Studies in Second Language Acquisition, 2016, **38**, 771–799.
doi:10.1017/S0272263115000467

EFFICACY OF RECASTS AND GESTURES ON THE ACQUISITION OF LOCATIVE PREPOSITIONS

Kimi Nakatsukasa
Texas Tech University

This study investigates whether gestures can be used during recasts to enhance the saliency of a target structure (locative prepositions) and to lead to better production of the target structure. Forty-eight low-intermediate English as a second language (ESL) students partook in communicative activities during which they received either no feedback (control), verbal recasts only (R), or recasts plus gesture (RG), and a subset of participants completed a stimulated recall session. Then the pretest, immediate, and delayed posttest scores of grammar and oral production tests were used to analyze the linguistic development. The results showed that no one commented on recasts or locative prepositions during the stimulated recall session and that there were no significant changes in grammar test scores in all conditions; however, the R and RG conditions performed significantly better in the production test than the control in the immediate posttest. Furthermore, the RG condition maintained the development in the delayed posttest, whereas the R condition did not.

The present study investigated whether gestures, when used along with verbal recasts—traditionally a form of implicit feedback—can promote noticing and facilitate L2 grammar acquisition. One central topic of a growing number of corrective feedback (CF) studies is what types of CF

This manuscript used part of the data from my dissertation project, which was funded by a Language Learning Dissertation Award. I also extend my gratitude to Drs. Shawn Loewen, Susan Gass, Paula Winke, and Debra Hardison for their valuable comments in the earlier version of the manuscript and to Drs. Luke Plonsky and Dominik Wolff for their feedback as well. In addition, I am truly grateful to the reviewers of *SSLA* for their thorough, constructive feedback. Lastly, I greatly appreciate the ESL teachers and students who collaborated with me on this study.

lead to L2 learning. Recasts, typically a rather implicit feedback, seem to be used by language teachers frequently (Ellis & Sheen, 2006; Han, 2002; Leeman, 2003; Loewen & Philp, 2006); however, their effectiveness has been questioned due to the lack of saliency of recasts.

Gestures have been a topic of interest in a number of fields, including developmental psychology and education. Researchers have reported that seeing gestures promotes cognitive development (e.g., Cook, Yip, & Goldin-Meadow, 2010; Goldin-Meadow, Cook, & Mitchell, 2009; Goldin-Meadow & Sandhofer, 1999; Goldin-Meadow & Singer, 2003). Studies on gestures have gained attention in the field of SLA recently. For example, teacher's gestures seem to help student comprehension (e.g., Allen, 2000; Lazaraton, 2004; Sime, 2006) and have pedagogical characteristics (e.g., Hudson, 2011; Zhao, 2007). To date, however, no interaction studies have addressed the extent to which exposure to pedagogical gestures alone promotes L2 learning. This intervention study examined whether pedagogical gestures can enhance the saliency of linguistic targets of recasts and lead to better L2 learning with regards to the acquisition of locative prepositions.

RECASTS AND GESTURAL STUDIES

This section first reviews the previous studies on CF, specifically about recasts, and then discusses noticing. The final part focuses on gesture studies with regards to gestures' functions in L2 learning and the use of gestures in language classrooms.

Recasts

Recasts, as a type of implicit feedback, are often the most commonly used type of CF in language classrooms (e.g., Doughty, 1994; Havranek, 1999; Lyster & Ranta, 1997; Mackey, Gass, & McDonough, 2000; also see Oliver, 1995). Investigators have found that recasts were favored because they did not interrupt the flow of communication and because they saved time spent on correction (Ellis & Sheen, 2006; Han, 2002; Leeman, 2003; Loewen & Philp, 2006).

Theoretically, recasts are thought to facilitate L2 development because they include accurate models, which enable learners to notice the gap between their interlanguage (IL) and targetlike production, resulting in IL development (Long, 2007). Specifically, Long (1996, 2007) argued that recasts promote form-function mapping because recasts are provided during the interaction when the learners' production is questioned. Indeed, Li (2010) found that the long-term benefits of implicit feedback,

such as recasts, compared favorably to those of explicit feedback. He meta-analyzed 33 studies on CF and found that although explicit feedback was more effective than implicit feedback in immediate and short-delayed posttests, the benefit remained substantially longer for implicit feedback.

However, the occurrence of noticeability in recasts is still debated. The majority of empirical studies (e.g., Mackey & Philp, 1998; Nassaji, 2009; Yang & Lyster, 2010), together with meta-analytic studies (e.g., Li, 2010; Mackey & Goo, 2007), have shown that recasts are developmentally effective. However, there have also been several studies that reported the opposite findings (e.g., Ammar & Spada, 2006; Lyster, 1998, 2001, 2004; Sheen, 2004, 2007; Slimani, 1992), and still others were inconclusive (e.g., Loewen & Nabei, 2007; Long, Inagaki, & Ortega, 1998).

Researchers accounted for such discrepancies by analyzing various interlearner and contextual factors affecting the noticeability of recasts and of the target structures. They identified four aspects that impact the efficacy of recasts. First, as for the individual differences, it was reported that recasts are beneficial when learners have a higher level of proficiency and are ready to learn a target structure (e.g., Ammar & Spada, 2006; Li, 2013; Mackey & Philp, 1998; Nicholas, Lightbown, & Spada, 2001; Philp & Mackey, 2010) and that cognitive factors such as working memory capacity also impact the effectiveness of recasts (e.g. Goo, 2012; Yilmaz, 2013). Second, in classroom contexts, the studies reported that recasts are more noticeable when a classroom is more form oriented, (e.g., Lyster & Mori, 2006; Sheen, 2004). Third, with regards to linguistic targets, lexical and phonological targets are more likely to be noticed (e.g., Lyster, Saito, & Sato, 2013; Mackey et al., 2000), but the correct noticeability of grammatical targets varies (e.g., Lyster et al., 2013). Lastly, as for the characteristics of oral recasts, it was reported that various features of recasts affect the saliency of recasts and thus the noticeability of recasts and linguistic targets (e.g., Kamiya, 2015; Loewen & Philp, 2006; Sheen, 2006). It has been suggested that the saliency and noticeability of recasts vary depending on features such as intonation, stress, and number of feedback moves (Loewen & Philp, 2006). If so, it is logical to assume that the nonlinguistic features may also influence recasts' saliency and subsequent L2 development.

Noticing

The crucial factor that determines the effectiveness of CF is the learners' cognitive processing that happens during the CF—notably, attention and awareness (e.g., Gass, 1997; Gass & Varonis, 1994; Hama & Leow, 2010;

Mackey et al., 2000; Philp, 2003; Robinson, 1995, 2001, 2003). The general consensus is that CF helps L2 learning because it provides opportunities for the learners (a) to bring attention to nontargetlike production, (b) to recognize the difference between their IL and targetlike production (to "notice the gap," in the words of Schmidt & Frota, 1986), and (c) to produce modified output following CF (Swain, 2005). In other words, the interaction draws learners' attention to the problems of their IL, which results in L2 development. Thus it is evident that recasts can be manipulated to improve efficacy; this effect results from learners' different levels of attention to and awareness of target structures. Yet, to date, studies that analyzed the saliency of feedback have relied on verbal cues, and the significance of nonverbal cues has yet to be determined.

The original noticing hypothesis (Schmidt, 1990, 1993, 1994, 1995) claims that noticing requires awareness and that SLA cannot occur without noticing. Schmidt (1995) further argued that only linguistic items that learners notice could result in intake. This notion—that noticing is required for L2 learning—caused researchers to question whether learning truly requires awareness (e.g., Carroll, 1999; Gass, 1997; Tomlin & Villa, 1994). Truscott (1999) specifically challenged the impact of noticing on communicative competence. He argued that noticing only helps development of metalinguistic knowledge. Because of these criticisms, in 2001, Schmidt (2001) proposed a weaker version of the noticing hypothesis. In this version, he still maintained that noticing requires a conscious apprehension and awareness of input; however, he accepted the notion of possible subliminal perception that does not result in subliminal learning.

In the past ten years, researchers have examined whether L2 learning could occur with or without attention and awareness, and the results have been inconclusive (e.g., see Gass, Svetics, & Lemelin, 2003; Hama & Leow, 2010; Williams, 2005). Today, it is generally conceded that noticing and awareness are helpful for L2 learning (e.g., Godfroid, Boers, & Housen, 2013; Leow, 2000; Long, 1996; Mackey, 2006; Mackey, Gass, & McDonough, 2000; Swain & Lapkin, 2002; Williams, 2005), even though unattended learning is not entirely impossible (e.g., Gass et al., 2003; Izumi, 2002). The question whether noticing at the level of consciousness is absolutely necessary for L2 learning is interesting; however, it is not the goal of the present study. Thus, in this study, I side with the general notion that noticing is overall helpful for L2 learning.

Gesture and SLA

This section summarizes three major areas of gestural studies in the field of SLA. Gestures—hand movements that are directly tied to speech

(McNeill, 1992)—play a crucial role in processes of interaction and communication. The majority of gestural studies have taken place in the field of education and psychology; however, recently, gestures have received attention in SLA research as well. Researchers have reported that when L2 speakers use gestures, their gestures often serve as a form of scaffolding to compensate for their speech, and such gestures facilitate L2 production (cited in Gullberg, 2010). In addition, researchers analyzed the gestures used in language classrooms and showed that students used the specific gestures of their target culture introduced by teachers (e.g., Peltier & McCafferty, 2010). Some studies also showed that teachers' gestures help L2 speakers with their comprehension (Allen, 1999; Sueyoshi & Hardison, 2005) and with memorization, including L2 vocabulary and expressions (Allen, 1995; Tellier, 2008).

Exposure to Gestures and L2 Learning. Prior to the recent gestural studies on L2 learning, studies in the field of psychology investigated the impact of gesturing and of exposure to gestures on learning (see Goldin-Meadow, 2004, for a summary). Some studies argued that the gestures used by teachers or adults are essential for learning because gestures (a) facilitate comprehension (e.g., Church, 1999; Golden-Meadow, Kim, & Singer, 1999; Perry, Birch, & Singleton, 1995; Valenzeno, Alibali, & Klatzky, 2003) and (b) help with the retention of information especially with regards to the viewers' short-term memory (R. L. Cohen & Otterbein, 1992; Feyereisen, 1998). Combining these findings, studies suggest that exposure to gestures plays an integral role in learning in general.

In the domain of SLA, some studies used intervention designs to assess whether gestures promote L2 comprehension (e.g., Church, Ayman-Nolley, & Mahootian, 2004; Sueyoshi & Hardison, 2005) and the acquisition of L2 vocabulary and expressions (Allen, 1995; Tellier, 2008). Overall, they found that the presence of gestures had a positive effect, as opposed to when gestures are absent. Sueyoshi and Hardison (2005) investigated the role of gestures and facial cues in L2 listening comprehension with 42 low-intermediate and advanced L2 English learners. They watched a lecture video in English in one of the following three conditions: (a) an audiovisual lecture with facial expressions and gestures, (b) an audiovisual lecture with facial expressions but without gestures, and (c) an audio lecture without facial expressions or gestures. According to the scores of the postlecture listening comprehension test, higher proficiency speakers who viewed the lecture with the facial cues answered the questions more accurately than those in the other two conditions. Lower proficiency speakers who watched the lecture with facial expressions and gestures answered the greatest number of questions correctly among the three conditions. This study suggests that gestures are more beneficial for lower proficiency speakers than for higher proficiency speakers.

These findings are not limited to L2 comprehension alone; researchers have found that teacher gestures help student vocabulary and expression learning. Allen (1995) and Tellier (2008) investigated how seeing and mimicking the gestures of teachers and researchers helped students better remember expressions and vocabulary. Tellier conducted a study with 20 monolingual French children to measure the effects of gestures on their memorization of eight English vocabulary items after watching a video recording of the vocabulary lesson with (a) equivalent illustrations or (b) equivalent gestures. The children in the gesture group also repeated the gestures used in the video while watching. The results indicated a positive impact for seeing gestures and gesturing; the children in the gesture group performed significantly better on vocabulary production and retention tests than the other children. Tellier concluded that this was because the exposure to gestures, in conjunction with children's gesturing, facilitated their memories. Allen also used a similar research design with regard to the acquisition of L2 French expressions and obtained similar results. The results from these two studies need to be interpreted with caution, especially when considering the efficacy of "seeing" the teachers' gestures, because the participants had opportunities to gesture in both studies. Although isolating the efficacy of seeing gestures is not possible, these studies indicate that being exposed to gestures facilitates vocabulary and expression acquisition.

Findings from recent neuroscience studies may account for the aforementioned beneficial role of gestures on L2 learning. The studies reported two major findings with regard to the processing of gestures: seeing gestures (a) triggers semantic processing (Holle & Gunter, 2007; Kelly, Kravitz, & Hopkins, 2004; Kelly, Ward, Creigh, & Bartolotti, 2007; Wu & Coulson, 2005) and (b) allows the brain to decrease the need for semantic aspects of language comprehension and to use the additional resources for information (Skipper, Goldin-Meadow, Nusbaum, & Small, 2007). If this is the case, L2 learners are able to process the information of the CF with less effort with regards to semantic processing when meaningful gestures are present.

Teachers' Gestures in Language Classrooms. The aforementioned studies showed that exposure to gestures is helpful for L2 learning, but understanding how gestures are actually used in the classroom is essential. A number of educational studies examined how teachers use gestures in math and science classrooms and reported two major findings: (a) teachers frequently use gestures in the classroom (Alibali, Flevares, & Goldin-Meadow, 1997; Crowder, 1996; Perry et al., 1995; Roth, 2001; Roth & Lawless, 2002a, 2002b) and (b) teachers use gestures to increase the attention of students, resulting in better learning (e.g., Flevares & Perry, 2001; Perry et al., 1995). In addition, researchers took further steps to

identify the efficacy of teacher gestures on learning in a classroom context (e.g., Perry et al., 1995) and found that the students who saw pointing gestures outperformed those who did not in problem-solving tasks; they were also more likely to produce correct responses following a teacher's verbal and gestural explanation as opposed to following a verbal-only explanation, possibly due to the enhanced content from gestures.

As for the studies in language classrooms, some observational studies described how teachers used gestures to provide pedagogical functions—to teach abstract concepts such as metaphors, verb tenses, and spatial relationships (e.g., Allen, 2000; Hudson, 2011; Lazaraton, 2004; Tellier, 2006; Wang, 2009; Zhao, 2007). For example, Hudson (2011) investigated how a teacher used gestures in an ESL classroom at a U.S. university and showed that the teacher incorporated gestures that involved body movement when teaching grammar, pronunciation, and vocabulary. Similarly, Tellier (2006) also reported that teachers use gestures to manage their classrooms, evaluate the responses of students, and explain grammar, pronunciation, and new vocabulary.

In sum, previous gestural studies in language classrooms showed that teachers gestured with a pedagogical purpose; however, the question regarding the impact of such gestures on actual learning in a classroom still remains unanswered. In the next section, I review studies that have quantitatively measured the effect of gestures on L2 learning.

Teacher Gestures, Corrective Feedback, and Learning. A few recent studies examined the extent to which teacher gestures in language classrooms impact language learning (Allen, 2000; Hudson, 2011; Lazaraton, 2004; Tellier, 2006; Wang, 2009; Zhao, 2007).

Davies (2006) conducted one of the first studies that incorporated paralinguistic features in interaction studies. He examined the ratio of various types of CF, including those that were provided purely paralinguistically (Davies did not explain what he meant by "purely paralinguistic," but presumably the paralinguistic cues included gestures, eye contact, facial expressions, etc.), and the ratio of learner uptake following CF. He found that when the feedback involved paralinguistic features, there was a higher ratio of uptake than when it did not involve any. This study provides information about how gestures are potentially used. However, to fully capture the efficacy of gestures during CF, an intervention study must be constructed with limited linguistic targets and task characteristics while still being administered in a real language classroom to maintain ecological validity. Lyster and Ranta (2013) highlighted the importance of carrying out classroom research with regards to CF—even though, as cautioned by Goo and Mackey (2013), classroom research may bring in various uncontrollable factors, such as uneven levels of proficiency and different social contexts. Still, conducting an

intervention study in a classroom setting may more realistically depict the effectiveness of different types of feedback than research carried out in a laboratory setting.

RESEARCH QUESTIONS

By bridging the literature on recasts and gestures, this study specifically investigated the effectiveness of recasts when used with pedagogical gestures for development in the use of locative prepositions. Specifically, the following questions were asked:

1. Do learners notice the linguistic target of recasts when verbal recasts involve gestures?
2. Are recasts more effective in the acquisition of locative prepositions when they are presented only verbally or with gestures?
3. How durable over time is the effectiveness of recasts on the acquisition of locative prepositions when they are presented only verbally versus with gestures?

METHODOLOGY

Participants

All the participants were enrolled in low-intermediate ESL classes at a large state university in the United States. A total of 48 participants from 10 classrooms agreed to participate in the study, and on completion of all the sessions they were given extra credit as determined by the instructors of each class. Four classrooms were assigned to the recast + gesture (RG) condition, another four classrooms were assigned to the verbal recast only (R) condition, and two classrooms were assigned to control groups. All the participants were new arrivals to the United States; most of the students (n = 40) were in their first semester, and a few (n = 8) were in their second.

A total of 27 participants for the RG condition, 13 for the R condition, and 8 for the control condition were included in the analysis.[1] Out of 48 participants, 28 were male, and 20 were female. The participants' L1s were Arabic (n = 15), Chinese (n = 29), Korean (n = 2), and Thai (n = 2). Their average age was 20.4 years (SD = 1.88).[2]

Target Structure

The selected target structure was English locative prepositions (*above*, *under*, *in*, on, and *next to*), which can be problematic for learners.

Prepositions are introduced early in the ESL curriculum; however, they remain one of the most difficult linguistic features to be acquired (Kao, 2001). This difficulty has led researchers to investigate patterns of preposition acquisition by L2 speakers (e.g., Jarvis & Odlin, 2000) and to develop instructional materials for language teachers (e.g., Celce-Murcia & Larsen-Freeman, 1998) and English learners (e.g., Lindstromberg, 2010). Jarvis and Odlin (2000) examined how L1 Swedish and Finnish speakers use English locative prepositions and argued that learners' omission and overgeneralization of prepositions are due to their L1 transfer.

In addition to the contrast between the L1 and L2 conceptualizations, learners also need to be aware of multiple dimensions of prepositions, such as the shape of the landmark (e.g., the landmark is seen as a container for *in*), the direction of the axis (e.g., vertical for *above*), and transitivity (e.g., "crashed *into*" is transitive, but "go *in*" can be either transitive or intransitive). Such complex dimensions of prepositions also seem to result in difficulty acquiring L2 prepositions.

Among the many meanings of each preposition, learners often find the locative ones to be particularly "depictable". It must be noted, however, that how "depictable" they are depends on each preposition. In addition, previous classroom studies showed that language teachers actually use gestures while teaching prepositional phrases (Hudson, 2011). Therefore, it seems logical to incorporate pedagogical gestures that depict geometric features when teaching locative prepositions.

Materials

This intervention study included the following five stages: (a) pretests, (b) treatment sessions, (c) immediate posttests, (d) verbal report, and (e) delayed posttests. The detailed procedure is presented in the next section.

Assessment Instruments. There were two types of assessments for the pretest, posttest, and delayed posttest: an oral production test and a grammar test targeting prepositions. The oral production test was designed to assess learners' spontaneous use of target structures in spontaneous speech, and the untimed grammar test was designed to assess their explicit knowledge.

Oral production test. Three versions of PowerPoint slides were developed to be randomly used in the pretest, posttest, and delayed posttest. Each version was composed of two practice questions, eight questions, and four distracters, to be used for the oral production tests.[3] A set of four slides was created for each question, which appeared in the format

shown in Figure 1. The prompt appearing on Slide 2 indicates which item participants needed to focus on for the study. In Slide 3, the second prompt appears, and the participants were asked to verbally respond. Each question elicited a preposition such as *above*, *under*, *in*, *on*, and *next to*. The responses were audio recorded individually using a voice recorder. In order to keep the participants from focusing only on the locations of the items, distracter questions—such as "How many strawberries do you see?"—were also asked. The scores were obtained by calculating the ratio of the correct use of prepositions in the obligatory context because the total number of prepositions was not controlled. The order of the versions was randomized for each participant. To verify whether or not the levels of difficulty of the three test versions were similar, the scores obtained from the three tests during pretests were calculated using a one-way ANOVA. The analysis showed that the tests' levels of difficulty were not different from one another (p = .984). This indicates that the level of difficulty was similar in all three versions.

Grammar test. The second section of the test was the preposition grammar test. This test was composed of 20 questions that included

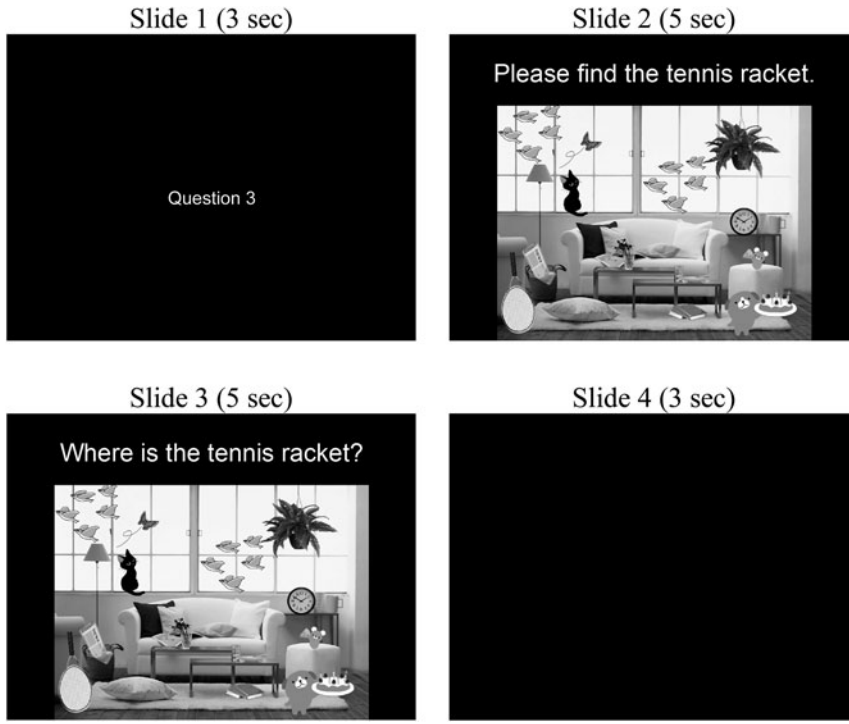

Figure 1. The timing and format of the questions from the preposition oral production test.

5 distracters. The participants looked at a picture and read a sentence to determine whether the sentence was an accurate description of the picture. If the sentence was correct, they were asked to circle TRUE, and if the sentence was incorrect, they were asked to circle FALSE and write a correct sentence. Out of the 15 target items, 10 (2 items per preposition) were accompanied by an incorrect (1 item per preposition) description. Participants were allowed to spend as much time as needed to complete this test.

Communicative Tasks. A total of two communicative tasks following a format of information gap activity were designed to elicit prepositions in class. The learners were required to use the prepositions in a communicative context to complete the tasks successfully. The tasks were constructed based on Ellis's (2003) definition of focus tasks—that is, a task that includes a gap and a clearly defined outcome. Learners needed to focus on meaning and on using their own linguistic resources.

First information gap activity. In the first task, the participants were divided into one of the two groups: a detective team and a burglar team. About six to eight students were in each team. The following story was provided as context for the activity: the burglars had stolen a diamond, and they needed to hide it somewhere in the house before the detectives came and found it. The detectives needed to find the diamond using the fewest possible questions. Each group was given 4 min to prepare for the task. A laminated, poster-sized print of a picture of a house was attached to the blackboard, and a picture of the house printed on letter-sized paper was given to each group. The names of the furniture were labeled to facilitate production. Then the detectives asked yes/no questions, such as "Is the diamond under the bed?" until they found it. To make sure that their prepositions were used correctly, the participants who were playing the role of detectives were told to point to locations using a laser pointer. Then the teams switched roles, and the group that asked fewer questions won the game. Each team asked about 15 to 20 questions in each round.

Second information gap activity. The second task was identical to the first one except that the participants were told find a key instead of a diamond. The two activities lasted about 30 min.

Verbal Report Session. I conducted stimulated recall sessions with nine participants to understand their perception of recasts to examine whether they recognized the corrective nature of recasts and the linguistic targets. Egi (2008) warned that participants who have taken a posttest and have taken part in stimulated recall may not be able to

provide accurate data during the stimulated recall session, and that participants who take part in stimulated recall prior to a posttest may have their posttest scores skewed. Thus the participants who took part in the stimulated recall session did not take part in any immediate posttest. The entire session was conducted in the learners' L2 and was audio recorded. The participants watched video segments of their own in-class tasks, and they were asked to comment on their thoughts about the interaction, following Gass and Mackey (2000). To elicit the participants' thoughts at that time, the following prompt was given: Tell me what you were thinking about during this interaction. The researcher first identified and selected the video segments that involved recasts for critical segments (75%). Then noncorrective interactions were added as distractors (25%). In Classroom A, 14 verbal recasts on prepositions were provided. In Classroom B, 8 recasts accompanied by gestures were provided.

Procedure

Each data collection took about 2 weeks and was repeated over the two semesters. The treatment was given during the class time, and the pretests, posttests, and verbal report session were done outside the class. On the day of Session 1, the participants completed the background questionnaire, oral production test, and then a grammar test. As mentioned earlier, there were three versions of each test, to avoid a learning effect. The results obtained from a one-way ANOVA revealed that there was no difference among the conditions with regards to the test results in the preposition grammar test, $F(2, 45) = 2.703$, $p = .136$, or the oral production test, $F(2, 45) = .980$, $p = .383$.

One to three days after the pretest, the participants completed the treatment session, which included the two information gap activities. The entire session was videotaped. When the participants did not use prepositions or used incorrect prepositions during the tasks, the researcher provided verbal recasts with or without gestures or provided no recasts, according to the condition to which the participants were assigned. To keep the quality of recasts consistent, the researcher did her best not to stress or emphasize the target structures. The immediate posttest, which included the oral production test and grammar test in a different version from the pretest, was administered a day after the treatment session. A subset of the participants, who were randomly chosen from those who did not complete either the pretest or the immediate posttest, participated in a stimulated recall session instead of the posttests. This posttest was repeated a week after the treatment session as a delayed posttest.

During the information gap activities, which were designed to elicit locative prepositions (*on*, *in*, *under*, *next to*, and *above*), the researcher used a set of specific gestures based on Tutton's (2011) illustration of native English speakers' gestures for *under*, *above*, *on*, and *next to* as used to describe locations. For example, for *in*, the right hand was positioned to show a container, and the left hand, while pointing, moved toward the container to show the concept of containment. Figure 2 illustrates what gestures were used for each preposition.

Analysis

Analysis of Learner Comments. Two major questions were asked in this study. The first research question explored whether students noticed the corrective nature of recasts. Overall, a total of 73 recasts were provided during the communicative tasks. The R condition received a total

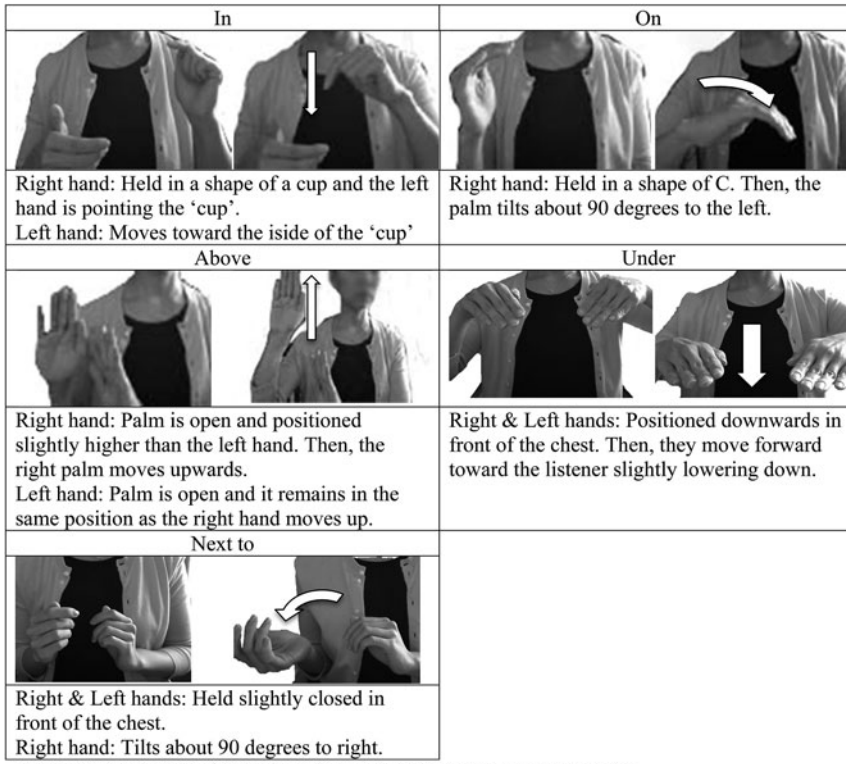

Figure 2. Illustrations of the gestures used in the preposition communicative tasks.

of 38 recasts (an average of 9.5 recasts per class [*SD* = 3.0]), and the RG condition received a total of 35 recasts (an average of 11.75 recasts per class [*SD* = 1.59]). Five participants from Classroom A (R condition) and four from Classroom B (RG condition) completed the stimulated recall session. In Classroom A, 14 verbal recasts on prepositions were provided. In Classroom B, 8 recasts accompanied by gestures were provided. Using a total of 274 comments made by the learners, the frequency of the comments that addressed the corrective nature of the interaction or linguistic errors was calculated.

Analysis of the Grammar and Production Tests. The second research question asked whether the learners used prepositions more readily when the recasts were provided along with gestures. I compared the scores obtained from the grammar test and oral production test in the pretest, immediate posttest, and delayed posttest, using a repeated-measures ANOVA, to identify if any of the groups performed significantly differently from the others. Before interpreting the results, Mauchly's sphericity test was used to verify whether the assumption of sphericity was violated. When it was violated, I used Greenhouse-Geisser adjusted scores. In addition, the effect size was calculated by η^2. The results obtained from η^2 were interpreted using the guideline provided by J. Cohen (1988). When η^2 was around .01, it was interpreted as a small effect size; .06, as a medium effect size; and .14 and above, as a large effect size. When there was a significant difference, I used a Bonferroni multiple post-hoc comparison to see which condition significantly differed from the others. Finally, I conducted one-way ANOVAs to examine at what point (e.g., immediately or at the delayed posttest) the significant change occurred. When there was a significant difference, I again used Bonferroni's multiple comparison to identify which group differed significantly from the others. Then I calculated Cohen's *d* to measure its effect size. The effect size was interpreted using the interpretation guideline in J. Cohen (1988). An effect size of .2 to .3 was interpreted as a small effect; around .5, as a medium effect; and above .8, as a large effect.

All the statistical analyses, except for Cohen's *d*, were conducted using IBM SPSS Statistics Version 21. This program does not allow the inclusion of missing data when conducting repeated-measures ANOVAs. In other words, if a participant had missed any part of the posttest, his or her data could not be included in the analysis. This resulted in a lower number of participants included in the repeated-measures ANOVA. Thus the analysis included 14 participants from the RG condition, 11 from the R condition, and 8 from the control group. However, this restriction did not apply for the one-way ANOVA, and thus all the eligible participants were included in the post-hoc analysis.

RESULTS

Research Question 1: Noticing of Recasts and Gestures

The first research question examined the learners' recognition of the corrective nature of recasts and the noticing of target structures when recasts were accompanied by gestures. Overall, 237 responses were obtained from the nine participants during the stimulated recall session. Based on the content analysis, five major categories were identified: (a) focused on the game results (33.8%), (b) noted the effort required for comprehension (15.2%), (c) did not remember (26.1%), (d) expressed anxiety (6.3%), (e) commented on peer evaluation (3.8%), and (f) other (9.7%).[4] However, none of the comments concerned the nature of error correction. Therefore, I was not able to identify a difference in the quality of noticing between the two experimental conditions at least at the level of conscious awareness.

Research Question 2: Gesture-Incorporated Recasts and L2 Development

The Effect of Types of Recasts and Explicit Linguistic Knowledge. The second research question asked whether the provision of recasts was overall more effective than when recasts were not provided. Table 1 shows the scores of the grammar test, and Figure 3 presents the results visually.

Mauchly's test of sphericity does not assume the sphericity of the data (p = .863); thus Greenhouse-Geisser's adjusted scores were used for the analysis. A repeated-measures ANOVA revealed that there was no significant time effect, $F(1.980, 59.402) = 2.368$, p = .103; no significant group effect, $F(2, 30) = 2.529$, p = .097; and no significant interactive effect of time and group, $F(3.960, 59.402) = .633$, p = .639. Their effect sizes—η^2 = .073, η^2 = .144, and η^2 = .041, respectively—were small to medium overall, except for the interactive effect of time and group.

The Effect of Types of Recasts on the Development of Oral Production. This analysis compares the learners' oral production test scores to identify

Table 1. Mean scores of the preposition grammar test

	Pretest	Immediate posttest	Delayed posttest
Control (n = 8)	9.82 (SD = 1.66)	9.36 (SD = 1.43)	9.27 (SD = 1.19)
Recast (n = 13)	10.18 (SD = 1.38)	9.88 (SD = 1.53)	10.59 (SD = 1.37)
Recast + gesture (n = 27)	10.07 (SD = 1.28)	10.60 (SD = 1.05)	10.07 (SD = 1.03)

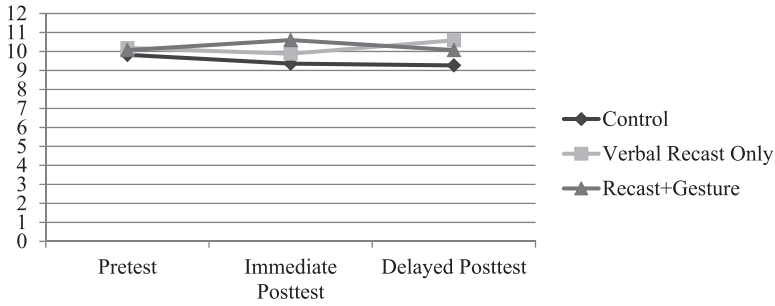

Figure 3. Mean scores of the preposition grammar test.

whether recasts are more beneficial for L2 learning when provided with gestures.

Table 2 shows the scores from the preposition oral production test, and Figure 4 presents the results visually. Figure 4 shows that the scores of the control condition were essentially static. Scores from the R and the RG conditions improved almost equally in the immediate posttest. However, the score of the R condition decreased in the delayed posttest, whereas the score of the RG condition remained the same.

The results from the repeated-measures ANOVA were used to verify if such variation in the performances by the learners in the three conditions was significant. Mauchly's test of sphericity found that the assumption of sphericity was violated (p = .089), so Greenhouse-Geisser's adjusted scores were used. The repeated-measures ANOVA revealed that there was a significant time effect, $F(1.765, 61.788) = 18.705$, p = .000; a significant group effect, $F(2, 35) = 6.356$, p = .004; and a significant interaction effect of time and group, $F(3.532, 61.788) = 3.829$, p = .010. Their effect sizes were consistently large as well, η^2 = .346, η^2 = .266, and η^2 = .180, respectively.

There was a significant difference across the three conditions and a large effect size. Bonferroni multiple comparisons showed that the RG condition outperformed the control condition significantly (p = .004). However, the differences between the R and control conditions (p = .253) and the RG and R conditions (p = .277) were not significant. I further

Table 2. Mean scores of the preposition oral production test

	Pretest	Immediate posttest	Delayed posttest
Control (n = 8)	56.07 (SD = 19.82)	59.72 (SD = 17.73)	58.83 (SD = 14.67)
Verbal recast only (n = 13)	53.79 (SD = 18.67)	79.82 (SD = 16.85)	69.93 (SD = 20.25)
Recast + gesture (n = 27)	62.89 (SD = 12.59)	83.45 (SD = 14.23)	85.25 (SD = 9.25)

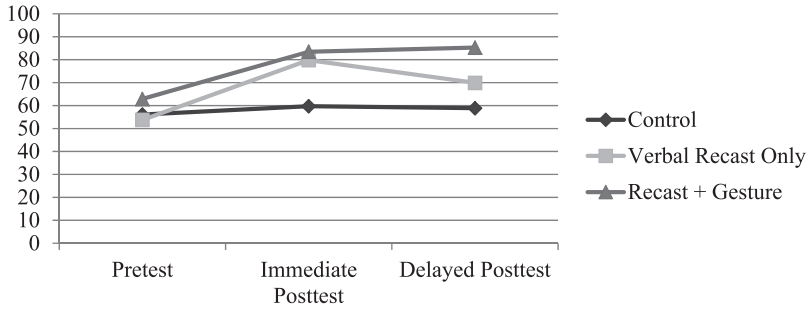

Figure 4. Mean scores of the preposition oral production test.

conducted a one-way ANOVA by using the pretest, posttest, and delayed posttest scores to assess when the learners started varying across the three conditions.[5] The results show that there was no significant difference in their pretest scores, $F(2, 45) = .032$, $p = .969$, and that the effect size was minimal, $\eta^2 = .001$. However, there was a significant difference in the scores from the immediate posttest, $F(2, 45) = 4.947$, $p = .011$, with a strong effect size, $\eta^2 = .180$. Specifically, the R and RG conditions both significantly diverged from the control condition at the time of the immediate posttest ($p = .011$ and $p = .032$, respectively). The effect on the conditions on the immediate posttest was large in the comparison both between the control and R conditions ($d = 1.16$) and between the control and RG conditions ($d = 1.47$). Yet there was no significant difference between the R and RG conditions ($p = 1.00$). The effect size was also small ($d = 0.23$).

Research Question 3: The Long-Term Effect of Gesture-Incorporated Recasts and L2 Development

Finally, the delayed posttest scores were compared. The results from a one-way ANOVA show that there was a significant difference in their delayed posttest scores, $F(2, 35) = 11.73$, $p = .000$, with a large effect size, $\eta^2 = .41$. However, there was no significant difference between the control and verbal conditions ($p = .292$), although there was a medium effect size ($d = .62$). Finally, the RG condition still significantly outperformed both the control condition ($p = .000$), with a large effect size ($d = 2.15$), and the R condition ($p = .015$), also with a large effect size ($d = 0.97$). In sum, the following observations were made:

- In the pretest, there was no significant difference among the three conditions. The effect size was small.

- In the immediate posttest, the R and RG conditions significantly outperformed the control condition. The effect size was large. There was no significant difference between the R and RG conditions, and the effect size was small.
- In the delayed posttest, the RG condition outperformed the R and control conditions significantly, with a large effect size. There was no significant difference between the R and control conditions—presumably due to the small sample size and low power—and the effect size was medium.

DISCUSSION

This intervention study investigated whether or not gestures, when used in addition to verbal recasts, can improve learning locative prepositions, by examining learners' comments on noticing, their explicit knowledge, and the production of locative prepositions.

Learner Noticing during Recasts

The first research question explored learners' noticing during recasts, and the results indicated that the learners did not exhibit any recognition of the corrective nature of either type of recast or noticing of the target structures.

The majority of the comments were about how engaged the learners were in the tasks, and there were no comments on the corrective nature of recasts or about target structures, regardless of the condition. This finding is surprising, given the previous studies used stimulated sessions as a methodology for eliciting learner noticing during classroom interaction (e.g., Carpenter, Jeon, MacGregor, & Mackey, 2006; Egi, 2007; Kim & Han, 2007; Mackey et al., 2000; Mackey, McDonough, Fujii, & Tatsumi, 2001; Polio, Gass, & Chapin, 2006; Roberts, 1995). This study's unexpected finding (i.e., that the learners provided no comments about noticing) may have been due to the linguistic types of the target structures (i.e., locative prepositions), as well as the type of CF (i.e., recasts).

A few studies that investigated learner noticing using stimulated recall sessions reported relatively low rates of correct identification of target structures when feedback was given in the form of recasts. Roberts (1995), for example, reported that learners were able to identify 33% of the errors in full recasts, which were also the kind of recasts used in the present study. Mackey et al. (2000) found that learners did not identify the nature of negative feedback (75% of the negative feedback came from recasts) when the linguistic targets concerned morphosyntax (13% in an ESL classroom and 24% in an Italian as a foreign language classroom). The present findings are in line with

the findings of those two studies in that the amount of comments on noticing (of linguistic targets) was relatively small in the aforementioned studies and nonexistent in the present study. In addition to the justification that is based on previous studies, it is important to address possible methodological issues, such as conducting the recall session in the learners' L2. This issue is revisited in the Limitations section.

Despite such issues, it is worthwhile to mention that the two experimental conditions outperformed the control without any indication of noticing. This could suggest that learning may have occurred without conscious awareness, which is in line with the weak version of the noticing hypothesis. This speculation, however, needs to be interpreted with caution because the absence of comments from the stimulated recall session does not prove that there was a lack of learners' noticing. Stimulated recall is a retrospective self-report. Thus it is possible that noticing occurred during the interaction, but it was not specifically verbalized during the stimulated recall session. This study was not specifically designed with the explicit intention to investigate the necessity of awareness for L2 learning, and the results of a few empirical studies that have examined the possibility of L2 learning without awareness vary to a great extent.

Lastly, to account for the lack of comments from the RG condition, it is important to address the different patterns of processing speech with and without gestures. Studies reported that human brains process verbal sentences differently when they are accompanied by meaningful gestures. Skipper et al. (2007) found that when meaningful gestures were present along with speech, listeners needed less effort to understand the contents of the speech compared to when meaningless gestures were present or when gestures were not present at all. If this is the case, learners in the RG condition may not have experienced noticing at the level of consciousness because they understood the contents of the recasts with less effort as a result of the presence of meaningful gestures. However, a separate study is needed to fully justify this argument.

The Short-Term Effect of Recasts with and without Gestures

The second research question addressed the efficacy of verbal recasts and verbal recasts plus gestures during the immediate posttest, specifically examining learners' spontaneous production of the target structures. The results revealed that both the R and RG conditions significantly surpassed the control condition in the immediate posttest. The learners had a high level of explicit knowledge about locative prepositions at the time of the pretest, which resulted in a ceiling effect that did not allow

for further development. Indeed, the items that were included in the tests targeted the most prototypical meanings of prepositions. Many learners came from countries in which English education is mandatory in junior high school and high school; therefore, it is not surprising that they were able to answer most questions correctly.

The effectiveness of recasts in both conditions seems to have resulted from the fact that the verbal recasts were not necessarily implicit, as the recasts were only provided following the nontargetlike use of locative prepositions—which possibly allowed the learners to pay attention to locative prepositions. This result is partially in line with those recast studies that illustrated the positive effect of recasts on L2 learning (e.g., Han, 2002; Mackey & Philp, 1998; Nassaji, 2009). These studies employed recasts targeting only specific grammatical features. For instance, Han (2002) examined L2 learner development of tense consistency, and Mackey and Philp (1998) investigated the development of question formation; both studies reported that the recast condition surpassed the nonrecast condition. Thus it is possible that targeting one structure may have enhanced the saliency of the recasts. Although there is no proof that this assumption aligns with learners' perception of recasts in the present study, this, arguably, made the recasts more explicit and resulted in better learning outcomes.

The Long-Term Effect of Recasts with and without Gestures

The final research question explored the long-term effect of verbal recasts with and without pedagogical gestures. The results of the spontaneous production assessment showed that the learners from the RG condition retained their development throughout the delayed posttest. However, the development of the R condition diminished after the immediate posttest. This section revisits this result from two perspectives: (a) less retention of learning following verbal recasts and (b) the relationship between memory and gestures.

The Long-Term Effect of Verbal Recasts. The finding that learners who received recasts only verbally did not maintain the development in the delayed posttest contradicts the findings from the meta-analyses by Li (2010) and Mackey and Goo (2007). The researchers in both studies reported that implicit feedback such as recasts provided a long-lasting effect on L2 learning, although it must be pointed out that the number of studies about long-term effects included in Li's meta-analysis was fairly limited. In this study, however, verbal recasts did not follow this pattern. A possible speculation is that the recasts used in the previous studies may have naturally incorporated visual cues such as gestures,

resulting in better performance in the delayed posttests. Indeed, Hudson's (2011) classroom research showed that language teachers do commonly use gestures while teaching grammar. These previous feedback studies did not provide nonverbal information; thus it is not possible to confirm this speculation. Other variables besides the presence or absence of gestures need to be considered with regard to the explicitness of recasts, as Loewen and Philp (2006) reported. It is not surprising that the recasts in the studies in the aforementioned meta-analysis incorporated corrective cues, such as intonation and stress, which may have allowed for a better learning outcome; in contrast, in this study verbal recasts were provided without such oral cues to make the data of the experimental conditions comparable to one another.

Gesture, Visual Input, and Memory Retention. To account for the continued retention of learning by the RG condition as compared to the diminished retention of the R condition, it is important to address the literature that reported the positive impact of looking at gestures on memory retention (e.g., Allen, 1995; R. L. Cohen & Otterbein, 1992; Feyereisen, 1998; Tellier, 2008). Specifically, Allen's (1995) and Tellier's (2008) studies reported that the teacher's gestures and the students' activity of gesturing improved memorization of L2 sentences and vocabulary, respectively. Although these studies did not isolate the effectiveness of receiving gestural visual input in particular, they provide a partial explanation for the retention of the target structures by the RG condition. In addition, several neuroscience studies have reported that seeing gestures facilitates processing of language. They reported that seeing cospeech gestures triggered semantic processing (Holle & Gunter, 2007; Kelly et al. 2004; Kelly et al. 2007; Wu & Coulson, 2005) and facilitated semantic comprehension of speech, which freed up additional resources for processing information (Skipper et al., 2007). It is noteworthy that the target structures of the current study were locative prepositions, which are *depictable* even through gestures. Therefore, learners may have been able to process the information of the CF with less effort with regards to semantic processing when the meaningful gestures were present.

Not limited to gestures as visual input, other studies in the field of psychology have reported the benefits for memory retention of visual input in general over verbal input (e.g., M. A. Cohen, Horowitz, & Wolfe, 2009). M. A. Cohen et al. (2009) examined whether participants' identification of new and old items varies in visual modality (visual images) and audio modality (sound clips) and found that the participants performed significantly better in the task that involved images than in the task that involved any sound type. This supports the contention that auditory memory is systematically inferior to visual memory. Extending the facilitative role of visual input in memorization, some researchers

have examined how multimodality promotes learning. The dual coding theory, proposed by Clark and Paivio (1991), argues that integrating verbal and nonverbal modalities reinforces learning because learners are left with more traces in the memory system after coding the information through different modalities. In short, in this study, learners in the RG condition may have surpassed those in the R condition in the delayed posttest of locative prepositions because they had processed recasts both aurally and visually. This appears to have helped them better retain the information of recasts and resulted in better learning outcomes than those of the participants who were only exposed to verbal recasts.

LIMITATIONS

The present study showed the long-term effect of recasts when gestures were incorporated when teaching locative prepositions. However, as with all studies, it had limitations.

First, although the stimulated recall protocol of this study strictly followed the recommendations of Gass and Mackey (2000), some methodological issues with the stimulated recall session need to be addressed. To begin with, in the present study, the number of learners who completed the stimulated recall session was limited, and stimulated recall sessions were conducted in English. Because of the relatively low proficiency of the participants, the learners might not have been fully able to express their thoughts (Polio & Chiu, 2007). Admittedly, conducting the stimulated recall sessions in their L1s would have prevented such issues; however, the fact that speakers with five different L1s were represented made such an approach unfeasible. Echoing Goo and Mackey (2013), who recommended incorporating triangulated approaches to feedback research, I suggest that researchers combine multiple measurements to fully capture learner noticing. Had the data collection method and context been different—for example, conducted via pair work—implementing additional measurements, such as think-aloud protocols, would have been possible. This study, however, was specifically carried out in a classroom to maintain ecological validity, and this factor limited the inclusion of additional online measures. For classroom studies, further qualitative analysis of learner-instructor interactions, for example, with a focus on the verbal and gestural uptake of learners, may serve as an additional method to assess learner noticing.

The second limitation is that the current study only investigated the learners' development of the use of locative prepositions. To make the results more generalizable, it is important to conduct additional studies on other grammatical features and linguistic targets. Some observational studies reported that gestures were used for teaching pronunciation in

language classrooms (e.g., Hudson, 2011; Tellier, 2006). Therefore, conducting an intervention study with regard to the acquisition of pronunciation would be worthwhile to fully understand the interactive effect of gestures and CF in other linguistic domains.

The third limitation of this study is that the total number of eligible participants was fairly limited. There were a total of 73 ESL learners who agreed to participate in the study initially. However, the attrition rate was relatively high because the learners needed to complete all the assessments outside the class, and each session lasted about 40 min. Because the participation was completely voluntary, the author was unable to control their attendance in the test sessions. The results from the parametric analysis might have been different had the number of participants been larger.

Finally, learners' perception of gestures was not assessed in this study. Despite the positive findings, because of the diverse backgrounds of the learners, some learners may have interpreted the meanings of gestures differently from what was intended. If a longitudinal study were to be conducted, introducing the meaning of gestures and using the relevant gestures on all occasions during the class would be an ideal way to avoid this issue.

CONCLUSIONS AND FUTURE STUDIES

This study examined whether pedagogical gestures used during recasts enhanced the saliency of recasts and whether they lead to learners' better noticing and L2 development of locative prepositions. In short, it was found that the R and RG conditions surpassed the control condition in the immediate posttest, but only the learners in the RG condition retained their improvement in the delayed posttest. Interestingly, regardless of the observed improvement, none of the learners exhibited a sign of noticing during the stimulated recall session.

Based on these findings, I propose that interaction studies need to take the impact of nonverbal features into consideration for more in-depth understanding on what factors contribute to L2 learning. Pedagogically, the findings from this study suggest that gestures have the potential for enhancing the long-term effectiveness of recasts on some linguistic structures. One aspect focused on in teacher preparatory programs is how to provide CF in language classrooms, and the findings of this study suggest the benefits of the inclusion of pedagogical gestures while providing feedback. Researchers in the field of SLA, along with language teachers, must explore teaching from a multimodal perspective to further explore effective language teaching.

For future studies, it is important to explore whether or not gestural CF assists the development of linguistic domains other than prepositions.

Some observational studies reported that gestures were used for teaching pronunciation in language classrooms (e.g., Hudson, 2011; Tellier, 2006). Therefore, conducting an intervention study with regard to the acquisition of pronunciation would be worthwhile to fully understand the interactive effect of gestures and CF in other linguistic domains.

Another aspect that needs further attention is the impact of individual and contextual differences. In the present study, all the participants were selected from a low-intermediate classroom from an ESL institution in the United States. However, other variables such as learners' level of proficiency and cultural backgrounds may impact L2 learning. For instance, exploring the impact of levels of proficiency and participants' attention to gestures during CF would be worthwhile. Learners' visual attention could vary culturally (e.g., whether or not they make eye contact with the instructor), and it would be advisable to explore whether this impacts comprehension of gestures and subsequent learning. Conducting more empirical studies on gesture and L2 learning could certainly provide important insights into these issues.

Received 27 October 2015
Accepted 10 November 2015
Final Version Received 10 November 2015

NOTES

1. A reviewer pointed out the uneven distribution of participants. The author collected additional data for the control and recast conditions specifically for this reason; however, it was beyond the author's control to manipulate the learners' participation in the study.

2. In the RG condition, the participants' L1s were Arabic (6), Chinese (19), and Thai (2). Among them, 17 were male and 10 were female students, and the average age was 20.4 years (*SD* = 1.81). In the R condition, the participants' L1s were Arabic (5), Chinese (6), and Korean (2), and there were 8 male and 5 female students. Their average age was 19.2 (*SD* = 1.09). In the control group, the participants' L1s were Arabic (4) and Chinese (4), and there were 3 male and 5 female students. The average age was 22.0 (*SD* = 2.00).

3. Originally, the items included 12 critical items and 8 distracters; however, due to time constraints, it was necessary to reduce the number of items.

4. In both the R and RG conditions, the most frequent comments were about focusing on the game (30.4% and 38.2%, respectively), followed by comments on the learner's lack of memory (26.7% and 25.4%, respectively) and comprehension effort (15.6% and 14.7%, respectively).

5. Following a reviewer's comment, the author conducted a separate post hoc analysis using a gain score. The results followed the same pattern as the one in the results section. The gain score from the pretest to immediate posttest showed that the R and RG conditions outperformed the control significantly (p = .044 and p = .029, respectively), but there was no difference between the R and RG conditions (p = 1.000). As for the gain score to the delayed posttest, only the RG condition significantly outperformed the control and R conditions (p = .009 and p = .052, respectively), and there was no significant difference between control and R conditions (p = 1.000).

REFERENCES

Alibali, M., Flevares, L., & Goldin-Meadow, S. (1997). Assessing knowledge conveyed in gesture: Do teachers have the upper hand? *Journal of Educational Psychology*, *23*, 183–194.

Allen, L. Q. (1995). The effects of emblematic features on the development and access of mental representations of French expressions. *Modern Language Journal*, *79*, 521–529.

Allen, L. Q. (1999). Functions of nonverbal communication in teaching and learning a foreign language. *The French Review*, *72*(3), 469–480.

Allen, L. Q. (2000). Form-meaning connections and the French causative. *Studies in Second Language Acquisition*, *22*, 69–84.

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

Carpenter, H., Jeon, K., MacGregor, D., & Mackey, A. (2006). Learners' interpretations of recasts. *Studies in Second Language Acquisition*, *28*, 209–236.

Carroll, S. E. (1999). Putting "input" in its proper place. *Second Language Research*, *15*, 337–388.

Celce-Murcia, M., & Larsen-Freeman, D. (1998). *The grammar book: An ESL/EFL teacher's course* (2nd ed.). Boston: Heinle.

Church, R. B. (1999). Using gesture and speech to capture transitions. *Cognitive Development*, *14*, 313–342.

Church, R. B., Ayman-Nolley, S., & Mahootian, S. (2004). The role of gesture in bilingual education: Does gesture enhance learning? *International Journal of Bilingual Education and Bilingualism*, *7*, 303–319.

Clark, J. M., & Paivio, A. (1991). Dual coding theory and education. *Educational Psychology Review*, *3*(3), 149–170.

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

Cohen, M. A., Horowitz, T. S., & Wolfe, J. M. (2009). Auditory recognition memory is inferior to visual recognition memory. *Proceedings of National Academy of Science U.S.A.*, *106*, 6008–6010.

Cohen, R. L., & Otterbein, N. (1992). The mnemonic effect of speech gestures: Pantomimic and non-pantomimic gestures compared. *European Journal of Cognitive Psychology*, *4*, 113–139.

Cook, S. W., Yip, T. K., & Goldin-Meadow, S. (2010). Gesturing makes memories that last. *Journal of Memory and Language*, *63*, 465–475.

Crowder, E. M. (1996). Gestures at work in sense-making science talk. *Journal of Learning Sciences*, *5*, 173–208.

Davies, M. (2006). Paralinguistic focus on form. *TESOL Quarterly*, *40*(4), 841–855.

Doughty, C. (1994). Fine-tuning of feedback by competent speakers to language learners. In J. E. Alatis (Ed.), *Georgetown University roundtable on language and linguistics* (pp. 96–108). Washington, DC: Georgetown University Press.

Egi, T. (2007). Recasts, learners' interpretation, and L2 development. In A. Mackey (Ed.), *Conversational interaction in second language acquisition* (pp. 249–267). Oxford: Oxford University Press.

Egi, T. (2008). Investigating stimulated recall as a cognitive measure: Reactivity and verbal reports in SLA research methodology. *Language Awareness*, *17*, 212–217.

Ellis, R. (2003). *Task-based language learning and teaching*. Oxford: Oxford University Press.

Ellis, R., & Sheen, Y. (2006). Re-examining the role of recasts in L2 acquisition. *Studies in Second Language Acquisition*, *28*, 576–600.

Feyereisen, P. (1998). Le rôle des gestes dans la mémorisation d'énoncés oraux. In S. Santi, I. Guaïtella, C. Cavé, & G. Konopczynski (Eds.), *Oralité et gestualité. Communication multimodale, interaction. Actes du colloque Orage 98* (pp. 355–360). Paris: L'Harmattan.

Flevares, L., & Perry, M. (2001). How many do you see? The use of nonspoken representations in first grade mathematics classes. *Journal of Educational Psychology*, *93*(2), 330–345.

Gass, S. (1997). *Input, interaction, and the language learner*. Mahwah, NJ: Lawrence Erlbaum.

Gass, S., & Mackey, A. (2000). *Stimulated recall methodology in second language research*. Mahwah, NJ: Erlbaum.

Gass, S., Svetics, I., & Lemelin, S. (2003). Differential effects of attention. *Language Learning*, *53*(3), 497–546.
Gass, S. M., & Varonis, E. M. (1994). Input, interaction, and second language production. *Studies in Second Language Acquisition*, *16*, 283–302.
Godfroid, A., Boers, F., & Housen, A. (2013). An eye for words. *Studies in Second Language Acquisition*, *35*, 483–517.
Goldin-Meadow, S. (2004). Gesture's role in the learning process. *Theory into Practice*, *43*, 314–321.
Goldin-Meadow, S., Cook, S. W., & Mitchell, Z. A. (2009). Gesturing gives children new ideas about math. *Psychological Science*, *20*(3), 267–272.
Golden-Meadow, S., Kim, S., & Singer, M. A. (1999). What the teacher's hands tell the student's mind about math. *Journal of Educational Psychology*, *91*, 720–730.
Goldin-Meadow, S., & Sandhofer, C. M. (1999). Gesture conveys substantive information to ordinary listeners. *Developmental Science*, *2*, 67–74.
Goldin-Meadow, S., & Singer, M. A. (2003). From children's hands to adults' ears: Gesture's role in the learning process. *Developmental Psychology*, *39*, 509–520.
Goo, J. (2012). Corrective feedback and working memory capacity in interaction-driven L2 learning. *Studies in Second Language Acquisition*, *34*, 445–474.
Goo, J., & Mackey, A. (2013). The case against the case against recasts. *Studies in Second Language Acquisition*, *35*, 127–165.
Gullberg, M. (2010). Methodological reflections on gesture analysis in second language acquisition and bilingualism research. *Second Language Research*, *26*, 75–102.
Hama, M., & Leow, R. (2010). Learning without awareness: Revisited: Extending Williams (2005). *Studies in Second Language Acquisition*, *32*, 465–491.
Han, Z. (2002). A study of the impact of recasts on tense consistency in L2 output. *TESOL Quarterly*, *36*, 542–572.
Havranek, G. (1999). The effectiveness of corrective feedback: Preliminary results of an empirical study. *Acquisition et Interaction en Langue Étrangère*, *2*, 189–206.
Holle, H., & Gunter, T. C. (2007). The role of iconic gestures in speech disambiguation: ERP evidence. *Journal of Cognitive Neuroscience*, *19*, 1175–1192.
Hudson, N. (2011). *Teacher gesture in a post-secondary English as a second language classroom: A sociocultural approach* (Doctoral dissertation). University of Nevada, Las Vegas, Nevada.
Izumi, S. (2002). Output, input enhancement, and the noticing hypothesis. *Studies in Second Language Acquisition*, *24*(4), 541–577.
Jarvis, S., & Odlin, T. (2000). Morphological type, spatial reference, and language transfer. *Studies in Second Language Acquisition*, *22*, 535–556.
Kamiya, N. (2015). The effectiveness of intensive and extensive recasts on L2 acquisition for implicit and explicit knowledge. *Linguistics and Education*, *29*, 59–72.
Kao, R. R. (2001). Where have the prepositions gone? A study of English prepositional verbs and input enhancement in instructed SLA. *IRAL*, *39*, 195–216.
Kelly, S. D., Kravitz, C., & Hopkins, M. (2004). Neural correlates of bimodal speech and gesture comprehension. *Brain and Language*, *89*, 253–260.
Kelly, S. D., Ward, S., Creigh, P., & Bartolotti, J. (2007). An intentional stance modulates the integration of gesture and speech during comprehension. *Brain and Language*, *101*, 222– 233.
Kim, J., & Han, Z. (2007). Recasts in communicative ESL classes: Do teacher intent and teacher interpretation overlap? In A. Mackey (Ed.), *Conversational interaction in second language acquisition: A collection of empirical studies* (pp. 378–403). Oxford: Oxford University Press.
Lazaraton, A. (2004). Gesture and speech in the vocabulary explanations of one ESL teacher: A microanalytic inquiry. *Language Learning*, *54*(1), 79–117.
Leeman, J. (2003). Recasts and second language development. *Studies in Second Language Acquisition*, *25*, 37–63.
Leow, R. (2000). A study of the role of awareness in foreign language behavior: Aware vs. unaware learners. *Studies in Second Language Acquisition*, *22*, 557–584.
Li, S. (2010). The effectiveness of corrective feedback in SLA: A meta-analysis. *Language Learning*, *60*, 309–365.
Li, S. (2013). The interactions between the effects of implicit and explicit feedback and individual differences in language analytic ability and working memory. *The Modern Language Journal*, *97*(3), 634–654.

Lindstromberg, S. (2010). *English prepositions explained* (2nd ed.). Amsterdam/Philadelphia: John Benjamins.
Loewen, S., & Nabei, T. (2007). Measuring the effects of oral corrective feedback on L2 knowledge. In A. Mackey (Ed.), *Conversational interaction in second language acquisition* (Vol. 339–360). Oxford: Oxford University Press.
Loewen, S., & Philp, J. (2006). Recasts in the adult L2 classroom: Characteristics, explicitness, and effectiveness. *Modern Language Journal*, *20*, 1–67.
Long, M. (1996). The role of the linguistic environment in second language acquisition. In W. C. Ritchie & T. K. Bhatia (Eds.), *Handbook of second language acquisition* (pp. 413–468). New York: Academic Press.
Long, M. (2007). *Problems in SLA*. Mahwah, NJ: Erlbaum.
Long, M., Inagaki, S., & Ortega, L. (1998). The role of implicit negative feedback in SLA: Models and recasts in Japanese and Spanish. *Modern Language Journal*, *82*, 357–371.
Lyster, R. (1998). Negotiation of form, recasts, and explicit correction in relation to error types and learner repair in immersion classrooms. *Language Learning*, *48*, 183–218.
Lyster, R. (2001). Negotiation of form, recasts, and explicit correction in relation to error types and learner repair in immersion classrooms. *Language Learning*, *51*, 265–301.
Lyster, R. (2004). Differential effects of prompts and recasts in form-focused instruction. *Studies in Second Language Acquisition*, *26*, 399–432.
Lyster, R., & Mori, H. (2006). Interactional feedback and instructional counterbalance. *Studies in Second Language Acquisition*, *28*, 321–341.
Lyster, R., & Ranta, L. (1997). Corrective feedback and learner uptake. *Studies in second language acquisition*, *19*(1), 37–66.
Lyster, R., & Ranta, L. (2013). Counterpoint piece: The case for variety in corrective feedback research. *Studies in Second Language Acquisition*, *35*, 167–184.
Lyster, R., Saito, K., & Sato, M. (2013). Oral corrective feedback in second language classrooms. *Language Teaching*, *46*, 1–40.
Mackey, A. (2006). Feedback, noticing and instructed second language learning. *Applied Linguistics*, *27*, 405–430.
Mackey, A., Gass, S., & McDonough, K. (2000). How do learners perceive interactional feedback? *Studies in Second Language Acquisition*, *22*, 471–497.
Mackey, A., & Goo, J. (2007). Interaction research in SLA: A meta-analysis and research synthesis. In A. Mackey (Ed.), *Conversational interaction in second language acquisition* (pp. 407–452). Oxford: Oxford University Press.
Mackey, A., McDonough, K., Fujii, A., & Tatsumi, T. (2001). Investigating learners' reports about the L2 classroom. *International Review of Applied Linguistics*, *39*, 285–308.
Mackey, A., & Philp, J. (1998). Conversational interaction and second language development: Recasts, responses and red herrings. *Modern Language Journal*, *82*, 338–356.
McNeill, D. (1992). *Hand and mind: What gestures reveal about thought*. Chicago: University of Chicago Press.
Nassaji, H. (2009). Effects of recasts and elicitations in dyadic interaction and the role of feedback explicitness. *Language Learning*, *59*(2), 411–442.
Nicholas, H., Lightbown, P., & Spada, N. (2001). Recasts as feedback to language learners. *Language Learning*, *51*, 719–758.
Oliver, R. (1995). Negative feedback in child NS-NNS conversation. *Studies in Second Language Acquisition*, *17*, 459–481.
Peltier, I. N., & McCafferty, S. G. (2010). Gesture and identity in the teaching and learning of Italian. *Mind, Culture, and Activity*, *17*(4), 331–349.
Perry, M., Birch, D., & Singleton, J. (1995). Constructing shared understanding: The role of nonverbal input in learning contexts. *Contemporary Legal Issues*, *6*, 213–235.
Philp, J. (2003). Constraints on "noticing the gap." *Studies in Second Language Acquisition*, *25*, 99–126.
Philp, J., & Mackey, A. (2010). Interaction research: What can socially informed approaches offer to cognitivists (and vice versa)? In R. Batstone (Ed.), *Sociocognitive perspectives on language use and language learning* (pp. 210–224). Oxford: Oxford University Press.
Polio, C., & Chiu, S. (2007). *Reactivity, veridicality, and language choice in L2 writing concurrent verbal protocols*. Paper presented at the American Association of Applied Linguistics Conference, Costa Mesa, CA.

Polio, C., Gass, S., & Chapin, L. (2006). Using stimulated recall to investigate speaker perceptions in native-nonnative speaker interaction. *Studies in Second Language Acquisition*, *28*, 237–267.
Roberts, M. A. (1995). Awareness and the efficacy of error correction. In R. Schmidt (Ed.), *Attention and awareness in foreign language learning* (pp. 163–182). Honolulu: University of Hawai'i, Second Language Teaching and Curriculum Center.
Robinson, P. (1995). Attention, memory, and the "noticing" hypothesis. *Language Learning*, *45*, 283–331.
Robinson, P. (2001). *Cognition and second language instruction*. Cambridge: Cambridge University Press.
Robinson, P. (2003). Attention and Memory during SLA, in C. J. Doughty & M. H. Long (Eds.), *The Handbook of Second Language Acquisition* (pp. 631–678), Blackwell Publishing Ltd, Oxford, UK.
Roth, W. (2001). Gestures: Their role in teaching and learning. *Review of Educational Research*, *71*, 365–392.
Roth, W., & Lawless, D. V. (2002a). Signs, deixis, and the emergence of scientific explanations. *Semiotica*, *138*, 95–130.
Roth, W., & Lawless, D. V. (2002b). When up is down and down is up: Body orientation, proximity, and gestures as resources. *Language in Society*, *31*, 1–28.
Schmidt, R. (1990). The role of consciousness in second language learning. *Applied Linguistics*, *11*, 129–158.
Schmidt, R. (1993). Awareness and second language acquisition. *Annual Review of Applied Linguistics*, *13*, 206–226.
Schmidt, R. (1994). Deconstructing consciousness in search of useful definitions for applied linguistics. *Consciousness in Second Language Learning*, *11*, 237–326.
Schmidt, R. (1995). Consciousness and foreign language learning: A tutorial on the role of attention and awareness in learning. In R. Schmidt (Ed.), *Attention and awareness in foreign language learning* (pp. 1–66). Honolulu: University of Hawaii at Manoa.
Schmidt, R. (2001). Attention. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 3–32). New York: Cambridge University Press.
Schmidt, R., & Frota, S. (1986). Developing basic conversational ability in a second language: A case study of an adult learner of Portuguese. In R. Day (Ed.), *Talking to learn: Conversation in a second language* (pp. 237–326). Rowley, MA: Newbury House.
Sheen, Y. (2004). Corrective feedback and learner uptake in communicative classrooms across instructional settings. *Language Teaching Research*, *8*(3), 263–300.
Sheen, Y. (2006). Exploring the relationship between characteristics of recasts and learner uptake. *Language Teaching Research*, *10*, 361–392.
Sheen, Y. (2007). Corrective feedback, individual differences, and acquisition of articles. In A. Mackey (Ed.), *Conversational interaction in second language acquisition* (pp. 301–360). Oxford: Oxford University Press.
Sime, D. (2006). What do learners make of teachers' gestures in the language classroom? *International Review of Applied Linguistics*, *44*, 211–230.
Skipper, J., Goldin-Meadow, S., Nusbaum, H., & Small, S. (2007). Speech-associated gestures, Broca's area, and the human mirror system. *Brain and Language*, *101*, 260–277.
Slimani, A. (1992). Evaluation of classroom interaction. In J. C. Anderson & A. Beretta (Eds.), *Evaluating second language education* (pp. 197–221). Cambridge: Cambridge University Press.
Sueyoshi, A., & Hardison, D. (2005). The role of gestures and facial cues in second language listening comprehension. *Language Learning*, *55*, 579–596.
Swain, M. (2005). The output hypothesis: Theory and research. *Handbook of Research in Second Language Teaching and Learning*, *1*, 471–483.
Swain, M., & Lapkin, S. (2002). Talking it through: Two French immersion learners' response to reformulation. *International Journal of Educational Research*, *37*(3), 285–304.
Tellier, M. (2006). *L'impact du geste pédagogique sur l'enseignement/apprentissage des langues étrangères: Etude sur des enfants de 5 ans* (Doctoral dissertation). Université Paris VII, Paris, France.
Tellier, M. (2008). The effect of gestures on second language memorization by young children. *Gesture*, *8*, 219–235.
Tomlin, R., & Villa, V. (1994). Attention in cognitive science and second language acquisition. *Studies in Second Language Acquisition*, *16*, 183–204.

Truscott, J. (1999). The case for "The case against grammar correction in L2 writing classes": A response to Ferris. *Journal of Second Language Writing*, *8*, 111–122.
Tutton, M. (2011). How speakers gesture when encoding location with English on and French sur. *Journal of Pragmatics*, *43*, 3431–3454.
Valenzeno, L., Alibali, M., & Klatzky, R. (2003). Teachers' gestures facilitate students' learning: A lesson in symmetry. *Contemporary Educational Psychology*, *28*, 187–204.
Wang, W. (2009). *The noticing and effect of teacher feedback in ESL classrooms* (Doctoral dissertation). Michigan State University, USA.
Williams, J. (2005). Learning without awareness. *Studies in Second Language Acquisition*, *27*, 269–304.
Wu, Y. C., & Coulson, S. (2005). Meaningful gestures: Electrophysiological indices of iconic gesture comprehension. *Psychophysiology*, *42*, 654–667.
Yang, Y., & Lyster, R. (2010). Effects of form-focused practice and feedback on Chinese EFL learner' acquisition of regular and irregular past tense forms. *Studies in Second Language Acquisition*, *32*, 235–263.
Yilmaz, Y. (2013). Relative effects of explicit and implicit feedback: The role of working memory capacity and language analytic ability. *Applied Linguistics*, *34*(3), 344–368.
Zhao, J. (2007). *Metaphors and gestures for abstract concepts in academic English writing* (Doctoral dissertation). University of Arizona, USA.