Studies in Second Language Acquisition, 2012, **34**, 591–626.
doi:10.1017/S0272263112000356

PEER INTERACTION AND CORRECTIVE FEEDBACK FOR ACCURACY AND FLUENCY DEVELOPMENT

Monitoring, Practice, and Proceduralization

Masatoshi Sato
Universidad Andres Bello, Chile

Roy Lyster
McGill University, Canada

This quasi-experimental study is aimed at (a) teaching learners how to provide corrective feedback (CF) during peer interaction and (b) assessing the effects of peer interaction and CF on second language (L2) development. Four university-level English classes in Japan participated ($N = 167$), each assigned to one of four treatment conditions. Of the two CF groups, one was taught to provide prompts and the other to provide recasts. A third group participated in only peer-interaction activities, and a fourth served as the control group. After one semester of intervention, the two CF groups improved in both overall accuracy and fluency, measured as unpruned and pruned speech rates, whereas the peer-interaction-only group outperformed the control group only on fluency measures. This study draws on monitoring in speech-production theory and the declarative-procedural model of skill-acquisition theory to interpret these results, thus contributing a new theoretical approach to CF research in the context

We are grateful to the teachers as well as the students who participated in the study. We thankfully acknowledge the valuable comments from the anonymous reviewers.

Address correspondence to Masatoshi Sato, Department of English, Universidad Andrés Bello, Santiago, Chile. E-mail: masatoshi.sato@unab.cl.

of peer interaction in which learners can be providers of CF. It is concluded that whereas peer interaction offered opportunities for repeated production practice, facilitating proceduralization, CF sharpened learners' ability to monitor both their own language production and that of their interlocutors.

Since form-focused instruction was proposed in the early 1990s (Lightbown & Spada, 1990; see also Spada, 1997, 2011), corrective feedback (CF) has been given much attention from both theoretical and pedagogical perspectives. Considered to trigger psycholinguistic processing that is conducive to second language (L2) development, CF has proven its general effectiveness with differential effects depending on age (Lyster & Saito, 2010), data-collection contexts (Li, 2010; Mackey & Goo, 2007), learning contexts (Li, 2010), treatment lengths (Li, 2010; Lyster & Saito, 2010), and linguistic targets (Mackey & Goo, 2007). Another important variable considered to mediate CF effectiveness is CF type, comprising two dimensions: (a) output-eliciting versus input-providing feedback (see Ammar & Spada, 2006; de Bot, 1996; Sheen & Ellis, 2011; Yang & Lyster, 2010) and (b) explicitness versus implicitness (see R. Ellis, 2006, 2009a; R. Ellis & Sheen, 2006; Nicholas, Lightbown, & Spada, 2001; Sato, 2011). Controlling the implicit-explicit dimension in the context of peer interaction, the current study examines the differential effects of prompts and recasts not only on accuracy development but also on fluency, on the premise that, if CF contributes to the process of automatization, its effect will also be observed in the speed of processing.

This study originated in claims that L2 development is triggered by interaction during which naturally occurring communication breakdowns lead to negotiation for meaning involving interactional feedback that facilitates mutual comprehension (e.g., Gass, 1997; Long, 1996; Mackey, 2007). The current study takes a further step and attempts to maximize this learning opportunity by explicitly teaching learners how to provide CF to one another during peer interaction activities and to investigate whether CF in peer interaction affects L2 processing in ways that are similar to the processing triggered by teachers' CF in L2 classrooms. Although classroom CF research has evolved from observational (e.g., Lyster & Ranta, 1997) to experimental (e.g., Ammar & Spada, 2006) designs and has shown positive effects for CF, what remains open for further investigation is more precisely how the different types of processing triggered by different interactional moves affect subsequent development. In the case of peer interaction, research has remained descriptive: No studies, to the authors' knowledge, have experimentally examined interactional patterns in peer interaction as variables to see how they contribute to L2 development (but see Adams, 2007, who showed the relationship between peer feedback and grammatical development). The current research takes this next step by investigating the relationship between various interactional moves and L2 development achieved over time.

CORRECTIVE FEEDBACK AND L2 DEVELOPMENT

Accuracy Development

Corrective feedback effectiveness has been explained in the study of L2 acquisition through reference to the noticing hypothesis (Schmidt, 1994, 2001), which claims that noticing is necessary and that "intake is what learners consciously notice" (Schmidt, 1990, p. 149). The importance of attention and awareness for L2 acquisition has been investigated by many researchers, especially within the framework of focus on form (e.g., Doughty, 2001; Doughty & Williams, 1998; Long & Robinson, 1998). Though the optimal level of consciousness while processing input is debatable, it is claimed that CF gives learners an opportunity to make a cognitive comparison between their interlanguage and the given input (R. Ellis, 1994) and to engage in focused input analysis (N. Ellis, 2005). This line of thought is especially applicable to input-providing CF such as recasts, but less to other output-prompting types of CF that do not provide targetlike models with which learners can compare their erroneous utterance. The benefits of prompts have instead been attributed to the opportunities they provide to learners to self-correct their initial erroneous utterances (e.g., R. Ellis, Basturkmen, & Loewen, 2001; Lyster & Ranta, 1997; Panova & Lyster, 2002). This modification move is claimed to contribute to L2 development on the basis of Swain's (1985) output hypothesis and the benefits of comprehensible output that is "not only conveyed, but that is conveyed precisely, coherently, and appropriately" (p. 249). During language production, learners may notice what they know and what they can actually say and may also test linguistic hypotheses whereby their attention is drawn to syntactic properties of the language (Swain, 1995, 2005). Because the noticing triggered by output and the resulting modified output (MO) have proven beneficial for improving accuracy (Izumi & Bigelow, 2000; Nobuyoshi & Ellis, 1993), the effectiveness of prompts has been attributed to this cognitive processing (e.g., Ammar & Spada, 2006; Lyster, 2004; McDonough, 2005).

Prompt effectiveness can also be explained by skill-acquisition theory (Anderson, 1983, 2005; Johnson, 1996), which postulates a gradual shift of knowledge from declarative to procedural (i.e., proceduralization), during which repeated practice plays an essential role (DeKeyser, 1998, 2001, 2007). Practice, which can broadly be defined as "specific activities in the second language, engaged in systematically, deliberately, with the goal of developing knowledge of and skills in the second language" (DeKeyser, 2007, p. 1), has shown its positive impact on L2 development in neurocognitive and neurolinguistic research as well (Paradis, 2004, 2009; Ullman, 2001, 2005), and some researchers have adapted the theory to CF effectiveness. The theory claims that CF at propitious moments is necessary to avoid entrenching wrong knowledge structures (Anderson & Schunn, 2000); therefore, prompts, which provide learners with opportunities to engage in "controlled practice in the context of communicative interaction" (Lyster & Izquierdo, 2009, p. 462), are effective both for developing accurate knowledge by restructuring their already existing knowledge and for enhancing the practice effect by pushing them to self-correct (Ranta & Lyster, 2007).

In addition to drawing on theories from cognitive psychology and SLA, the current study proposes a new perspective for CF research: the integration of monitoring in Levelt's (1989, 1999) speech-production model with declarative-procedural models (see Ullman, 2004). Although Levelt's studies have been adopted by some SLA researchers to discuss focus on form (e.g., Doughty, 2001) or the output hypothesis (e.g., Izumi, 2003; see also de Bot, 1992, who adapted the model for L2 speech), the current study focuses on the monitor in terms of its functionality for both production and comprehension during peer interaction. In such contexts, learners may monitor their own grammatical errors during and after the production process, and their input may also contain errors that trigger their monitoring for comprehension. As argued by Crookes (1991) and Towell, Hawkins, and Bazergui (1996), each stage in the modular model (i.e., conceptualizer, formulator, and articulator) consists of declarative and procedural knowledge, and speech production itself can be considered as repeated practice (see also Muranoi, 2007). That is, slow L2 speech can be due to incomplete proceduralization and, through practice, L2 processing can become parallel and incremental—that is, automatized (Segalowitz, 2003).

Particularly relevant to CF in this model is monitoring in conjunction with computing speech and parsing input in the comprehension system. Arguing that monitoring functions in several stages of production, Levelt (1983) proposed the perceptual theory of monitoring. A speaker can compare the preverbal message to his or her original intention, examine the outcome of the phonological process at the articulatory stage, and finally detect errors in parsed speech. In this regard, Kormos (2006) argued that the notion of monitoring interfaces with SLA theories in several ways. First, because it involves checking not only internal but also external speech, it can contribute to noticing the gap between insufficiently automatized knowledge and the given input (Schmidt, 1994). Second, noticing itself triggered by monitoring is beneficial for L2 development while both processing input (Robinson, 1995) and producing output (Swain, 1985). This process of detecting an error, rehearsing the error-free solution, and, thus, reducing the error rate contributes to proceduralization by storing correct linguistic representations in long-term memory (Anderson, 1995).

Most importantly, though without any reference to negative evidence or CF, it is claimed that monitoring functions in a similar way to comprehension processes via the acoustic-phonetic processing module in Levelt's model, as also shown by Indefrey and Levelt's (2004) meta-analysis of neuroimaging studies (see also Hulstijn & Hulstijn, 1984; Morrison & Low, 1983). That is, the functionality of monitoring the execution of corrected utterances can be triggered by an external factor in the process of "making repairs upon the confirmation or clarification requests of the interlocutor" (Kormos, 2006, p. 135), the only difference being who perceives the error. Additionally, the perceptual theory of monitoring postulates that detecting linguistic errors in one's own speech (either covertly or overtly) is essentially the same as detecting errors in others' speech with "the same parsing procedures and sources of knowledge" (Levelt, 1983, p. 97). The monitoring theory, therefore, may apply well to an interactional context in which both interactants monitor their own speech as well as that of their conversational partner.

In sum, monitoring in tandem with dual models of skill acquisition may provide an additional way of understanding CF effectiveness. During interaction with a teacher or another learner, monitoring at the internal level may lead to self-initiated corrections. In the case in which a learner fails to edit the error on his or her own, CF substitutes the monitoring and helps the learner notice the gap. At this point, the comprehension system of the learner may be provided with relevant information, enabling comparison of the erroneous utterance (via recasts), or with an opportunity to test another hypothesis (via prompts). It may then be the case that prompts more than recasts engage the learner once again in the speech-production process in a way that activates another sequence of monitoring. This means that, at the same time, the learner is afforded an opportunity to engage in repeated practice, which leads to proceduralization. What is fruitful in this process for L2 acquisition is that, by producing MO, the learner may automatize more accurate grammatical knowledge. Additionally, in the case in which a learner is the one who detects errors in another's speech, this process may contribute to improving his or her monitor. On the basis of this theoretical perspective, the current study examines the effects of recasts and prompts on accuracy development.

Fluency Development

Research on CF within the framework of form-focused instruction was motivated to some extent by the results of Canadian French immersion programs in which learners fell short of achieving nativelike accuracy. Partly due to this historical development, researchers have continued to investigate CF efficacy mainly on learning grammar (see Segalowitz & Lightbown, 1999). However, when L2 development is accounted for by automatization, some change in the speed of processing should also be expected. This is because accessing procedural knowledge (or implicit knowledge; see R. Ellis, 2009b; Ortega, 2009, for discussion concerning procedural and implicit knowledge) requires fast retrieval from memory (DeKeyser, 2003; N. Ellis, 2002; R. Ellis, 2005; Hulstijn, 2005), which can be accounted for by the power law of practice (Newell & Rosenbloom, 1981). By engaging in practice, execution of a skill (e.g., accurate L2 speech) eventually becomes automatized such that attention is no longer necessary. Following these claims, some studies investigated CF efficacy on the development of implicit knowledge (e.g., R. Ellis, 2007; R. Ellis, Loewen, & Erlam, 2006) and also on changes in reaction time (e.g., Ammar, 2008; Lyster & Izquierdo, 2009). However, to the authors' knowledge, no studies have examined the effect of CF on fluency development in spontaneous speech (i.e., utterance fluency)—that is, on L2 performance rather than competence (see Douglas, 2001; Segalowitz, 2010; Towell & Dewaele, 2005).

Though a direct link between cognitive fluency (i.e., processing speed) and utterance fluency has not been established, an observable change in speech production is thought to reflect underlying cognitive processing (Derwing, Munro, & Thomson, 2008; Segalowitz, 2003; Segalowitz & Freed, 2004). Hulstijn and de Graaff (1994) also made a connection between implicit knowledge and oral output, claiming that fluent performance is the "phenomenological experience of not being aware of the way in which information is being processed and how these processes are monitored" (p. 98). Given that CF effectiveness may extend to fluency development by facilitating the automatization process that concerns the processing speed, the present study investigates the effects of CF on improving speech rate (SR), which has been shown to influence how listeners perceive fluency (Cucchiarini, Strik, & Boves, 2002; Lennon, 1990).

PEER INTERACTION AND L2 DEVELOPMENT

From a cognitive perspective, peer interaction has been examined primarily as conversational exchanges in which communication breakdowns trigger negotiation for meaning (e.g., Gass, 2003; Pica, 1994). Findings indicate that, in comparison to interaction with native speakers, learners interacting with other learners tend to engage in more such negotiations (Porter, 1986; Varonis & Gass, 1985), during which they use interactional moves claimed to benefit L2 development, such as input modifications (when learners are highly proficient: García Mayo & Pica, 2000) and interactional feedback (e.g., Adams, 2007; Gass & Varonis, 1989; Soler, 2002). This may lead to subsequent MO (e.g., Sato & Lyster, 2007). In addition to these interactional moves, learners tend to self-correct more while interacting with one another than when they interact with native speakers (e.g., Buckwalter, 2001; Sato, 2007; Shehadeh, 2001). Such self-corrections are considered to be "overt manifestations of the monitoring process" (Kormos, 2006, p. 123), hypothesized to facilitate L2 processing in the same way as MO triggered by CF (Kormos, 1999). Hence, if production practice is viewed as essential for L2 acquisition, peer interaction, in general, provides optimal conditions.

However, peer interaction has several weaknesses, especially in terms of grammatical development. First, although the directionality of error-treatment sequences is usually toward correctness (e.g., Bruton & Samuda, 1980; Porter, 1986), another type of negotiation whose purpose is to work on grammatical errors, as opposed to communication breakdowns—that is, negotiation of form (Lyster, 1994, 2002)—barely occurs unless the task itself is designed to make it happen (e.g., consciousness-raising tasks, Fotos & Ellis, 1991, and dictogloss, Kowal & Swain, 1997; but see Adams, 2007). Instead, learners often avoid negotiation and solely focus on task completion (e.g., Loschky & Bley-Vroman, 1993). Presumably, this is because they do not provide one another with interactional moves that indicate errors. Second, although some studies reported negotiation of form in peer interaction, showing that learners can pay attention to form and signal it to their partners (Mackey, Oliver, & Leeman, 2003; McDonough & Mackey, 2000; Sato, 2007), their interactional feedback is usually made up of simple segmentations of their partners' erroneous utterances (Pica, Lincoln-Porter, Paninos, & Linnell, 1996; Sato & Lyster, 2007). This is not quality feedback because it lacks a corrective force to signal that there is an error or to give opportunities to produce MO (e.g., Toth, 2008). Last but not least, learners' perceptions of one another may hinder the effectiveness of peer interaction. If learners do not believe in one another's linguistic ability, feedback may be missed, unnoticed, or ignored, and its effectiveness may be discarded (Yoshida, 2008).

In summary, peer interaction provides good opportunities for L2 learning because learners tend to produce a lot of language; that is, they engage in contextualized production practice. Their attention, however, rarely goes to linguistic form, and they do not indicate grammatical problems to one another unless they are provided with such guidance from the task. Thus, classroom peer interaction lacks several important elements to be conducive to L2 development: (a) autonomous attention to linguistic forms, (b) quantity and quality in feedback, and (c) positive perceptions of peer interaction itself. To overcome these weaknesses, the current study examines pedagogy in which learners are taught how to provide CF to one another during communicative activities.

INTERACTIONAL MOVES AND L2 DEVELOPMENT

Whereas the overall positive effects of CF on L2 development are well documented (e.g., Li, 2010; Lyster & Saito, 2010; Russell & Spada, 2006), what remains less established is the direct causal link between CF and L2 development. Does the CF move itself trigger the processing that alters interlanguage representations, or is the move following the CF (i.e., uptake) the trigger? For those who see CF as a type of input entailing new linguistic information or negative evidence that engages learners in a cognitive comparison, the effectiveness lies mainly in the psycholinguistic process of noticing. For example, Mackey (2006) showed that learners who noticed the corrective nature of feedback improved more than those who did not report noticing (but see Bigelow, delMas, Hansen, & Tarone, 2006). Another perspective, especially when discussing prompt effectiveness, is that pushed MO is linked to L2 development. There is an ongoing debate on whether or not uptake, which is a modification move that follows CF, can be a reliable measure of L2 learning. On the one hand, uptake following a recast might simply be a repetition of what has been provided by the interlocutor and so should not be considered as a predictor of acquisition but rather a facilitator of subsequent acquisition (Gass, 2003; Mackey & Philp, 1998; McDonough, 2005; McDonough & Mackey, 2006). On the other hand, Loewen (2005) found that successful uptake was a significant predictor of accuracy test scores, although Loewen and Philp (2006), drawing on the same database, later specified that this was the case for prompts but not for recasts (see also Nabei & Swain, 2002). The current study adds to this discussion by conducting correlational analyses on L2 development scores and interactional moves that encompass both CF and MO.

THE CURRENT STUDY

In this study, CF effectiveness was examined by explicitly teaching L2 learners how to provide CF to one another during meaning-focused activities. Recasts and prompts were separately investigated given their differential effects on processing (e.g., Yang & Lyster, 2010). The theoretical underpinnings of the current study entail monitoring and practice rather than noticing as defined in either Schmidt's noticing hypothesis or Swain's output hypothesis. This is because learners in the present study were both CF providers and receivers. Therefore, the monitor of not only the learner who received CF but also the learner who provided CF could be affected by noticing his or her partner's errors. Thus, the notion of a cognitive comparison between interlanguage and perceived input may not sufficiently account for the results from such a design. Moreover, the target population of the study had advanced grammatical knowledge and engaged in extensive practice activities during which CF was provided. Proceduralization is consequently a more suitable concept than that of learners noticing the gap and testing hypotheses as they produce comprehensible output. Finally, the study investigates CF effectiveness as changes in fluency via performance phenomena (i.e., temporal aspects of speech), which entails processing speed and, thus, automatization that results from practice. On the basis of these theoretical foundations, the following research questions were formulated:

1. Does CF embedded in peer interaction improve L2 accuracy and fluency? If so, how do recasts and prompts differentially contribute to development?
2. Does peer interaction alone improve L2 accuracy and fluency?
3. How are interactional moves associated with L2 accuracy and fluency development?

METHODOLOGY

Participants

This study was conducted at a reputable university in Japan. The Japanese learners of English in the present study had followed typical learning patterns of foreign language development: Years of language classes with grammar-translation methods had led them to become functional readers and writers yet poor speakers (see Miyagi, Sato, & Crump, 2009; O'Donnell, 2005). A background questionnaire—a modified version of the language contact profile developed by Freed, Dewey, Segalowitz, and Randall (2004)—was administered at the beginning of the semester, which confirmed that no students in the current study had had excessive exposure to English (e.g., being part of a bilingual household or living abroad for a long time) or had been taught English through the medium of English prior to the classes in which the intervention was implemented. These uniform characteristics of the participants' learning background and proficiency level can be partly explained by the administrative system that exempted high-proficiency learners from taking English classes based on their TOEFL scores—that is, 500 points in the paper-based test. Because the present study conceptualizes grammatical knowledge as declarative and spontaneous production as procedural knowledge, this group of learners was considered ideal for a study designed to examine the effects of peer interaction and CF on proceduralization.

Participants were majoring in economics, business, human development, or biology and were enrolled in one of four sections of a required second-year English course (N = 167; mean age = 19.5). Two of the classes were taught how to provide CF to one another: Learners in the peer-interaction-plus-prompt class (PI-prompt; n = 41) were taught how to provide prompts, and the peer-interaction-plus-recast class (PI-recast; n = 46) was taught how to give recasts. Another class was given peer-interaction activities only (PI-only; n = 42). A fourth class served as the control group (n = 38). Therefore, the design teases apart the effects of (a) peer interaction, (b) CF, and (c) types of CF.

Classroom Intervention

The classes were team taught by two male teachers. One was an American who had more than 20 years of teaching experience, and the other (the first author) was Japanese and had 5 years of teaching experience and nativelike English proficiency; the latter teacher taught all classes except the control class, in which he was merely an observer. Conducted entirely in English, each class was held for 1.5 hr per week over a 10-week period. During each class, an average of 60 min were devoted to the intervention (totaling 10 hr), and the rest of the time was used for other regular activities specified in the course syllabus (e.g., writing reports and preparing for exams). The three experimental groups were given peer-interaction instruction and engaged in fluency-focused activities, and, additionally, the PI-prompt and PI-recast groups received CF training. During the first 3 weeks, an additional 10 min were used for the CF training (see Figure 1).

Peer-Interaction Instruction. A peer-interaction activity, based on Nation's (1989) fluency-focused activity, was administered for approximately 40 min every week throughout the intervention period, starting from the first week. This was an information-exchange activity in which learners worked in pairs and changed partners multiple times. Each time they had new partners, they needed to tell their new partner what their previous partner had said, all in the same amount of time. Thus, although the delivery time did not change, the amount of information

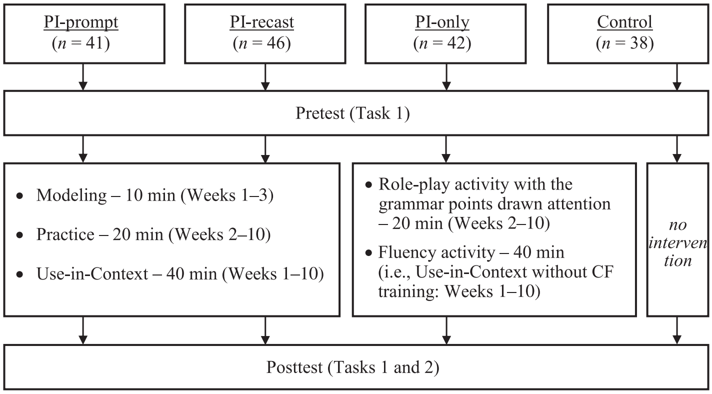

Figure 1. Outline of the intervention.

they needed to communicate kept increasing, requiring them to speak faster as the activity developed. For instance, in one of the classes, a 6-min scene from the movie *Babel* (Iñárritu & Arriaga, 2006) was divided into three parts. The class was also divided into three groups, each of which was responsible for one of the three parts. For the first two 3-min rounds, students retold what they had seen to students who were responsible for the same segment. For the third and fourth rounds, they switched to partners whose parts were different from theirs. They were also asked to include their opinions and interpretations of the movie segment. Therefore, by the end of the final round, they had talked about all three parts of the scene with the different descriptions and opinions of 12 people. Other materials included music videos and songs, into which the same activity format was integrated. This type of activity is intended to (a) draw learners' attention to meaning instead of form as they need to convey the information from the previous partners in each round, (b) develop their confidence as they repeat the same sentences several times, and thus (c) engage them in practice that contributes to automatization (DeKeyser, 2001; de Jong & Perfetti, 2011).

CF Training. Because providing CF was expected to be new to the students, the present study followed a well-established instructional framework whereby CF was taught through a sequence of preparation, practice, and expansion (Chamot, Barnhardt, El-Dinary, & Robbins, 1999; Oxford, 1990). These three stages were correspondingly designed as modeling, practice, and use-in-context.

In the modeling phase, which lasted for the first 3 weeks for 10 min per class, the two teachers demonstrated a mini role play in which they provided CF to each other (either prompts or recasts, depending on the group). One of the teachers made errors and the other provided CF, taking care to ensure that the responses to CF were evenly distributed in terms of modification moves. That is, about half of the uptake that followed CF entailed MO, whereas the other half involved repetition of the same error, so that the learners were trained how to give CF rather than how to respond to CF. In the recast class, students were simply told that CF is a reaction to grammatical errors that supplies the correct version, although there were degrees of explicitness in the recasts shown during the role-plays (see Egi, 2007; R. Ellis & Sheen, 2006; Lyster, 1998; Nicholas et al., 2001; Sato, 2011; Sheen, 2006). In contrast, students in the prompt group were told that the role of CF is to let the interlocutor know that he or she made an error and to give a chance to self-correct. More specifically, they were shown two types of prompts: clarification requests (e.g., *Pardon?*) and metalinguistic feedback (e.g., *You need past tense.*). These two types were chosen to parallel possible variation in the explicitness of recasts. Nonetheless, all CF happened to be relatively explicit because of its pragmatically unnatural occurrence in the context of peer interaction. The two CF types, therefore, can be best distinguished in the present study along the dimension of input-providing and output-promoting feedback. The role-play presented by the two teachers in the modeling stage followed the same format as the student role-plays in the practice stage, which is described next.

In the practice stage, learners were given a role-play scenario, which had been designed to have them practice giving CF. This activity took approximately 20 min per class and was implemented from Week 2 to Week 10. First, the class was divided into groups of three. Each member was either the speaker, feedback provider, or observer, and they took turns playing each role. In addition to a scenario, they were given a list of sentences, each of which contained an error that they had to include during the role-play. There were 10 different scenarios and accompanying error lists, each focusing on a different linguistic feature (e.g., past tense, subject-verb agreement, pronouns; see Appendix A). Each learner in the group had a different error list because detecting errors would otherwise have been too easy. A few minutes were given as planning time to create original stories that incorporated the given error sentences and that included their own sentences so as to make the errors harder to detect. While the speaker was telling a story, the feedback provider's role was to detect the errors and give CF. The observer checked which errors were detected or missed and gave a report to the group after the role-play. Learners in the PI-only group did not engage in these same practice tasks but were given a similar peer-interaction activity in which the same grammatical features as those in the scenarios were given attention. This was done (a) to equalize time spent on the task across groups and (b) to avoid possible effects of the grammar-teaching element in the practice stage (i.e., selected grammatical features were highlighted in the scenarios) and, thus, to tease apart the effects of CF.

Finally, in the use-in-context stage, students in the PI-prompt and PI-recast groups were encouraged to use the CF technique in a more authentic communicative context—that is, in the activity given to all the experimental groups (see the Peer-Interaction Instruction section). Although the practice activity was rather decontextualized and often shifted attention exclusively to form, students' attention was primarily drawn to meaning at the use-in-context stage by providing them with authentic visual materials and embedding an information-exchange element. Thus, students in the CF groups engaged in the sequence of modeling, practice, and use-in-context, whereas the PI-only group engaged in grammar-focused role-play activities and the use-in-context activity without any embedded CF training.

Testing and Scoring

To assess L2 development and interactional patterns during peer interaction as well as to investigate relationships between interaction and development, speech data were elicited by means of two different tasks. Task 1 was administered at both the beginning and the end of the semester, whereas Task 2 was administered at the end of the semester. These tasks were implemented in a classroom where computers and headsets were available to all participants and used interactive learning software called *CaLabo* to collect data from more than 40 participants at the same time. The speech samples were transcribed using the CHAT transcription and coding format under the CHILDES system (MacWhinney, 2009).

Task 1: L2 Development. To elicit learners' spontaneous production, a picture-description task was employed in which they were asked to narrate an event in chronological order. Similar but different sets of pictures were used for the pretest and posttest (i.e., the same characters but different events). Participants were given 2 min to plan, but, to obtain pure speech samples, they were not allowed to take notes. The first minute of each one of the 3-min samples (from the beginning of a given speech sample to the end of the clause closest to the 1-min point: M = 59.11 s) was used for scoring to equalize data sources for accuracy and fluency.

Whereas classroom experimental CF studies have targeted specific linguistic features—for example, possessive determiners (Ammar & Spada, 2006), past-tense forms (R. Ellis et al., 2006), grammatical gender (Lyster, 2004), and articles (Sheen, 2007)—overall accuracy scores were used in the present study for several reasons. First, the classroom intervention in this longitudinal study was given throughout one academic semester. Therefore, focusing on multiple grammatical structures was thought to be more realistic to meet educational objectives. Second, the participants had good grammatical knowledge, and the intervention was designed to foster conditions for proceduralization of their explicit knowledge. That is, CF was expected to promote not only noticing of specific features but also the monitoring function in general. Third, as learners engaged in many communicative activities, practice effects were hypothesized to emerge in different aspects of production that would not be captured if only one grammatical structure were targeted. Following Foster and Skehan's (1996) procedure, overall accuracy was defined as freedom from error operationalized as error-free clauses in which "there is no error in syntax, morphology, or word order" (p. 304). Scores were calculated as percentage scores representing the ratio of correct clauses to the total number of clauses. Another researcher independently scored 15% of the data set: The interrater reliability using Cronbach's alpha yielded $\alpha = .89$.

Temporal measures of oral fluency were employed in the current study for two main reasons. First, SR has been found to be strongly related to perceived fluency (e.g., Cucchiarini et al., 2002), and, second, it can be a measurable performance phenomenon considered to manifest underlying cognitive processing (Segalowitz, 2010). Two types of SRs—namely, unpruned and pruned SRs—were calculated on the basis of Lennon's (1990) finding that pruning speech samples gives more reliable scores of fluency. Unpruned SRs were obtained by dividing the total number of words by the total delivery time in milliseconds. The number was then multiplied by 60 to obtain words-per-minute scores. Pruned SRs were calculated in the same manner but after (a) subtracting the sum of filled (e.g., *mm*, *ah*) and unfilled pauses (the cutoff was 200 ms) from the delivery time and (b) excluding repetitions and hesitation markers from the word count. In the current study, self-corrections were not considered as dysfluency markers, following suggestions that they are unrelated to perceived fluency (e.g., Freed, 1995; Iwashita, Brown, McNamara, & O'Hagan, 2008).

Task 2: Interactional Moves. Task 2 was designed to examine learners' interactional patterns and employed a decision-making task to provide a communicative interactional context (Pica, Kanagy, & Falodun, 1993). They were given a short description of a famous person who had certain problems to deal with (e.g., the prime minister of Japan) and asked to decide, in pairs, on a gift for him. Three minutes of individual planning time were given, and they were asked to agree on their gift in those 3 min.

Five interactional moves were accounted for, all of which have been claimed to be related to L2 development. First, CF entailed negotiation of grammatical errors—that is, negotiation of form (Lyster, 2002)—and included the two types of feedback that learners had been taught to use (i.e., recasts or prompts). Second, interactional feedback (IF) entailed negotiation for message comprehensibility problems—that is, negotiation for meaning (Long, 1996). Third, MO following CF entailed repair moves containing more accurate versions of erroneous utterances, and, fourth, MO following IF contained more comprehensible versions of problematic utterances. Fifth, self-initiated MO (SMO) was identified when a learner successfully corrected a grammatical error on his or her own without any provision of CF, following Kormos's (1999) definition of SMO as instances in which the learner detects that "the output has been erroneous or inappropriate, halts the speech flow, and finally executes a correction" (p. 315; see also Sato, 2007). Using the Kappa statistics, interrater reliability was calculated by having the second coder identify the five interactional moves in 25% of the interaction data (i.e., 78 instances). The reliability score reached Kappa = .83 ($p < .001$), which is considered as "almost perfect" agreement (Landis & Koch, 1977, p. 165).

Statistical Analysis

Statistical analyses were conducted using version 17.0 of the Statistical Package for Social Sciences (SPSS). First, to examine differential effects of the intervention on L2 development among the four classes ($N = 147$; learners who did not take the posttest were excluded), three ANCOVAs were conducted on the three types of scores from Task 1, using the pretest scores as covariates. All the statistical assumptions were met, including the assumption of regression of slopes, meaning that ANCOVAs were initially run with models that included interaction between the independent variable and covariate. After confirming that there were no significant interaction effects, another set of ANCOVAs was computed that excluded interaction. The results from the second set of ANCOVAs will be reported.

For interactional moves (i.e., Task 2), a MANOVA was employed, rather than a repeated-measures ANOVA or multiple univariate ANOVAs for each move. Although MANOVAs are not conventionally used in L2 research (Larson-Hall, 2010), it was thought appropriate to investigate group differences (i.e., the independent variable) in different interactional moves (i.e., the five dependent variables), given that (a) the intervention may affect all moves together, (b) all the moves arguably contribute to L2 development according to the literature, and (c) the CF was provided not by the learner producing MO but nonetheless by a learner in the same group (see Keselman et al., 1998; Stevens, 2002). To meet one of the assumptions of the parametric test—that is, independence of observations (Hatch & Lazaraton, 1991)—mean scores were calculated from each pair. That is, although there were 148 participants, the sample size was 74. For post hoc single comparisons for analyses of variance (ANCOVAs and MANOVAs), Bonferroni corrections were made (Kappel, 1991), and the alpha level for all tests of significance was set at .05.

Finally, to explore the relationships between interactional moves and L2 development, three 6 × 6 correlation matrices were created with Pearson correlation coefficients (two-tailed), using gain scores of the three analyses of speech (i.e., accuracy, unpruned SR, and pruned SR) for each individual (n = 147). Statistically significant correlations were decided after adjusting α levels with a modified Bonferroni procedure. More specifically, the Holm procedure was used to reduce the likelihood of Type I error. In this procedure, obtained p values are ranked from most significant to least significant and then compared with adjusted significance levels obtained by dividing the original α level (i.e., .05) by the number of comparisons for each level, which sequentially decreases. In this study, 15 comparisons were made, so the first adjusted α level was .0033, the second was .0036 (.05 divided by 14), the third was .0038 (.05 divided by 13), and so forth. These comparisons continued until an observed p value exceeded the corresponding adjusted α level (Olejnik, Li, Supattathum, & Huberty, 1997). To determine which interactional moves were most closely related to L2 development, a multiple regression analysis was originally thought to be appropriate (see Loewen, 2004; Loewen & Philp, 2006); however, simple bivariate correlational analyses were chosen because of the high correlations between some interactional moves (multicollinearity)—namely, between feedback and the following modification moves.

RESULTS

Effectiveness of CF and Peer Interaction

Table 1 presents the group means and standard deviations of the three measurements for each condition over time, and Figure 2 displays the group means graphically. The ANCOVA of the accuracy scores yielded a significant group effect, $F(3, 142) = 43.783$, partial $\eta^2 = .481$, $p < .001$, revealing significant differences among the four groups at the posttest. The Bonferroni post hoc comparisons revealed that the PI-prompt group outperformed both the PI-only group (d = 1.69) and the control group (d = 2.11) and that the PI-recast group outperformed the PI-only (d = 1.73) and control (d = 2.16) groups, with large effect sizes (Cohen, 1988). No significant difference was detected between the two CF groups. The ANCOVA on the unpruned SRs also detected a significant main effect, $F(3, 142) = 6.930$, partial $\eta^2 = .128$, $p < .001$. The post hoc analyses detected significant differences between the three experimental

Table 1. Group means and standard deviations for accuracy scores and unpruned and pruned speech rate (SR)

		Pretest		Posttest	
Groups		*M*	*SD*	*M*	*SD*
PI-prompt (n = 39)	(accuracy)	29.73	22.44	66.07	24.34
PI-recast (n = 40)		22.48	21.73	66.20	23.44
PI-only (n = 34)		23.49	17.70	27.35	21.73
Control (n = 34)		23.24	22.31	19.55	19.81
PI-prompt	(unpruned SR)	43.75	12.96	50.80	10.87
PI-recast		41.16	9.78	49.94	10.32
PI-only		43.94	12.87	50.77	9.74
Control		44.00	11.67	41.00	15.73
PI-prompt	(pruned SR)	111.81	26.03	133.91	23.62
PI-recast		109.13	19.67	134.42	20.33
PI-only		112.91	14.36	130.38	21.17
Control		114.37	19.44	98.10	19.41

groups and the control group, with all three treatment groups outperforming the control group (PI-prompt: d = .74; PI-recast: d = .69; PI-only: d = .76) but with no significant differences among them. Similar results were obtained from the pruned SRs. A group effect reached significance, $F(3, 142) = 25.947$, partial $\eta^2 = .354$, $p < .001$, and all treatment groups proved significantly different from the control group (PI-prompt: d = 1.67; PI-recast: d = 1.85; PI-only: d = 1.61). Effect sizes of the SR results indicated that the instruction had a larger impact on the pruned SRs than on the unpruned SRs.

Interactional Moves

Pillai's Trace Criteria was chosen for the multivariate F test because the assumption of homogeneity of covariances was violated. This test was statistically significant, $F(15, 204) = 4.081$, partial $\eta^2 = .231$, $p < .001$, and indicated that the groups differed when all interactional moves were combined and compared with one another. The univariate tests detected significant main effects in CF, $F(3, 70) = 27.821$, partial $\eta^2 = .544$, $p < .001$; MO following CF, $F(3, 70) = 31.724$, partial $\eta^2 = .576$, $p < .001$; and SMO, $F(3, 70) = 12.961$, partial $\eta^2 = .357$, $p < .001$. These three moves were thus submitted to post hoc analyses using Games-Howell for pairwise comparisons because the tests of homogeneity of variances reached significance. The single comparisons revealed that, regarding CF, the PI-prompt group was significantly different from the PI-only (d = 1.16)

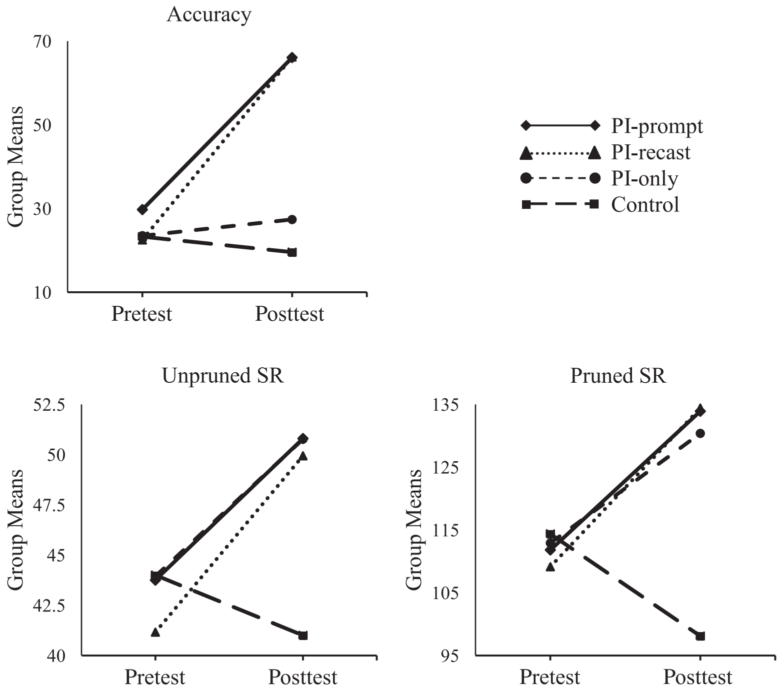

Figure 2. Group means on accuracy scores, unpruned SRs, and pruned SRs over time.

and control (d = 1.72) groups. The PI-recast group outperformed all the other groups (PI-prompt: d = .90; PI-only: d = 2.20; control: d = 2.83). The contrast between the PI-only and control groups was not significant. As for MO following CF, exactly the same patterns emerged: PI-prompt and PI-only: d = 1.40; PI-prompt and control: d = 1.92; PI-recast and PI-prompt: d = .97; PI-recast and PI-only: d = 2.30; PI-recast and control: d = 2.76. Finally, post hoc tests for SMO revealed that both the PI-prompt and PI-recast groups outperformed the PI-only and control groups: PI-prompt and PI-only: d = 1.11; PI-prompt and control: d = 1.54; PI-recast and PI-only: d = 1.35; PI-recast and control: d = 1.69. The PI-only group was not statistically different from the control group. The group means and standard deviations appear in Table 2, and Figure 3 displays the means of the interactional moves.

Relationships between Interactional Moves and L2 Development. Pairs of variables that yielded p values smaller than .05, correlation coefficients, effect sizes, and adjusted significance levels are displayed in Table 3. The correlation analyses detected seven significantly correlated pairs when accuracy gain scores were included—six in the matrix with

Table 2. Group means and standard deviations for interactional moves

	IF		CF		MO-IF		MO-CF		SMO	
Groups	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
PI-prompt (*n* = 19)	.32	.42	.73	.61	.29	.42	.63	.47	1.63	.96
PI-recast (*n* = 20)	.30	.50	1.28	.64	.30	.50	1.13	.58	2.10	1.27
PI-only (*n* = 17)	.38	.45	.18	.30	.26	.40	.12	.22	.76	.59
Control (*n* = 18)	.36	.51	.00	.00	.25	.43	.00	.00	.50	.45

Note. MO-IF = MO following IF; MO-CF = MO following CF; SMO = self-initiated MO

unpruned SRs and six in the matrix with pruned SRs. After adjusting α levels, however, there remained five for accuracy, three for unpruned SRs, and five for pruned SRs (marked with asterisks in the table). Again, owing to multicollinearity, feedback and the following modification moves were found to be highly correlated. Although relationships between feedback and immediate learner responses are of theoretical significance, our primary focus in this study is on interactional moves in relation to L2 development.

The interactional moves found to be correlated with accuracy development were MO following CF, $r = .31$, $df = 145$, $p < .000$, and CF, $r = .31$, $df = 145$, $p < .000$, with medium effect sizes. These positive correlations indicate that (a) greater frequency of MO after CF led to larger differences between pretest and posttest scores on accuracy and (b) greater frequency of CF led to more improvement. Regarding the two types of SRs, none of the interactional moves was found to be significantly correlated with unpruned SRs. However, two moves reached significance with pruned SRs: MO following CF, $r = .32$, $df = 145$, $p < .001$, and CF, $r = .31$,

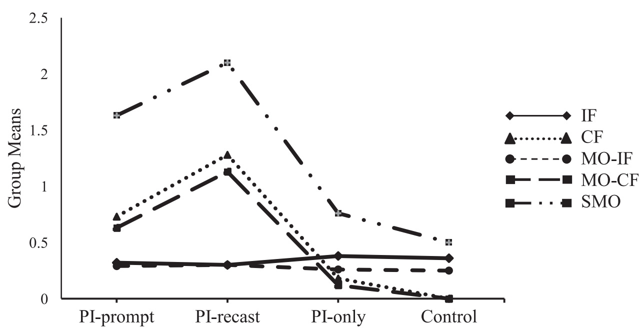

Figure 3. Group means on interactional moves at the end of the semester.

Table 3. Correlations between interactional moves and L2 development

Correlations		*r*	R^2	*p* value	Adjusted *α*
CF × MO-CF	(accuracy)	.97	.94	.000*	.0033
IF × MO-IF		.92	.85	.000*	.0036
Accuracy × MO-CF		.31	.10	.000*	.0038
Accuracy × CF		.31	.10	.000*	.0042
CF × SMO		.27	.07	.001*	.0045
MO-CF × SMO		.23	.05	.005	.0050
Accuracy × SMO		.20	.04	.020	.0056
CF × MO-CF	(unpruned SR)	.95	.91	.000*	.0033
IF × MO-IF		.88	.77	.000*	.0036
CF × SMO		.27	.07	.001*	.0038
MO-CF × SMO		.23	.05	.005	.0042
SR(Unp) × MO-CF		.21	.04	.009	.0045
SR(Unp) × CF		.19	.04	.023	.0050
CF × MO-CF	(pruned SR)	.95	.91	.000*	.0033
IF × MO-IF		.88	.77	.000*	.0036
SR(Pru) × MO-CF		.32	.10	.000*	.0038
SR(Pru) × CF		.31	.10	.000*	.0042
CF × SMO		.27	.07	.001*	.0045
MO-CF × SMO		.23	.05	.005	.0050

Note. * = significantly correlated pairs.

$df = 145$, $p < .001$. These results indicate that the more a learner received CF and the more MO he or she produced following CF, the greater the extent to which his or her pruned SR improved.

Summary of Results

In answer to the first and second research questions, CF embedded in peer interaction has positive impacts on accuracy development. Given that all treatment groups—including the PI-only group—improved unpruned and pruned SRs, it can be said that peer interaction contributes to fluency development. Effect size analyses revealed that CF had large effects for accuracy, medium effects on unpruned SRs, and large effects on pruned SRs.

The preliminary analyses for the third research question (i.e., frequencies of interactional moves) found, first, that the instructional treatment had a significant impact on the frequency of interactional moves in general. Second, the CF groups, as expected, produced more CF than the PI-only and control groups; also, the recast group produced more than the prompt group. The CF groups also produced more SMO. The correlational analyses revealed that (a) CF is related to accuracy development, (b) none of the interactional moves was related to change in unpruned SR, and (c) CF and following MO are related to development of pruned SRs.

DISCUSSION

CF and L2 Development

The current study provides evidence of CF efficacy on accuracy development via improved monitoring. Given that peer interaction alone contributed to fluency development, it may seem that CF is effective only on accuracy development. It is important to note, however, that CF did not impede fluency development but rather facilitated monitoring, which contributed to both more accurate and faster processing. That is, giving CF to one another triggered an ideal proceduralization loop in which learners started to rely on procedural knowledge during spontaneous production. The learners in the present study already had adequate grammatical knowledge at the beginning of the intervention, and proceduralization was observed in increased fluency. Corrective feedback enabled learners to reassess nontarget structures retrieved from long-term memory by reprocessing them in working memory with the help of declarative knowledge. It is necessary to note that the learners in this study were not only CF receivers but also providers. Given this design, the results indicate that the monitor of the learners in the CF groups became more sophisticated. In Levelt's model, which is not purported to have any developmental implications, improvement in accuracy can be explained as a result of sophistication in the prearticulatory monitoring loop. Those learners came to detect errors and edit their potential erroneous utterances before finally articulating speech. It can also be said, by adding a developmental perspective, that proceduralization was accelerated in the formulator where syntactic computation is executed.

From a SLA perspective, this result may suggest that learners receiving CF succeeded in engaging in focus on form (see Loewen, 2011, for a review of focus on form). That is, learners are more able to draw their attention to form while maintaining their primary focus on meaning when (a) their explicit knowledge exceeds their automatized online production skills, provided they have sufficient memory span to shift their attention to linguistic forms (see Green & Hecht, 1992; VanPatten, 1990), and (b) the given task is not cognitively too demanding (see Bygate, 1996; Seedhouse, 1997). The fact that their fluency improved as well indicates that CF from other learners does not impose a heavy cognitive load; rather, it helps them to establish accurate form-meaning mappings (VanPatten, 2000). Some suggest that CF is effective only on accuracy development because it breaks the communicative flow (Harmer, 1998), and many teachers believe this claim (Basturkmen, Loewen, & Ellis, 2004). Though the present study did not investigate whether or not peer CF interrupts communication, it showed that CF does not impede fluency development.

The only variable causing fluency development in the PI-only group is practice. They were given peer-interaction activities in which they repeatedly produced output. It is relevant to discuss this finding in light of the output hypothesis, which claims that learners notice their errors by producing comprehensible output. The results of the present study showed, instead, that mere engagement in peer-interaction activities is not enough to restructure interlanguage. It is noteworthy that learners in the PI-only group were told to pay attention to the same grammatical features on which the learners in the CF groups gave CF to one another. This instruction alone was insufficient to facilitate their noticing process in the current study. This finding suggests that the extent of noticing and possible opportunities for hypothesis testing can be enhanced by CF (Adams, 2007). That is, just like teachers' CF, peer CF offered external support that helped learners to improve their monitoring and created opportunities to engage in practice.

It is important to reiterate here that the learners in the CF groups were also CF providers; the fact that they noticed errors in their classmates' speech may have contributed to their accuracy and fluency development. It may seem, therefore, that the functions of noticing specified in the output hypothesis and the noticing hypothesis may account for the grammatical development observed in the present study. However, these hypotheses are only applicable insofar as learners received CF and engaged in production, which together may have encouraged cognitive comparison of the input and their own output. The results of the current study suggest instead that the combination of monitoring and repeated practice provides a more convincing explanation, as it accounts for (a) noticing in both comprehension (decoding) and production (encoding) and (b) the practice effect observed in fluency development.

In terms of types of CF, the findings of the present study do not conform to previous teacher CF studies (e.g., Ammar, 2008; Ammar & Spada, 2006; Loewen & Philp, 2006; Lyster, 2004; Sheen, 2007). Recasts and prompts functioned in a comparable way by contributing to accuracy development to a similar extent. According to previous CF accounts, two possibilities are conceivable. First, the learners in both groups had access to input containing information that enabled them to notice their nontarget output. Second, in addition to opportunities to notice their errors, learners in both groups had opportunities to repair them after the CF. Neither of these possibilities adequately accounts for the findings, however, precisely because development was measured on overall accuracy. That is, it is implausible that learners received CF on all types of erroneous utterances that they repaired and, as a result, learned how to produce correctly by the time of the posttest. A better explanation seems to be changes in the quality and speed of their monitoring, in the course of which the accuracy and speed of retrieving their explicit knowledge improved.

Additionally, the interaction data indicate a possible practice effect. The learners in the CF groups modified their erroneous utterances after CF most of the time regardless of types. In the PI-prompt group, 24 out of 28 prompts (85.71%) were followed by MO, and 45 out of 51 recasts led to successful MO in the recast group (88.24%). These high MO rates may be explained by certain characteristics of peer interaction: Learners were in a comfort zone in which they had more time to think and to correct their initial erroneous utterances (Sato & Lyster, 2007). It is again improbable that development in overall accuracy is entirely due to CF or MO. However, if the interactional data collected at the end of the semester are considered indicative of how learners provided and responded to CF during class over the semester, MO can be considered equivalent to a systematic production practice that contributed to L2 development. Therefore, these findings suggest that practice effects were observed in two ways. First, learners in the experimental groups all engaged in repeated practice during the peer-interaction activities for one semester, which resulted in fluency development. Second, learners in the CF groups had opportunities to practice correct forms after receiving CF (irrespective of CF type) and thus developed both fluency and accuracy. Their monitoring abilities may have been enhanced in such a way as to utilize existing grammatical knowledge rapidly in spontaneous speech (Hulstijn, 2002) or implicit knowledge that previously existed as explicit knowledge (DeKeyser, 2003), as reflected in overall accuracy development. Another way to explain the improvement is that learners proceduralized their declarative knowledge, which seems to be a more suitable concept to describe L2 development. Furthermore, if in fact MO accounts for more accurate and faster production, this study may add a possible explanation for teacher CF studies that showed greater effectiveness for prompts over recasts (but see DeKeyser, 2010, for an interpretation of recasts as receptive practice).

Interactional Moves and L2 Development

The first finding, in terms of interactional moves, was that L2 learners can be trained to become CF providers and that, unless they are explicitly instructed to do so, they do not engage in negotiation of form (see Appendix B for examples of peer CF). It is important to remember that the communicative task in this study was purely meaning based, according to Williams's (1999) task continuum, and it did not affect learners' interactional patterns as expected. The reason that learners in the PI-only group did not signal grammatical errors to their partners is not known. It might have been socially awkward to point out classmates' errors. They might not have known how to point out errors, or their monitor might not have been functioning to the same extent as the monitor of those whose instructional treatment encouraged its use. In other words, the language awareness of those who were trained to give CF—and thus received CF—was increased. Additionally, the frequency of recasts was greater than that of prompts, and MO was more frequent after recasts than after prompts—although these differences were not reflected in the L2 development scores. This does not conflict with findings in teacher CF research: Teachers tend to provide recasts more than prompting types of CF (e.g., R. Ellis et al., 2001; Lyster & Ranta, 1997). It is likely the case that, for pragmatic reasons, it was also easier for learners to give recasts than to give prompts.

The change in monitoring by the CF groups is especially apparent in the results pertaining to the amount of SMO. Learners who were taught how to provide CF corrected grammatical errors on their own more than those who were not, which shows that they autonomously attended to form during spontaneous speech. From these three interactional moves (CF, MO, and SMO), we speculate that learners' monitoring was activated both internally and externally. This claim is further supported by the finding that learners' SMO was significantly correlated with the frequency of CF given by their partners. That is, SMO and MO prompted by CF may entail cognitively similar processing (Kormos, 2006). It is interesting that these learners developed not only accuracy but also fluency. In the literature, there are mixed findings as to whether self-corrections are a dysfluency indicator (see Freed, 1995; Iwashita et al., 2008; Lennon, 1990; Riggenbach, 1991). Because self-corrections were neither excluded when pruning speech samples nor examined against SRs within the samples, this study does not contribute to this discussion directly but does indicate that learners who learned to self-correct became more fluent than learners whose monitoring was not the target of any instruction. Further investigation into the causal relationship between SMO and L2 development is warranted.

It is interesting to discuss this interactional move in terms of Anderson's skill acquisition model, according to which fluent performance requires proceduralized knowledge because retrieval from declarative knowledge through interpretive mechanisms is slow and limited in capacity. That the learners produced SMO yet developed overall accuracy and fluency may indicate that learners' grammatical knowledge was not yet fully proceduralized (also known as partial proceduralization; see DeKeyser, 2010). With respect to Anderson's three stages of proceduralization, some of their knowledge may have been in between the autonomous and associative stages in which they were still depending on partly declarative and partly procedural knowledge. As is the case in other cognitive research, it is methodologically difficult to identify at what point SRs are fast enough to claim that learners are using automatized knowledge (see Segalowitz & Segalowitz, 1993).

Of particular interest in terms of relationships between interactional moves and L2 development is how the two types of negotiation are related to L2 development. Whereas negotiation of form sequences entailing CF was significantly correlated with gains in both accuracy and fluency scores, negotiation for meaning was not related to L2 development at all. Interaction studies in general have attributed L2 development to negotiation for meaning because it creates more comprehensible input, interactional feedback, and MO. However, this study suggests otherwise: Merely working on communication breakdown seems to be less influential on L2 development than collaboratively working on grammatical problems (see Aston, 1986; Foster, 1998). As for accuracy, negotiation for meaning in the present study did not contribute to development insofar as only the CF groups outperformed the control group, even though they did not engage in negotiation for meaning significantly more than the PI-only group. For fluency development, one may claim that the improvement of the PI-only group was due to negotiation for meaning; however, again all treatment groups engaged in negotiation for meaning to a comparable degree. Given also that negotiation for meaning was not correlated with the gain scores, we can therefore conclude that it is not related to fluency development either. In contrast, the frequency of negotiation of form sequences is related to L2 development. This development must be due in part to the negotiation of form or some processing routine manifested in the frequency of negotiation of form. More precisely, it can be claimed that proceduralization resulted from the repeated practice afforded by both peer interaction and peer CF.

Finally, an additional finding in terms of interactional moves and L2 development was the difference between unpruned and pruned SRs. The analyses did not find any significant correlations between gains in unpruned SRs and interactional moves; in contrast, CF and the following modification moves were related to gain scores of pruned SR. Therefore, it can be argued that these two SRs measure different constructs. More specifically, if quantification of CF and MO frequencies taps into L2 learners' cognitive processing, then pruning speech samples may provide more reliable information about the underlying cognitive processing of spontaneous speech production. It can be further argued that CF and peer interaction contribute not only to performance but also to cognitive fluency development. Caution is needed, however, because correlations do not indicate any causality. That is, it may be the case that those who became faster in their cognitive processing due to the practice effect produced more MO in response to CF.

LIMITATIONS

The combination of monitoring and practice used in the present study to interpret CF efficacy has certain limitations insofar as it may be restricted to interactional contexts in which learners are not only CF receivers but also providers. As argued by de Bot (1992), there are some issues when applying this theory to L2 speech (e.g., it is not known whether the same production system is employed in first language [L1] and L2 production). Related to this is that the monitoring theory accounts for parallel (fast) processing, and thus it may not be the ideal model for the learners in the present study, who are still struggling to access their grammatical knowledge in spontaneous production. Finally, it is important to remember that the participants had adequate grammatical knowledge on which to draw to provide CF and to produce MO not only after recasts but also after prompts. This design raises the question as to whether the instruction would be effective for learners without explicit knowledge because, according to the declarative-procedural model, “there needs to be knowledge to be practiced” (DeKeyser, 2010, p. 161). The current study suggests that this theoretical approach adds another way of accounting for CF effectiveness to the extent that (a) it applies well to both receiving and providing CF and (b) it explains not only MO but also SMO. Further studies are warranted to test this claim.

Additionally, this empirical study, with its novel instructional approach, had some design issues. For example, the interaction data were collected outside the classrooms at the end of the intervention because it was too noisy to collect individual speech samples from approximately 20 learner pairs concurrently engaged in communicative tasks (see Sato & Perry, 2011). Moreover, the data were obtained from interactions that lasted only 3 min, thus yielding low frequencies. In view of these limitations, the present study calls for future research in which interactional moves are cumulatively collected over time and examined against L2 gain scores. For the same reason, the present study is limited in drawing any definitive conclusions about practice effects resulting from CF. Finally, it must be noted that accuracy was measured as overall percentages of correct clauses. Future research will benefit by conducting more focused linguistic analyses.

CONCLUSION

In this study, L2 learners were trained to be CF providers and then given communicative tasks to use the strategy. This study distinguishes itself, however, from the strategy-training and task-based language teaching research in several ways. First, the primary purpose of the study was to examine CF as a way of enhancing L2 development rather than examining (a) mastery of the strategy itself (Naughton, 2006), (b) change in participatory patterns (Bajarano, Levine, Olshtain, & Steiner, 1997), or (c) the relationship between strategy use and proficiency (Lai, 2009; Lam, 2009). Second, the communicative task in this study was not designed to elicit any particular linguistic features: It did not give learners any new linguistic information, nor did it help them reach explicit understanding of linguistic properties, nor did it require specific structures for the task to be completed. Instead, a fluency-based task was employed in this study for promoting practice in which CF was integrated to help learners attend to form.

Given these differences, the current study contributes to SLA research in several ways. Although there are studies in which CF was taught as a peer-interaction strategy (Cohen, Weaver, & Li, 1998; Dörnyei, 1995; Nakatani, 2005), this is the first study to examine its effectiveness on L2 development with fine-grained analyses drawing on SLA methodology. The study also offers a new perspective for CF effectiveness drawing on a speech-production model and skill-acquisition theory. Moreover, the study is the first to examine CF efficacy on the development of utterance fluency, which was measured over time analyzing temporal aspects. Particular to this study is that instructional intervention rather than naturalistic exposure to the target language was the variable causing developmental change. Additionally, the study revealed, through experimental design, that peer interaction has a positive impact on fluency development and thus offers insight into the relationship between specific interactional moves and L2 development.

Conducted in intact classrooms (university-level required English classes) over one academic semester, this study has strong ecological validity. The finding that teaching CF to L2 learners is effective and feasible is encouraging especially for foreign language settings in which students do not have much chance to either interact in the target language or benefit from the effects of CF provided by teachers. A L2 is often taught in the students' and teacher's L1, and, even if it is not, the class is usually too large for learners to engage in communication and receive CF from the teacher (see Block, 2003). From another perspective, peer interaction creates an ideal context for transfer-appropriate processing (Morris, Bransford, & Franks, 1977; Segalowitz, 2000) given that interaction between nonnative speakers is becoming a more likely real-life situation that learners will encounter (Ridder, Vangehuchten, & Gómez, 2007).

Giving L2 learners opportunities to engage in negotiation of form can also be recommended because CF was found to be correlated with both accuracy and fluency development. It is important to note, however, that the learners in the present study were already form oriented and needed to proceduralize their rule-based knowledge to access it during spontaneous production. In such cases, it may not be a good pedagogical option to have them become too analytical during meaningful interaction, as also claimed by Lyster and Mori's (2006) counterbalance hypothesis (see also Sato, 2011, for conceptualization of form orientation). Instead, it seems more important to give learners activities in which they engage in meaningful repeated practice. Of importance to further research into CF is the present study's finding that CF provided by peers is a feasible option for helping learners to attend to form in effective ways during peer interaction.

(*Received 30 September 2011*)

REFERENCES

Adams, R. (2007). Do second language learners benefit from interacting with each other? In A. Mackey (Ed.), *Conversational interaction in second language acquisition: A collection of empirical studies* (pp. 29–51). Oxford: Oxford University Press.

Ammar, A. (2008). Prompts and recasts: Differential effects on second language morphosyntax. *Language Teaching Research*, *12*, 183–210.

Ammar, A., & Spada, N. (2006). One size fits all? Recasts, prompts and L2 learning. *Studies in Second Language Acquisition*, *28*, 543–574.

Anderson, J. (1983). *The architecture of cognition*. Cambridge, MA: Harvard University Press.

Anderson, J. (1995). *Learning and memory: An integrated approach*. New York: Wiley.

Anderson, J. (2005). *Cognitive psychology and its implications* (6th ed.). New York: Worth.

Anderson, J., & Schunn, C. (2000). Implications of the ACT-R learning theory: No magic bullets. In R. Glaser (Ed.), *Advances in instructional psychology* (Vol. 5, pp. 1–33). Mahwah, NJ: Erlbaum.

Aston, G. (1986). Trouble-shooting in interaction with learners: The more the merrier? *Applied Linguistics*, *7*, 128–143.

Bajarano, Y., Levine, T., Olshtain, E., & Steiner, J. (1997). The skilled use of interaction strategies: Creating a framework for improved small-group communicative interaction in the language classroom. *System*, *25*, 203–214.

Basturkmen, H., Loewen, S., & Ellis, R. (2004). Teachers' stated beliefs about incidental focus on form and their classroom practices. *Applied Linguistics*, *25*, 243–272.

Bigelow, M., delMas, R., Hansen, K., & Tarone, E. (2006). Literacy and the processing of oral recasts in SLA. *TESOL Quarterly*, *40*, 665–689.

Block, D. (2003). *The social turn in second language acquisition*. Washington, DC: Georgetown University Press.

Bruton, A., & Samuda, V. (1980). Learner and teacher roles in the treatment of oral error in group work. *RELC Journal*, *11*, 49–63.

Buckwalter, P. (2001). Repair sequences in Spanish L2 dyadic discourse: A descriptive study. *Modern Language Journal*, *85*, 380–397.

Bygate, M. (1996). Effects of task repetition: Appraising the developing language of learners. In D. Willis & J. Willis (Eds.), *Challenge: Change in language teaching* (pp. 136–146). London: Heinemann.

Chamot, A., Barnhardt, S., El-Dinary, P., & Robbins, J. (1999). *The learning strategies handbook*. London: Longman.

Cohen, A. (1998). Second language learning and language use strategies: Defining terms. In A. Cohen (Ed.), *Strategies in learning and using a second language* (pp. 3–23). London: Longman.

Cohen, A., Weaver, S., & Li, T.-Y. (1998). The impact of strategies-based instruction on speaking a foreign language. In A. Cohen (Ed.), *Strategies in learning and using a second language* (pp. 107–156). London: Longman.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Mahwah, NJ: Erlbaum.

Crookes, G. (1991). Second language speech production research: A methodologically oriented review. *Studies in Second Language Acquisition*, *13*, 113–132.
Cucchiarini, C., Strik, H., & Boves, L. (2002). Quantitative assessment of second language learners' fluency: Comparisons between read and spontaneous speech. *Journal of the Acoustical Society of America*, *111*, 2862–2873.
de Bot, K. (1992). A bilingual production model: Levelt's "speaking" model adapted. *Applied Linguistics*, *13*, 1–24.
de Bot, K. (1996). The psycholinguistics of the output hypothesis. *Language Learning*, *46*, 529–555.
de Jong, N., & Perfetti, C. (2011). Fluency training in the ESL classroom: An experimental study of fluency development and proceduralization. *Language Learning*, *61*, 533–568.
DeKeyser, R. (1998). Beyond focus on form: Cognitive perspectives on learning and practicing second language grammar. In C. Doughty & J. Williams (Eds.), *Focus on form in classroom second language acquisition* (pp. 42–63). New York: Cambridge University Press.
DeKeyser, R. (2001). Automaticity and automatization. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 125–151). New York: Cambridge University Press.
DeKeyser, R. (2003). Implicit and explicit learning. In C. Doughty & M. Long (Eds.), *Handbook of second language acquisition* (pp. 313–348). Oxford: Blackwell.
DeKeyser, R. (Ed.). (2007). *Practice in a second language: Perspectives from applied linguistics and cognitive psychology*. New York: Cambridge University Press.
DeKeyser, R. (2010). Practice for second language learning: Don't throw out the baby with the bathwater. *International Journal of English Studies*, *10*, 155–165.
Derwing, T., Munro, M., & Thomson, R. (2008). A longitudinal study of ESL learners' fluency and comprehensibility development. *Applied Linguistics*, *29*, 359–380.
Dörnyei, Z. (1995). On the teachability of communication strategies. *TESOL Quarterly*, *29*, 55–85.
Doughty, C. (2001). Cognitive underpinnings of focus on form. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 206–257). New York: Cambridge University Press.
Doughty, C., & Williams, J. (1998). Issues and terminology. In C. Doughty & J. Williams (Eds.), *Focus on form in classroom second language acquisition* (pp. 1–11). New York: Cambridge University Press.
Douglas, D. (2001). Performance consistency in second language acquisition and language testing research: A conceptual gap. *Second Language Research*, *17*, 442–456.
Egi, T. (2007). Interpreting recasts as linguistic evidence: The role of linguistic target, length, and degree of change. *Studies in Second Language Acquisition*, *29*, 511–537.
Ellis, N. (2002). Frequency effects in language processing: A review with implications for theories of implicit and explicit language acquisition. *Studies in Second Language Acquisition*, *24*, 143–188.
Ellis, N. (2005). At the interface: Dynamic interactions of explicit and implicit language knowledge. *Studies in Second Language Acquisition*, *27*, 305–352.
Ellis, R. (1994). A theory of instructed second language acquisition. In N. Ellis (Ed.), *Implicit and explicit learning of languages* (pp. 79–114). San Diego, CA: Academic Press.
Ellis, R. (2005). Measuring implicit and explicit knowledge of a second language: A psychometric study. *Studies in Second Language Acquisition*, *27*, 141–172.
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 A. Mackey (Ed.), *Conversational interaction in second language acquisition: A collection of empirical studies* (pp. 339–360). Oxford: Oxford University Press.
Ellis, R. (2009a). Corrective feedback and teacher development. *L2 Journal*, *1*, 3–18.
Ellis, R. (2009b). Implicit and explicit learning, knowledge and instruction. In R. Ellis, S. Loewen, C. Elder, R. Erlam, J. Philp, & H. Reinders (Eds.), *Implicit and explicit knowledge in second language learning, testing and teaching* (pp. 3–25). Bristol, UK: Multilingual Matters.
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). Reexamining the role of recasts in second language acquisition. *Studies in Second Language Acquisition*, *28*, 575–600.

Foster, P. (1998). A classroom perspective on the negotiation of meaning. *Applied Linguistics*, *14*, 1–23.
Foster, P., & Skehan, P. (1996). The influence of planning and task type on second language performance. *Studies in Second Language Acquisition*, *18*, 293–323.
Fotos, S., & Ellis, R. (1991). Communication about grammar: A task-based approach. *TESOL Quarterly*, *25*, 605–628.
Freed, B. (1995). What makes us think that students who study abroad become fluent? In B. Freed (Ed.), *Second language acquisition in a study abroad context* (pp. 123–148). Amsterdam: Benjamins.
Freed, B., Dewey, D., Segalowitz, N., & Randall, H. (2004). The language contact profile. *Studies in Second Language Acquisition*, *26*, 349–356.
García Mayo, M., & Pica, T. (2000). Interaction among proficient learners: Are input, feedback and output needs addressed in a foreign language context? *Studia Linguistica*, *54*, 272–279.
Gass, S. (1997). *Input, interaction, and the second language learner*. Mahwah, NJ: Erlbaum.
Gass, S. (2003). Input and interaction. In C. Doughty & M. Long (Eds.), *Handbook of second language acquisition* (pp. 224–255). Oxford: Blackwell.
Gass, S., & Varonis, E. (1989). Incorporated repairs in nonnative discourse. In M. Eisenstein (Ed.), *The dynamic interlanguage: Empirical studies in second language variation* (pp. 71–86). New York: Plenum Press.
Green, P., & Hecht, K. (1992). Implicit and explicit grammar: An empirical study. *Applied Linguistics*, *13*, 168–184.
Harmer, J. (1998). *How to teach English: An introduction to the practice of English language teaching*. London: Longman.
Hatch, E., & Lazaraton, A. (1991). *The research manual: Design and statistics for applied linguistics*. Rowley, MA: Newbury House.
Hulstijn, J. (2002). Towards a unified account of the representation, processing and acquisition of second language knowledge. *Second Language Research*, *18*, 193–223.
Hulstijn, J. (2005). Theoretical and empirical issues in the study of implicit and explicit second-language learning. *Studies in Second Language Acquisition*, *27*, 129–140.
Hulstijn, J., & de Graaff, R. (1994). Under what conditions does explicit knowledge of a second language facilitate the acquisition of implicit knowledge? A research proposal. *AILA Review*, *11*, 97–113.
Hulstijn, J., & Hulstijn, W. (1984). Grammatical errors as a function of processing constraints and explicit knowledge. *Language Learning*, *34*, 23–43.
Iñárritu, A. G. (Director), & Arriaga, G. (Writer). (2006). *Babel* [Motion picture]. United States: Central Films.
Indefrey, P., & Levelt, W. (2004). The spatial and temporal signatures of word production components. *Cognition*, *92*, 101–144.
Iwashita, N., Brown, A., McNamara, T., & O'Hagan, S. (2008). Assessed levels of second language speaking proficiency: How distinct? *Applied Linguistics*, *29*, 24–49.
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.
Izumi, S., & Bigelow, M. (2000). Does output promote noticing and second language acquisition? *TESOL Quarterly*, *34*, 239–278.
Johnson, K. (1996). *Language teaching and skill learning*. Oxford: Blackwell.
Kappel, G. (1991). *Design and analysis: A researcher's handbook* (3rd ed.). Englewood Cliffs, NJ: Prentice Hall.
Keselman, H., Huberty, C., Lix, L., Olejnik, S., Cribbie, R., Donahue, B., et al. (1998). Statistical practices of educational researchers: An analysis of their ANOVA, MANOVA, and ANCOVA analyses. *Review of Educational Research*, *68*, 350–386.
Kormos, J. (1999). Monitoring and self-repair in L2. *Language Learning*, *49*, 303–342.
Kormos, J. (2006). *Speech production and second language acquisition*. Mahwah, NJ: Erlbaum.
Kowal, M., & Swain, M. (1997). From semantic to syntactic processing: How can we promote metalinguistic awareness in the French immersion classroom? In R. Johnson & M. Swain (Eds.), *Immersion education: International perspectives* (pp. 284–309). New York: Cambridge University Press.

Lai, Y.-C. (2009). Language learning strategy use and English proficiency of university freshmen in Taiwan. *TESOL Quarterly*, *43*, 255–280.

Lam, W. (2009). Examining the effects of metacognitive strategy instruction on ESL group discussions: A synthesis of approaches. *Language Teaching Research*, *13*, 129–150.

Landis, J., & Koch, G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, *33*, 159–174.

Larson-Hall, J. (2010). *A guide to doing statistics in second language research using SPSS*. London: Routledge.

Lennon, P. (1990). Investigating fluency in EFL: A quantitative approach. *Language Learning*, *40*, 387–417.

Levelt, W. (1983). Monitoring and self-repair in speech. *Cognition*, *14*, 41–104.

Levelt, W. (1989). *Speaking: From intention to articulation*. Cambridge, MA: MIT Press.

Levelt, W. (1999). Language production: A blueprint of the speaker. In C. Brown & P. Hagoort (Eds.), *Neurocognition of language* (pp. 83–122). Oxford: Oxford University Press.

Li, S. (2010). The effectiveness of corrective feedback in SLA: A meta-analysis. *Language Learning*, *60*, 309–365.

Lightbown, P., & 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. (2004). Uptake in incidental focus on form in meaning-focused ESL lessons. *Language Learning*, *54*, 153–187.

Loewen, S. (2005). Incidental focus on form and second language learning. *Studies in Second Language Acquisition*, *27*, 361–386.

Loewen, S. (2011). Focus on form. In E. Hinkel (Ed.), *Handbook of research in second language teaching and learning* (Vol. 2, pp. 576–592). London: Routledge.

Loewen, S., & Philp, J. (2006). Recasts in the adult English L2 classroom: Characteristics, explicitness, and effectiveness. *Modern Language Journal*, *90*, 536–556.

Long, M. (1996). The role of the linguistic environment in second language acquisition. In W. Ritchie & T. Bhatia (Eds.), *Handbook of second language acquisition* (pp. 413–468). San Diego, CA: Academic Press.

Long, M., & Robinson, P. (1998). Focus on form: Theory, research, and practice. In C. Doughty & J. Williams (Eds.), *Focus on form in classroom second language acquisition* (pp. 15–41). New York: Cambridge University Press.

Loschky, L., & Bley-Vroman, R. (1993). Grammar and task-based methodology. In G. Crookes & S. Gass (Eds.), *Tasks and language learning* (pp. 123–167). Bristol, UK: Multilingual Matters.

Lyster, R. (1994). La négociation de la forme: stratégie analytique en classe d'immersion [Negotiation of form: An analytic strategy in immersion classrooms]. *Canadian Modern Language Review*, *50*, 446–465.

Lyster, R. (1998). The ambiguity of recasts and repetition in L2 classroom discourse. *Studies in Second Language Acquisition*, *20*, 51–81.

Lyster, R. (2002). The importance of differentiating negotiation of form and meaning in classroom interaction. In P. Burmeister, T. Piske, & A. Rohde (Eds.), *An integrated view of language development: Papers in honor of Henning Wode* (pp. 381–397). Trier, Germany: Wissenschaftlicher Verlag Trier.

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

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. (1997). Corrective feedback and learner uptake: Negotiation of form in communicative classrooms. *Studies in Second Language Acquisition*, *19*, 37–66.

Lyster, R., & Saito, K. (2010). Oral feedback in classroom SLA: A meta-analysis. *Studies in Second Language Acquisition*, *32*, 265–302.

Mackey, A. (2006). Feedback, noticing and instructed second language learning. *Applied Linguistics*, *27*, 405–430.

Mackey, A. (2007). Introduction: The role of conversational interaction in second language acquisition. In A. Mackey (Ed.), *Conversational interaction in second language acquisition: A collection of empirical studies* (pp. 1–26). Oxford: Oxford University Press.
Mackey, A., & Goo, J. (2007). Interaction research in SLA: A meta-analysis and research synthesis. In A. Mackey (Ed.), *Conversational interaction in second language acquisition: A collection of empirical studies* (pp. 407–452). Oxford: Oxford University Press.
Mackey, A., Oliver, R., & Leeman, J. (2003). Interactional input and the incorporation of feedback: An exploration of NS-NNS and NNS-NNS adult and child dyads. *Language Learning*, *53*, 35–66.
Mackey, A., & Philp, J. (1998). Conversational interaction and second language development: Recasts, responses, and red herrings? *Modern Language Journal*, *82*, 338–356.
MacWhinney, B. (2009). *The CHILDES project: Tools for analyzing talk: Vol. 2: Transcription format and programs*. Department of Psychology, Carnegie Mellon University.
McDonough, K. (2005). Identifying the impact of negative feedback and learners' responses on ESL question development. *Studies in Second Language Acquisition*, *27*, 79–103.
McDonough, K., & Mackey, A. (2000). Communicative tasks, conversational interaction, and linguistic form: An empirical study of Thai. *Foreign Language Annals*, *33*, 82–91.
McDonough, K., & Mackey, A. (2006). Responses to recasts: Repetitions, primed production, and linguistic development. *Language Learning*, *56*, 693–720.
Miyagi, K., Sato, M., & Crump, A. (2009). To challenge the unchallenged: Potential of non-"standard" Englishes for Japanese EFL learners. *JALT Journal*, *31*, 261–273.
Morris, C., Bransford, J., & Franks, J. (1977). Levels of processing versus transfer appropriate processing. *Journal of Verbal Learning & Verbal Bahavior*, *16*, 519–533.
Morrison, D., & Low, G. (1983). Monitoring and the second language learner. In J. Richards & R. Schmidt (Eds.), *Language and communication* (pp. 228–250). London: Longman.
Muranoi, H. (2007). Output practice in the L2 classroom. In R. DeKeyser (Ed.), *Practice in a second language: Perspectives from applied linguistics and cognitive psychology* (pp. 51–84). New York: Cambridge University Press.
Nabei, T., & Swain, M. (2002). Learner awareness of recasts in classroom interaction: A case study of an adult EFL student's second language learning. *Language Awareness*, *11*, 43–63.
Nakatani, Y. (2005). The effects of awareness-raising instruction on oral communication strategy use. *Modern Language Journal*, *89*, 76–91.
Nation, P. (1989). Improving speaking fluency. *System*, *17*, 377–384.
Naughton, D. (2006). Cooperative strategy training and oral interaction: Enhancing small group communication in the language classroom. *Modern Language Journal*, *90*, 169–184.
Newell, A., & Rosenbloom, P. (1981). Mechanisms of skill acquisition and the law of practice. In J. Anderson (Ed.), *Cognitive skills and their acquisition* (pp. 1–55). Mahwah, NJ: Erlbaum.
Nicholas, H., Lightbown, P., & Spada, N. (2001). Recasts as feedback to language learners. *Language Learning*, *51*, 719–758.
Nobuyoshi, J., & Ellis, R. (1993). Focused communication tasks and second language acquisition. *ELT Journal*, *47*, 203–210.
O'Donnell, K. (2005). Japanese secondary English teachers: Negotiation of educational roles in the face of curricular reform. *Language, Culture and Curriculum*, *18*, 300–315.
Olejnik, S., Li, J., Supattathum, S., & Huberty, C. (1997). Multiple testing and statistical power with modified Bonferroni procedures. *Journal of Educational and Behavioral Statistics*, *22*, 389–406.
Ortega, L. (2009). *Understanding second language acquisition*. London: Hodder Education.
Oxford, R. (1990). *Language learning strategies: What every teacher should know*. Rowley, MA: Newbury House.
Panova, I., & Lyster, R. (2002). Patterns of corrective feedback and uptake in an adult ESL classroom. *TESOL Quarterly*, *36*, 573–595.
Paradis, M. (2004). *A neurolinguistic theory of bilingualism*. Amsterdam: Benjamins.
Paradis, M. (2009). *Declarative and procedural determinants of second languages*. Amsterdam: Benjamins.
Pica, T. (1994). Research on negotiation: What does it reveal about second-language learning conditions, processes, and outcomes? *Language Learning*, *44*, 493–527.

Pica, T., Kanagy, R., & Falodun, J. (1993). Choosing and using tasks for second language instruction and research. In G. Crookes & S. Gass (Eds.), *Tasks and language learning: Integrating theory and practice* (pp. 9–34). Bristol, UK: Multilingual Matters.

Pica, T., Lincoln-Porter, F., Paninos, D., & Linnell, J. (1996). Language learners' interaction: How does it address the input, output, and feedback needs of L2 learners? *TESOL Quarterly*, *30*, 59–84.

Porter, P. (1986). How learners talk to each other: Input and interaction in task-centered discussions. In R. Day (Ed.), *Talking to learn: Conversation in second language acquisition* (pp. 200–222). Rowley, MA: Newbury House.

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: Perspective from applied linguistics and cognitive psychology* (pp. 141–160). New York: Cambridge University Press.

Ridder, I., Vangehuchten, L., & Gómez, M. (2007). Enhancing automaticity through task-based language learning. *Applied Linguistics*, *28*, 309–315.

Riggenbach, H. (1991). Toward an understanding of fluency: A microanalysis of nonnative speaker conversations. *Discourse Processes*, *14*, 423–441.

Robinson, P. (1995). Attention, memory, and the "noticing" hypothesis. *Language Learning*, *45*, 283–331.

Russell, J., & Spada, N. (2006). The effectiveness of corrective feedback for the acquisition of L2 grammar. In J. Norris & L. Ortega (Eds.), *Synthesizing research on language learning and teaching* (pp. 133–162). Amsterdam: Benjamins.

Sato, M. (2007). Social relationships in conversational interaction: A comparison between learner-learner and learner-NS dyads. *JALT Journal*, *29*, 183–208.

Sato, M. (2011). Constitution of form-orientation: Contributions of context and explicit knowledge to learning from recasts. *Canadian Journal of Applied Linguistics*, *14*, 1–28.

Sato, M., & Lyster, R. (2007). Modified output of Japanese EFL learners: Variable effects of interlocutor vs. feedback types. In A. Mackey (Ed.), *Conversational interaction in second language acquisition: A collection of empirical studies* (pp. 123–142). Oxford: Oxford University Press.

Sato, M., & Perry, B. (2011). How to transform second language learners into interactive corrective feedback providers. In M. Skier & M. Walsh (Eds.), *Pan-SIG 2010 conference proceedings* (pp. 188–205). Kyoto, Japan: JALT.

Schmidt, R. (1990). The role of consciousness in second language learning. *Applied Linguistics*, *11*, 129–158.

Schmidt, R. (1994). Deconstructing consciousness in search of useful definitions for applied linguistics. *AILA Review*, *11*, 11–26.

Schmidt, R. (2001). Attention. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 3–32). New York: Cambridge University Press.

Seedhouse, P. (1997). Combining form and meaning. *ELT Journal*, *51*, 337–344.

Segalowitz, N. (2000). Automaticity and attentional skill in fluent performance. In H. Riggenbach (Ed.), *Perspectives on fluency* (pp. 200–219). Ann Arbor: University of Michigan Press.

Segalowitz, N. (2003). Automaticity and second languages. In C. Doughty & M. Long (Eds.), *Handbook of second language acquisition* (pp. 382–408). Oxford: Blackwell.

Segalowitz, N. (2010). *Cognitive bases of second language fluency*. London: Routledge.

Segalowitz, N., & Freed, B. (2004). Context, contact, and cognition in oral fluency acquisition: Learning Spanish in at home and study abroad contexts. *Studies in Second Language Acquisition*, *26*, 173–199.

Segalowitz, N., & Lightbown, P. (1999). Psycholinguistic approaches to SLA. *Annual Review of Applied Linguistics*, *19*, 43–63.

Segalowitz, N., & Segalowitz, S. (1993). Skilled performance, practice, and the differentiation of speed-up from automatization effects: Evidence from second language word recognition. *Applied Psycholinguistics*, *14*, 369–385.

Sheen, Y. (2006). Exploring the relationship between characteristics of recasts and learner uptake. *Language Teaching Research*, *8*, 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.

Sheen, Y., & Ellis, R. (2011). Corrective feedback in language teaching. In E. Hinkel (Ed.), *Handbook of research in second language teaching and learning* (Vol. 2, pp. 593–610). London: Routledge.

Shehadeh, A. (2001). Self- and other-initiated modified output during task-based interaction. *TESOL Quarterly*, *35*, 433–457.

Soler, E. (2002). Relationship between teacher-led versus learners' interaction and the development of pragmatics in the EFL classroom. *International Journal of Educational Research*, *37*, 359–377.

Spada, N. (1997). Form-focused instruction and second language acquisition: A review of classroom and laboratory research. *Language Teaching*, *30*, 73–87.

Spada, N. (2011). Beyond form-focused instruction: Reflections on past, present and future research. *Language Teaching*, *44*, 225–236.

Stevens, J. (2002). *Applied multivariate statistics for the social sciences* (4th ed.). Mahwah, NJ: Erlbaum.

Swain, M. (1985). Communicative competence: Some roles of comprehensible input and comprehensible output in its development. In S. Gass & 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 G. Cook & B. Seidlhofer (Eds.), *Principle and practice in applied linguistics: Studies in honour of H. G. Widdowson* (pp. 125–144). Oxford: Oxford University Press.

Swain, M. (2005). The output hypothesis: Theory and research. In E. Hinkel (Ed.), *Handbook of research in second language teaching and learning* (pp. 471–483). Mahwah, NJ: Erlbaum.

Toth, P. (2008). Teacher- and learner-led discourse in task-based grammar instruction: Providing procedural assistance for L2 morphosyntactic development. *Language Learning*, *58*, 237–283.

Towell, R., & Dewaele, J. (2005). The role of psycholinguistic factors in the development of fluency amongst advanced learners of French. In J. Dewaele (Ed.), *Focus on French as a foreign language: Multidisciplinary approaches* (pp. 210–239). Bristol, UK: Multilingual Matters.

Towell, R., Hawkins, R., & Bazergui, N. (1996). The development of fluency in advanced learners of French. *Applied Linguistics*, *17*, 84–119.

Ullman, M. (2001). The neural basis of lexicon and grammar in first and second language: The declarative/procedural model. *Bilingualism: Language and Cognition*, *4*, 105–122.

Ullman, M. (2004). Contributions of memory circuits to language: The declarative/procedural model. *Cognition*, *92*, 231–270.

Ullman, M. (2005). A cognitive neuroscience perspective on second language acquisition: The declarative/procedural model. In C. Sanz (Ed.), *Mind and context in adult second language acquisition* (pp. 141–178). Washington, DC: Georgetown University Press.

VanPatten, B. (1990). Attending to content and form in the input: An experiment in consciousness. *Studies in Second Language Acquisition*, *12*, 287–301.

VanPatten, B. (2000). Processing instruction as form-meaning connections: Issues in theory and research. In A. Valdman & J. Lee (Eds.), *Form and meaning: Multiple perspectives* (pp. 43–68). Boston: Heinle & Heinle.

Varonis, E., & Gass, S. (1985). Non-native/non-native conversations: A model for negotiation of meaning. *Applied Linguistics*, *6*, 71–90.

Williams, J. (1999). Learner-generated attention to form. *Language Learning*, *49*, 583–625.

Yang, Y., & Lyster, R. (2010). Effects of form-focused practice and feedback on Chinese EFL learners' acquisition of regular and irregular past tense forms. *Studies in Second Language Acquisition*, *32*, 235–263.

Yoshida, R. (2008). Learners' perception of corrective feedback in pair work. *Foreign Language Annals*, *41*, 525–541.

APPENDIX A

A SAMPLE SCENARIO AND ERROR LIST (POLICE REPORT)

Grammatical target: Past tense

Scenario: You witnessed a bank robbery. While you were at the bank, a robber suddenly came in with a knife and took money from the bank. A police officer is now interviewing you. You are asked to describe what you saw.

Error	Correct
I **am** withdrawing some money.	I was withdrawing some money.
The robber **is** wearing a mask.	The robber was wearing a mask.
He **has** a knife.	He had a knife.
Everyone **is** scared.	Everyone was scared.
He **says** "give me the money."	He said "give me money."
I **am** withdrawing some money.	I was withdrawing some money.
The robber **is** yelling.	The robber was yelling.
He **puts** the knife on the counter.	He put the knife on the counter.
The bank teller **is** killed.	The bank teller was killed.
He **runs** away from the scene.	He ran away from the scene.
I **am** withdrawing some money.	I was withdrawing some money.
A boy **is** screaming.	A boy was screaming.
He **sees** me.	He saw me.
An old lady **is** kicked.	An old lady was kicked.
He **shoot** the banker.	He shot the banker.

APPENDIX B

SAMPLE INTERACTION DATA

Topic: Mr. Okada (head coach of the Japanese national soccer team during the World Cup 2010)

Excerpt 1: Wataru (W) vs. Yusuke (Y)

Y:	*But, actually, he won in World Cup. So, he can have confidence. He must have confidence. If he have confidence,*	
W:	***If he have? He have?***	**Prompt**
Y:	*Oh!* ***If he had*** *confidence, we can believe him.*	**MO**
W:	*Your opinion is good.*	

Excerpt 2: Shun (S) vs. Kei (K)

K:	*But, his team enter the final league.*	
S:	***Pardon me?***	**Prompt**
K:	*Oh,* ***entered.*** *Thank you. And he is very great. I want to give him trophy.*	**MO**
S:	*OK. My present is Messi.*	

Excerpt 3: Sawako (S) vs. Mayu (M)

M:	*When Okada Coach lose the game*	
S:	***Lost, I think.***	**Recast**
M:	*Ah.* ***When he lost the game****, all the people in Japan accused him. But, when the World Cup ended, we like him.*	**MO**
S:	*We all like him.*	

Excerpt 4: Sachiko (S) vs. Mai (M)

M:	*In TV, I was heard that Mr. Okada have*	
S:	***Mr. Okada has***	**Recast**
M:	***Has, Mr. Okada has daughter.***	**MO**