	.59	.19
Control	.47	.17	.49	.19	.53	.20

S: xx [Silence]

T: They (Elicitation)

S: They are all very happy.

Ss: They were.

S: They were all very happy.

b *Equivalence of the groups (pre-test results).* The descriptive statistics of the pre-test scores of the three groups' accurate use of four structures can be found in the main analyses. The one-way ANOVAs showed the three groups were equivalent in the pre-test performances with regard to the learners' accurate use of each structure: $F(2, 87) = 1.16, p = .32$ for wh-questions; $F(2, 77) = .23, p = .80$ for interrogative questions¹; $F(2, 87) = .98, p = .38$ for irregular past tense verbs; and $F(2, 87) = .05, p = .96$ for regular past tense verbs.

2 Main analyses

The main analysis for each target structure involved the performance of a mixed-design two-way ANOVA ($p \leq .05$). If the mixed ANOVA found any significant effects, post hoc multiple comparisons, which applied a Bonferroni adjustment to the p value to avoid Type I error (Pallant, 2010), were conducted to detect the significant differences. Parametric statistics, such as ANOVAs, were justified for the analyses based on the large sample size ($n = 30$ in each group) and the results of the normality test (i.e. Shapiro–Wilk test). Partial eta-square (η^2) is reported for the mixed ANOVA to show the effect size of a significant difference ($\eta^2 = .01$ as small effect, $\eta^2 = .06$ as moderate effect, and $\eta^2 = .14$ as large effect). Cohen's d is used for any differences found in the post hoc analyses ($d < 0.2$ as small effect, $d > 0.5$ as medium effect, and $d > 0.8$ as large effect).

a *Effects of feedback on accurate use of wh-questions.* Table 3 summarizes the descriptive statistics for pre-, immediate post-, and delayed post-test results for learners' accurate use of wh-questions. In the mixed ANOVA, the additional assumptions of homogeneity of variances and inter-correlations were met. The results indicated a significant time effect, $F(2, 86) = 13.35, p \leq .000, \eta^2 = .24$, and a significant group effect, $F(2, 87) = 4.19, p = .018, \eta^2 = .09$. However, the interaction between group and time was not significant, $F(4, 172) = 1.16, p = .33, \eta^2 = .03$.

Table 4. Descriptive statistics of pre-, immediate post- and delayed post-tests for interrogative questions.

Groups	Pre-test		Immediate post-test		Delayed post-test	
	M	SD	M	SD	M	SD
Recasts ($n = 28$)	.69	.16	.73	.18	.79	.16
Negotiated prompts ($n = 28$)	.66	.22	.76	.19	.77	.21
Control ($n = 24$)	.68	.20	.63	.21	.58	.24

Subsequent post hoc comparisons found that the recast group significantly improved its scores from the pre-test to the immediate post-test ($p = .001$) with medium effects ($d = .73$), and from the pre-test to the delayed post-test ($p \leq .000$) with large effects ($d = .83$). However, neither the negotiated prompt nor the control group significantly increased their scores from the pre-test to the post-tests. Further comparisons regarding the group effect revealed that the differences lay in the comparison of the recast group vs. the control group at the immediate post-test ($p = .009$, $d = .87$) and at the delayed post-test ($p = .025$, $d = .71$).

In summary, the analyses revealed four significant differences with medium or large effect sizes, and all the differences related to the recast group. Therefore, it can be concluded that recasts were the only CF that benefited the learners' development of wh-questions.

b Effects of feedback on accurate use of interrogative questions. Table 4 presents the descriptive statistics of how the learners in the three groups produced interrogative questions in the pre-, immediate post-, and delayed post-tests. Satisfying the additional assumptions of homogeneity of variances and inter-correlations, the results showed that there was a significant group effect, $F(2, 77) = 4.30$, $p = .017$, $\eta^2 = .10$, and a significant group \times time interaction, $F(4, 152) = 3.77$, $p = .006$, $\eta^2 = .090$, but there was no significant effect for time, $F(2, 76) = 1.32$, $p = .27$, $\eta^2 = .03$.

Post hoc comparisons revealed that the recast group performed significantly better at the delayed post-test compared to the group's pre-test ($p = .012$, $d = .59$) and the control's delayed post-test ($p = .001$, $d = 1.01$). Instead, the negotiated prompt group significantly outperformed at the immediate post-test with relation to the group's pre-test ($p = .038$, $d = .49$) and to the control group's two post-tests ($p = .047$, $d = .65$ and $p = .003$, $d = .85$, respectively).

It can be seen that both experimental groups made significant improvement at one of the two post-tests and outperformed the control group at one post-test or two, with the effect sizes from medium to large. As there were similar significant differences and similar effect sizes for both experimental groups, it is then logical to conclude that both CF types were similarly beneficial for the learners' development of interrogative questions.

c Effects of feedback on accurate use of irregular past tense verbs. Table 5 summarizes the descriptive statistics for pre-, immediate post-, and delayed post-test results

Table 5. Descriptive statistics of pre-, immediate post- and delayed post-tests for irregular past tense verbs.

Groups (<i>n</i> = 30 each)	Pre-test		Immediate post-test		Delayed post-test	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Recasts	.46	.18	.53	.20	.60	.18
Negotiated prompts	.47	.21	.52	.22	.50	.23
Control	.41	.18	.37	.20	.41	.18

regarding the learners' accurate use of irregular past tense verbs. The assumptions of homogeneity of variances and inter-correlations were satisfied. The results showed a significant time effect, $F(2, 86) = 5.04, p = .008, \eta^2 = .11$, a significant group \times time interaction, $F(4, 172) = 3.14, p = .016, \eta^2 = .068$, and a significant group effect, $F(2, 87) = 4.94, p = .009, \eta^2 = .102$.

The results demonstrate that the recast group was the only group that showed significant gains in scores from the pre-test to the delayed post-test ($p = .000, d = .80$). Although the recast group achieved considerable gains in scores from the pre-test to the immediate post-test, no significance was reached. With regard to the group differences, the results of both feedback groups were significantly different from those of the control group at the immediate post-test (recasts: $p = .01, d = .80$; negotiated prompts: $p = .015, d = .72$). The recast group also significantly outperformed the control at the delayed post-test ($p = .002$) with large effects ($d = 1.02$). In contrast, the results of the negotiated prompt group were not significantly different from those of the control group at the delayed post-test.

In summary, the recast group showed three significant differences in relation to time and group, all with large effect sizes, whereas the negotiated prompt group showed one significant difference in relation to group at the time of the immediate post-test, with medium effect size. As such, the answer to research question 3 is that both feedback types benefited the learners' development of irregular past tense verbs in narrative storytelling, but recasts were more effective than negotiated prompts.

d Effects of feedback on accurate use of regular past tense verbs. Table 6 provides the descriptive statistics for pre-, immediate post-, and delayed post-test results regarding the learners' accurate use of regular past tense verbs. Again, there was satisfaction of the assumptions about homogeneity of variances and inter-correlations. The mixed ANOVA yielded a significant time effect, $F(2, 86) = 4.38, p = .015, \eta^2 = .092$. There was no significant difference for group \times time interaction, $F(4, 170) = 1.06, p = .38, \eta^2 = .02$, or group effect, $F(2, 87) = 1.14, p = .33, \eta^2 = .025$.

Post hoc comparisons showed that the negotiated prompt group significantly improved in test scores from the pre-test to the immediate post-test ($p = .012$) and maintained the improvement at the delayed post-test ($p = .027$), both with medium effects ($d = .46, d = .52$). Although the recast group showed gains in scores from the pre-test to the post-tests, this group only approached significance ($p = .06$).

Table 6. Descriptive statistics of pre-, immediate post- and delayed post-tests for regular past tense verbs.

Groups (<i>n</i> = 30 each)	Pre-test		Immediate post-test		Delayed post-test	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Recasts	.52	.23	.62	.19	.62	.17
Negotiated prompts	.54	.22	.64	.18	.65	.21
Control	.54	.26	.57	.24	.53	.26

Because the negotiated prompt group achieved significant differences in the multiple comparisons and the recast group did not, the answer to research question 4 is that only negotiated prompts were effective in the learners' accurate use of regular past tense verbs.

V Discussion

The present study investigated the effectiveness of recasts and negotiated prompts in accuracy development of questions and past tense verbs in a pedagogical context that focuses on grammar and accuracy. It found that recasts were effective in the accuracy development of wh-question and interrogative questions and irregular past tense, but not in the development of regular past tense. In comparison, negotiated prompts were found to be helpful for the accuracy development of interrogative questions and irregular and regular past tense verbs, but not helpful for the development of wh-questions. Most of the time, the recasts achieved large effect sizes, but the negotiated prompts achieved only medium effect sizes. See the summary of results in Online Appendix 4. The findings of the present study will be discussed in relation to questions, past tense verbs, and input-providing recasts vs. output-providing negotiated prompts.

I Target structures

a Questions. Several studies found a beneficial role of recasts and prompts for the development of questions (Mackey & Philp, 1998; McDonough, 2005; McDonough & Mackey, 2006). For example, Mackey and Philp (1998) found that recasts helped learners not only to produce a higher stage of questions but also to increase the production of already-acquired question types. McDonough (2005) provided evidence that clarification requests facilitated the production of a higher stage of questions when the learners successfully modified their initial, non-target-like utterances. The present study supports the beneficial role of the two CF types in question development by finding that both recasts and negotiated prompts yielded positive impacts on production of more accurate interrogative questions.

The present study found that wh-questions with inversions were not amenable to negotiated prompts but highly amenable to recasts. As shown in Table 2, for each treatment group, there were similar numbers of CF episodes targeting this question type

(11 episodes for each type of CF). This similarity dispels any role of feedback frequency in the differing effects of recasts and negotiated prompts on *wh*-questions with inversion. The differing effects may then result from the interaction of two factors. One factor relates to *wh*-questions with inversion being more complex than interrogative questions regarding the number of elements that need to be moved and changed, although both question types are classified as rule-based learning. The other factor relates to the input-providing nature of recasts, which helps to reduce the cognitive load in a learner's processing of complex structures (Mifka-Profozic, 2013). These two factors in conjunction resulted in recasts being beneficial for the complex structure of *wh*-questions that involve inversion. Negotiated prompts, on the other hand, cannot play a scaffolding role due to their non-provision of input, thereby leading to a lesser or nil effectiveness in this question type.

It is of interest to note that compared to *wh*-questions, interrogative questions received irregular provision of recasts and negotiated prompts (5 episodes for each type of CF) but were responsive to the feedback treatment. This finding may result from the relative simplicity of interrogative questions, and a few episodes of feedback may be sufficient for the learners to gain improvement. This finding, along with what was found for *wh*-questions, seems to indicate that frequency of feedback did not contribute to the effectiveness of both feedback types in question forms. Indeed, Mifka-Profozic (2013) also found that recasts were the only CF that worked for the acquisition of a more complex structure, but both recasts and clarification requests facilitated the acquisition of a simple structure.

b Past tense verbs. The findings that recasts were effective only for irregular past tense verbs but that negotiated prompts were effective for both irregular and regular types are congruent with the findings of Yang and Lyster (2010). Such findings may be explained in terms of the salience of structure and the manner of CF provision. Yang and Lyster state that irregular past tense verbs are salient due to their distinctive pronunciations from the base verbs and high frequency of use. In this study, the salience of irregular past tense verbs was enhanced by a majority of instances of recasts and negotiated prompts being provided to this structure (see Table 2: 11 episodes for each type of CF). In contrast, regular past tense verbs have low salience because the addition of a voiceless *-ed* is less noticeable and semantically redundant (DeKeyser, 1998; R. Ellis, 2007). Adding to its low salience, regular past tense was provided with fewer instances of recasts (5 episodes) and negotiated prompts (7 episodes). Given the differences in salience and frequency of recasts, it may be little wonder that recasts were effective in the development of irregular past tense verbs but not in the development of regular past tense verbs.

Compared to the consistent provision of implicit recasts (e.g. full reformulation, single move, no prosodic stress, and no uptake opportunities), the provision of negotiated prompts tended to be explicit. Most of the episodes contained multiple prompt moves, and two-thirds of the moves were elicitation, which is a type of CF toward the explicit end on the implicit/explicit feedback scale (Lyster et al., 2013). Moreover, each of the negotiated prompts afforded an opportunity for uptake, and if a learner could not produce a successful uptake after being prompted several times, the teacher asked for peer help. For regular past tense verbs with low salience, the explicitness of negotiated prompts

was likely to draw the learners' attention to their problematic production of these verbs, despite the feedback on this structure being infrequently provided. Given irregular past tense verbs owning high salience and being frequently provided with feedback, the explicitness of negotiated prompts would, of course, make the learners more aware of their problematic production of these irregular past tense verbs. As such, in the current study, negotiated prompts were found to be effective for both types of past tense verbs. However, the impact that negotiated prompts yielded on the two types of past tense verbs did not seem to exhibit any distinctive difference.

In addition to the difference in salience between the two kinds of past tense verbs, salience pertains to the comparison of questions and the past tense. Researchers (N. Ellis, 2004; VanPatten, 1996; VanPatten et al., 2004) argue that questions carry higher communicative value and are perceptually more salient than morphological forms, such as the past tense. This difference seems to offer an explanation to the finding that although recasts were implicit, they were effective for both question types and that because negotiated prompts tended to be explicit, they were effective for the learners' increased control in the use of the two types of past tense verbs.

2 Input-providing and output-promoting CF

Despite the finding that the effects of both recasts and negotiated prompts were structure dependent, another trend emerged. That is, recast effectiveness was usually associated with large effect sizes, whereas the effectiveness of negotiated prompts was usually with medium effect sizes. These results were in contrast to Yang and Lyster's study (2010), which found that large effect sizes were associated mainly with the prompt group. It then can be argued that recasts showed a certain advantage over negotiated prompts in this study, whereas prompts were advantageous to recasts in Yang and Lyster's study.

The advantage of recasts as found in the present study may result from their provision of input. Because the learners were irregularly exposed to language tasks for communication and were not explicitly informed of what their errors were through metalinguistic clues, they may have needed the input in recasts to scaffold their noticing of a gap in their production, conducting of a cognitive comparison and realizing of form and function mapping (Doughty, 2001; Long, 1996, 2007). As discussed earlier, Mifka-Profozic (2013) emphasized the scaffolding role that recasts play in the FL context where the instructional focus is on grammar and accuracy. However, we have to acknowledge at the same time that the target-like utterances in recasts may serve as exemplars for some learners (Leeman, 2003; Nicholas et al., 2001).

On the other hand, output-promoting negotiated prompts may be disadvantageous for some of our learners. Batstone (2002) believes that pushed output can be face-threatening and risk-taking, and that L2 learners need time to develop the risk-taking behavior that facilitates pushed output. Because the learners in the present study were given little time for speaking up in their English classes, they may have been anxious and unwilling to take risks when pushed for output. Therefore, there is no guarantee that they were able to access the rule-based knowledge system through prompts and apply it in their responses, even if they had ample grammatical knowledge of the target structures. As shown in

examples 4 and 5 earlier, the learners could not self-correct their initial utterances until the peer students provided the corrections.

The finding that our learners had difficulty in modifying their output following negotiated prompts is different from that in Yang and Lyster's study (2010). Yang and Lyster found that prompts consistently elicited successful uptake from their learners, and then attributed the greater effectiveness of prompts partly to the self-repair being elicited. The L1 Chinese learners in Yang and Lyster were second-year English-major university students, whose classes were all conducted in English during their four-year study and who had the intended purpose of becoming competent in their all-round use of the L2. After being exposed to an extensive amount of input and being conscious of practicing their spoken English, they may have developed the risk-taking behavior when pushed for output. With the development of risk-taking behavior, the provision of an explicit prompt in the form of metalinguistic clues and the receiving of form-focused instruction, it may not be surprising that the learners in Yang and Lyster were able to modify their output after prompts and then made substantial gains in learning.

VI Conclusions

The current study investigated the effects of recasts and negotiated prompts on the development of questions and past tense verbs among 90 adult learners in an FL environment in China. The results show that both CF types were effective for the accuracy development of three structures, but not effective for one structure. However, the results show that recasts generated strong effects, but negotiated prompts yielded mainly medium effects.

The positive effects of CF that were found in this study, especially recasts, on accuracy development in an FL learning context with grammar-focused instruction encourage teachers to devise various strategies to provide CF in their teaching. Information on the role of CF in accuracy development could be provided to all teachers in their in-service and pre-service training. Types of CF can be chosen according to target structures and instructional focuses. Recast feedback can be brief and given in a single move and non-segmentally when the structure is salient and the instructional focus is on grammar and accuracy. However, for negotiated prompt feedback, such as clarification requests and even elicitation, the corrective intention may need to be made clear in certain contexts, such as the context of the present study, so that students will be encouraged to revise their initial utterance.

The results of the present study have enhanced understanding of the effectiveness of recasts and negotiated prompts in L2 classroom learning; however, the study had certain limitations, which can be addressed in future research. First, a limitation exists in the involvement of a control group, which took the three versions of the test without receiving any task-based treatment lessons. The positive outcomes for the two experimental groups may have resulted partly from the provision of CF, but also partly from the extra practice that the three treatment lessons provided. Future research needs to include a control group that has the lack of provision of CF as the only variable differentiating it from the experimental groups; therefore, any gains in the post-test scores can be purely credited to the feedback treatment.

A second limitation lies in the delivery of recasts and negotiated prompts. As the present study was carried out in a classroom context, pedagogical issues were considered in the operationalization of the feedback; for example, allowing the teacher's occasional negotiation for meaning with the learners in the recast group when he did not hear or understand the learners' utterances, or allowing the teacher to seek peer correction when the learners were not able to self-correct after being prompted several times. These considerations resulted in the occasional provision of clarification requests within the recast group, and of recasts and explicit feedback within the negotiated prompt group, but the two experimental groups were obviously distinguishable with regard to the vast majority of the CF they received.

Furthermore, because the provision of CF depended on the learners' non-target-like production of the target structures, this study could not predict which target structure the learners would use more and/or which target structure the learners would produce more errors in, especially for the differences in production between *wh*-questions and interrogative questions and between the regular and irregular past tense verbs. This unpredictability subsequently resulted in more instances of recasts and negotiated prompts being provided to *wh*-questions and irregular past tense verbs and fewer instances of both CF types to interrogative questions and regular past tense verbs, although there were similar instances of both CF types being provided to the same structure. While the provision of similar instances of recasts and negotiated prompts on one structure did not affect the aim of the study to compare the effects of the CF on the structure, the unevenness of CF provision across structures may contribute to the CF being effective for one structure but not the other structure. While eliciting learners' natural production of different target structures, future classroom studies need to work out a method to control and balance the provision of CF so that feedback frequency can be dismissed as being a contributing factor to the efficacy of the feedback. Furthermore, as indicated in the discussion section that feedback frequency did not seem to contribute to CF efficacy on the question form and may have affected CF efficacy on the past tense form, future studies need to clarify the interacting effects between feedback frequency, feedback types and target structures.

The present study also excluded metalinguistic clues in the operationalization of the prompt feedback, which is different from most previous experimental classroom research. With this current operationalization, recasts became a favorable CF technique for *wh*-questions and irregular past tense verbs in an FL learning environment with grammar-focused instruction. It would be interesting to test the effects of recasts against metalinguistic clues in the same learning environment. It would also be interesting for future research to follow the current operationalization to test the effects of recasts and negotiated prompts on different structures in varied classroom contexts with learners at different proficiency levels.

Third, the testing instrument used in the study was only of one type, the oral production test, due to the limited availability of time for the volunteer learners. Although our learners had grammatical knowledge of the target structures, we were not able to measure their explicit knowledge through written tests and grammaticality judgment tests. Future studies that compare the efficacy of recasts and prompts in L2 classroom learning should employ a variety of testing measures in their methodological design.

Acknowledgements

We would like to thank all participant learners and the instructor who joined this study. We also thank the anonymous reviewers and the editors for their helpful comments on the earlier drafts of this paper.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This research received the Postgraduate Travel Grant from the University of Queensland, Australia.

Note

1. Ten learners were removed from the analysis regarding the effects of recasts and negotiated prompts on interrogative questions, because the learners' production of this question type was not consistent across the three tests. Specifically, they did not produce interrogative questions in one of the pre-test, the immediate post-test or the delayed post-test, and then the data related to their production of this question type became outliers in the ANOVA analyses. Of the 10 learners, two learners were from each of the experimental groups and six from the control group. Other than the 10 learners, the rest of the learners produced interrogative questions in a relatively consistent manner.

ORCID iD

Noriko Iwashita  <https://orcid.org/0000-0002-1402-7365>

Supplemental material

Supplemental Appendix for this article is available online.

References

- Ammar, A., & Spada, N. (2006). Recasts, prompts, and L2 learning. *Studies in Second Language Acquisition*, 28, 543–574.
- Anderson, J.R. (1993). *Rules of the mind*. Hillsdale, NJ: Lawrence Erlbaum.
- Anderson, J.R. (2000). *Learning and memory: An integrated approach*. New York: John Wiley.
- Batstone, R. (2002). Contexts of engagement: A discourse perspective on 'intake' and 'pushed output'. *System*, 30, 1–14.
- Chen, Q., & Klenowski, V. (2009). *Assessment and curriculum reform in China: The college English test and tertiary English as a foreign language education*. In Proceedings of the 2008 AARE International Education Conference, Queensland University of Technology, Brisbane, Australia.
- Collentine, J., & Freed, B.F. (2004). Learning context and its effects on second language acquisition. *Studies in Second Language Acquisition*, 26, 153–171.
- Cortazzi, M., & Jin, L. (1996a). Cultures of learning: Language classrooms in China. In Coleman H. (Ed.), *Society and the language classroom* (pp. 169–205). Cambridge: Cambridge University Press.
- Cortazzi, M., & Jin, L. (1996b). English teaching and learning in China. *Language Teaching*, 29, 61–80.
- Cortazzi, M., & Jin, L. (2001). Large classes in China: 'Good' teachers and interaction. In Watkins D.A., & J.B. Biggs (Eds.), *Teaching the Chinese learner: Psychological and pedagogical per-*

- spectives (pp. 115–134). Camberwell, VIC / Hong Kong: Comparative Education Research Centre, the University of Hong Kong.
- DeKeyser, R. (1998). Beyond focus on form: Cognitive perspectives on learning and practicing second language grammar. In Doughty C., & J. Williams (Eds.), *Focus on form in classroom second language acquisition* (pp. 42–63). Cambridge: Cambridge University Press.
- Doughty, C. (2001). Cognitive underpinnings of focus on form. In Robinson P. (Ed.), *Cognition and second language instruction* (pp. 125–151). New York: Cambridge University Press.
- Ellis, N. (2004). The processes of second language acquisition. In VanPatten B., Williams J., Rott S., & Overstreet M. (Eds.), *Form–meaning connections in second language acquisition* (pp. 49–76). Mahwah, NJ: Lawrence Erlbaum.
- Ellis, R. (2006). Researching the effects of form-focused instruction on L2 acquisition. *AILA Review*, 19, 18–41.
- Ellis, R. (2007). The differential effects of corrective feedback on two grammatical structures. In Mackey A. (Ed.), *Conversational interaction in second language acquisition* (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.
- Goo, J., & Mackey, A. (2013). The case against the case against recasts. *Studies in Second Language Acquisition*, 35, 127–165.
- Gooch, R., Saito, K., & Lyster, R. (2016). Effects of recasts and prompts on L2 pronunciation development: Teaching English /ɪ/ to Korean adult EFL learners. *System*, 60, 117–127.
- Han, Z. (2002). A study of the impact of recasts on tense consistency in L2 output. *TESOL Quarterly*, 36, 543–572.
- Han, Z., & Kim, J.H. (2008). Corrective recasts: What teachers might want to know. *Language Learning Journal*, 36, 35–44.
- Jin, L., & Cortazzi, M. (1998). Dimensions of dialogue: Large classes in China. *Interactional Journal of Educational Research*, 29, 739–761.
- Leeman, J. (2003). Recasts and second language development: Beyond negative evidence. *Studies in Second Language Acquisition*, 25, 37–63.
- Li, S. (2010). The effectiveness of CF in SLA: A meta-analysis. *Language Learning*, 60, 309–365.
- Liu, M., Xu, W., & Ran, Q. (2015). An empirical study of writing feedback analysis of non-English majors in China with natural language processing technologies. *International Journal of e-Education, e-Business, e-Management and e-Learning*, 5, 85–93.
- Loewen, S., & Nabei, T. (2007). Measuring the effects of oral corrective feedback on L2 knowledge. In Mackey A. (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 English L2 classroom: Characteristics, explicitness and effectiveness. *The Modern Language Journal*, 90, 536–556.
- Long, M. (1996). The role of the linguistic environment in second language acquisition. In Ritchie W., & T. Bhatia (Eds.), *Handbook of second language acquisition* (pp. 413–468). San Diego, CA: Academic Press.
- Long, M. (2007). *Problems in SLA*. Mahwah, NJ: Lawrence Erlbaum.
- Lyster, R. (1998). Recasts, repetition, and ambiguity in L2 classroom discourse. *Studies in Second Language Acquisition*, 20, 55–85.
- Lyster, R. (2004). Differential effects of prompts and recasts in form-focused instruction. *Studies in Second Language Acquisition*, 26, 399–432.
- Lyster, R. (2015). The relative effectiveness of corrective feedback in classroom interaction. In Markee N. (Ed.), *The handbook of classroom discourse and interaction* (pp. 213–228). Oxford: Wiley-Blackwell.

- 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., & Ranta, L. (2013). Counterpoint piece: The case for variety in CF research. *Studies in Second Language Acquisition, 35*, 167–184.
- Lyster, R., Saito, K., & Sato, M. (2013). Oral corrective feedback in second language classrooms. *Language Teaching, 46*, 1–40.
- Mackey, A. (1999). Input, interaction and second language development: An empirical study of question formation in ESL. *Studies in Second Language Acquisition, 21*, 557–587.
- Mackey, A., & Philp, J. (1998). Conversational interaction and second language development: Recasts, responses, and red herrings? *Modern Language Journal, 82*, 338–356.
- McDonough, K. (2005). Identifying the impact of negative feedback and learner's responses on ESL question development. *Studies in Second Language Acquisition, 27*, 79–103.
- McDonough, K., & Mackey, A. (2006). Responses to recasts: Repetitions, primed production, and linguistic development. *Language Learning, 56*, 693–720.
- Mifka-Profozic, N. (2013). *The effectiveness of corrective feedback and the role of individual differences in language learning*. Frankfurt/Mein: Peter Lang.
- Nicholas, H., Lightbown, P., & Spada, N. (2001). Recasts as feedback to language learners. *Language Learning, 51*, 719–758.
- Pallant, J. (2010). *SPSS survival manual: A step by step guide to data analysis using SPSS*. 4th edition. Maidenhead: McGraw-Hill Education.
- Philp, J. (2003). Constraints on 'noticing the gap': Non-native speakers' noticing of recasts in NN–NNS interaction. *Studies in Second Language Acquisition, 25*, 99–126.
- Pica, T. (1983). Methods of morpheme quantification: Their effect on the interpretation of second language data. *Studies in Second Language Acquisition, 6*, 69–78.
- Robinson P., (Ed.). (2001). *Cognition and second language instruction*. Cambridge: Cambridge University Press.
- Russell, J., & Spada, N. (2006). The effectiveness of CF for second language acquisition. In Norris J., & L. Ortega (Eds.), *Synthesizing research on language learning and teaching* (pp. 131–164). Philadelphia, PA / Amsterdam: John Benjamins.
- Schmidt, R. (1990). The role of consciousness in second language learning. *Applied Linguistics, 11*, 129–158.
- Schmidt, R. (1993). Awareness and second language acquisition. *Annual Review of Applied Linguistics, 13*, 206–226.
- Schmidt, R. (1995). Consciousness and foreign language learning: A tutorial on the role of attention and awareness in learning. In Schmidt R.W. (Ed.), *Attention and awareness in foreign language learning* (pp. 1–63). Honolulu, HI: University of Hawai'i.
- Sheen, Y. (2004). CF 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.
- Spada, N., & Lightbown, P.M. (1993). Instruction and the development of questions in L2 classrooms. *Studies in Second Language Acquisition, 15*, 205–224.
- Swain, M. (1985). Communicative competence: Some roles of comprehensible input and comprehensible output in its development. In Gass S.M., & C. Madden (Eds.), *Input in second language acquisition* (pp. 235–253). Rowley, MA: Newbury House.

- Swain, M. (1995). Three functions of output in second language learning. In Cook C., & B. Seidlhofer (Eds.), *Principle and practice in applied linguistics: Studies in honor of H. G. Widdowson* (pp. 125–144). Oxford: Oxford University Press.
- Swain, M. (2005). The output hypothesis: Theory and research. In Hinkel E. (Ed.), *Handbook of research in second language teaching and learning* (pp. 471–483). Mahwah, NJ: Lawrence Erlbaum.
- van de Guchte, M., Braaksma, M., Rijlaarsdam, G., & Bimmel, P. (2015). Learning new grammatical structures in task-based language learning: The effects of recasts and prompts. *Modern Language Journal*, 99, 246–262.
- VanPatten, B. (1996). *Input processing and grammar instruction in second language acquisition*. Norwood NJ: Ablex.
- VanPatten, B., Williams, J., & Rott, S. (2004). Form–meaning connections in second language acquisition. In VanPatten B., Williams J., Rott S., & Overstreet M. (Eds.), *Form–meaning connections in second language acquisition* (pp. 1–26). Mahwah, NJ: Lawrence Erlbaum.
- Yang, Y., & Lyster, R. (2010). Effects of form-focused practice and feedback on Chinese learners' acquisition of regular and irregular past tense forms. *Studies in Second Language Acquisition*, 32, 235–263.