

## CHAPTER 5

### Gender and recasts

#### Analysis of males' and females' L2 development following verbal and gesture-enhanced recasts

Kimi Nakatsukasa  
Texas Tech University

This study investigated the influence of learners' gender on the effectiveness of verbal recasts and gesture-enhanced recasts. A total of 40 adult English-as-a-second-language learners participated. In the first condition, the learners received verbal recasts during two communicative tasks. The results showed no difference across gender. In the second condition, the learners received gesture-enhanced recasts. Females significantly outperformed males with a strong effect size in the immediate to delayed posttests gain score but not in the gain score of pretest to immediate posttests. This indicates that the long-term effectiveness of recasts seems to be influenced by learners' gender. The results are discussed in relation to learners' memory and the exposure to visual cues.

**Keywords:** communicative tasks, gesture, learner gender, recasts

#### Introduction

This study investigates whether or not the effectiveness of verbal recasts and visually-enhanced recasts vary between male and female learners by bridging two gaps in the existing literature. The first gap deals with the impact of gender differences in second language (L2) learning. Psychological and neurolinguistic studies have illustrated that gender plays an integral role in L2 learning because there is a significant difference in males' and females' attitude toward language learning and in short-term and long-term memory. Specifically, with regard to males' and females' cognitive systems, it has been reported that females surpass males in the use of both memory systems. However, whether or not such differences in cognition influence interlanguage (IL) development differently in communicative tasks with recasts in a classroom setting has not been investigated. The second contribution of this study deals with the effectiveness of visually-enhanced recasts. For example, examining recasts has shown

that seeing visual cues during recasts helped learners' retention of learned target items (Nakatsukasa, 2016). This coincides with studies in psychology, which have argued that visual cues such as gestures facilitate learners' memory retention and lead to long-term effects of learning (e.g., Macedonia & Klimesch, 2014). However, whether or not the exposure to visual cues works equally for males and females is yet to be explored. To investigate these issues, this paper examines male and female adult English-as-a-second-language (ESL) learners' L2 development following verbal recasts and gesture-enhanced recasts.

## Background

This section will detail the aforementioned gap in the field, that is, lack of analysis of learners' gender in the effectiveness of corrective feedback, by first summarizing the existing research on recasts, gesture, as well as gender and memory.

### Recasts

In language classrooms, instructors often employ recasts as a common type of corrective feedback (e.g., Doughty, 1994; Havranek, 1999; Lyster & Ranta, 1997; Mackey, Gass, & McDonough, 2000; see also Oliver, 1995) given that they do not interrupt the flow of communication and because they save time spent on correction (Ellis & Sheen, 2006; Han, 2002; Leeman, 2003; Loewen & Philp, 2006). Recasts include target-like models without explicitly telling learners that their initial production was non-target like (e.g., Ellis & Sheen, 2006). Long (2007) argued that this feature encourages learners to notice the gap between their interlanguage (IL) and target-like production, which, then, leads to IL development. Furthermore, recasts promote form-function mapping because they are given in a moment when a learner's production is questioned (Long, 1996, 2007). The effectiveness of implicit feedback, including recasts, was confirmed in Li's (2010) meta-analysis. Examining 33 corrective feedback studies, Li found that the benefit of implicit feedback was retained substantially longer than the explicit feedback, even though the explicit feedback had a stronger immediate effect.

It is worth mentioning that the noticeability of recasts is still debated. A number of empirical studies (e.g., Mackey & Philp, 1998; Nassaji, 2009; Yang & Lyster, 2010) and recent meta-analyses (e.g., Li, 2010; Mackey & Goo, 2007) have shown that recasts facilitate L2 development. Several studies reported the opposite findings (e.g., Ammar & Spada, 2006; Lyster, 2001, 2004; Sheen, 2004; Slimani, 1992), and still others were inconclusive (e.g., Loewen & Nabai, 2007; Long, Inagaki, & Ortega, 1998).

### *Characteristics of recasts*

Studies investigating characteristics of oral recasts provide further explanation about the reasons for the inconsistency of findings on the effectiveness (or ineffectiveness) of recasts. For example, some studies have reported that the saliency and noticeability of recasts and their linguistic targets are crucial factors that determine the effectiveness of recasts (e.g., Ellis & Sheen, 2006; Kamiya, 2015; Loewen & Philp, 2006). Specifically, in their 17-hour-long observational study, Loewen and Philp (2006) suggested that discourse features, such as intonation, stress, and a number of feedback moves, influence the saliency and noticeability of recasts. Following this argument, it is logical to assume that these nonlinguistic features may also influence recasts' saliency and subsequent L2 development.

### *Recasts and individual differences*

In addition to the aforementioned characteristics of recasts, another factor to consider is the influence of learners' individual differences on the effectiveness of recasts. Studies have found the following factors to have a significant impact on the efficacy of recasts: learners' proficiency levels and readiness to learn a target structure (e.g., Ammar & Spada, 2006; Li, 2010; Mackey & Philp, 1998; Nicholas, Lightbown, & Spada, 2001; Philp & Mackey, 2010), cognitive factors such as working memory capacity (e.g., Goo, 2012; Yilmaz, 2013), and psychological factors such as anxiety (Sheen, 2008). Yet, to the best of my knowledge, no work has been done in relation to learners' gender.

### *Gesture and L2 learning*

In the field of SLA, descriptive studies have illustrated how gestures have been used in L2 classrooms, specifically in relation to teaching pronunciation (e.g., Hudson, 2011; Smotrova, 2015), vocabulary items (e.g., Smotrova & Lantolf, 2013), and grammatical features (e.g., Hudson, 2011). In addition, experimental studies have reported that seeing the instructor's gestures helps learners with vocabulary acquisition (e.g., Macedonia & Klimesch, 2014; Tellier, 2008, 2010), expression learning (Allen, 1995), and comprehension (e.g., Sime, 2006; Sueyoshi & Hardison, 2005; see, however, Kelly & Lee, 2012, and Kelly, McDevitt, & Esch, 2009, for the mixed findings in terms of learning L2 pronunciation). Macedonia and Klimesch (2014) presented learners with 36 vocabulary items in an audio-visual condition (reading, hearing, and speaking) or gesture condition (audio-visual plus seeing gestures) over 14 months. They reported that the gesture condition surpassed the audio-visual condition and argued that the target items were better integrated to long-term memory because of the increased number of modalities. Studies in the field of

psychology also support the significance of multimodality (e.g., Cohen, Horowitz, & Wolfe, 2009). For example, 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.

Even though the general consensus is that the instructor's gestures help L2 learning, so far, only a few studies have addressed the use of gestures during corrective feedback, including recasts. Overall, the descriptive studies have illustrated how a language instructor incorporates gestures during the provision of feedback (e.g., Faraco & Kida, 2008; Wang & Loewen, 2015). For example, Wang and Loewen (2015) observed about 65 hours of an ESL classroom and identified that more than 60% of corrective feedback was accompanied by various nonverbal behaviors. Specifically, as for the explicitness of feedback, they found that the more explicit feedback was accompanied by nonverbal behaviors. In a similar line of study, Davies (2006) also observed how frequently nonverbal cues were used during focus-on-form episodes (FFEs) and reported that 47% of FFEs were accompanied by nonverbal cues.

These descriptive studies have illustrated that instructors, indeed, use nonverbal features during the provision of corrective feedback; however, the number of experimental studies has been extremely limited. As for the empirical studies, Nakatsukasa (2016) is the only study thus far that has investigated the effectiveness of gestures accompanied by corrective feedback. Comparing the effectiveness of verbal recasts with gestures (gesture-enhanced recasts) and without gestures during the acquisition of prepositions in an ESL classroom, the study found that learners who received verbal recasts with gestures showed better linguistic development, and the learning was maintained better in a delayed posttest.

### Gender and L2 learning

In the past decade, the field of SLA has investigated what types of learner factors influence the acquisition of L2. The reported factors include biological factors, such as gender, working memory, age, aptitude, and learners' psychological factors, including motivation and anxiety (e.g., Bowden, Sanz, & Stafford, 2005; Matsuda & Gobel, 2004; Robinson, 2010).

Gender, which is the key variable in the present study, seems to influence L2 learning both psychologically and cognitively. As for learners' psychology, many studies addressed gender as a significant variable in relation to learners' motivation based on societal bias (e.g., Kissau, 2006; Kissau, Kolano, & Wang, 2010; Kissau & Turnbull, 2008; Pajares & Valiante, 2001; Pajares, Valiante, & Cheoung, 2007).

learning strategies (e.g., Green & Oxford, 1995; Oxford, Nyikos, & Ehrman, 1988), and goals (e.g., Koul, Roy, Kaewkuekool, & Plaisawaschai, 2009). With regard to learners' L2 ability, overall, females' superior performance than males has been reported, including with respect to L2 comprehension (Payne & Lynn, 2011), classroom performance (Matsuda & Gobel, 2004), and grammar learning (Wucherer & Reiterer, 2016). Regarding phonological recognition and production, however, several studies have reported that males outperformed females (Lim, 1994; Reiterer, Hu, Erb, Rota, Nardo, Grodd, Winkler, & Ackermann, 2011; Wucherer & Reiterer, 2016). These studies illustrate that the overall differences between males and females in academic performance are due to both psychological and social factors.

### Gender and memory

In addition to the aforementioned psychological and social factors, researchers have tapped into different mechanisms in males' and females' long-term and short-term memory. Overall, the studies report an advantage for females in both memory systems, as detailed in the following paragraph. Because the present study deals with the acquisition of locative prepositions, which requires learners to remember the association of each lexical item and its meaning, and also because visual cues seem to play an integral role in learners' memory. In this section, I review studies which investigated the relationship between memory and gender.

Both short-term and long-term memory play an integral role in language learning. Some neurolinguistic studies have attempted to account for females' linguistic advantage by examining whether short- and long-term memory is gender specific, the studies have shown females to be better than men on the tasks that involve short-term memory. The tasks in the previous studies required participants to retain information for a brief period of time, including numbers (Jensen & Reynolds, 1983; Kail & Siegel, 1978, as cited in Kaushanskaya, Marian, & Yoo, 2011) and vocabulary words (Bleeker, Bolla-Wilson, Agnew, & Meyers, 1988; Kramer, Delis, Kaplan, O'Donnell, & Prifitera, 1997). As for long-term memory, researchers have compared males' and females' use of declarative memory, which is one of the two types of long-term memory and requires explicit learning. The studies that Ullman and colleagues (e.g., Hartshorne & Ullman, 2006; Ullman, Miranda, & Travers, 2007) have conducted have shown females to be better at linguistic tasks that use a declarative memory system, such as lexical retrieval and grammar formation.

To summarize the existing research, the following issues need further investigation: First, gesture has been used frequently during corrective feedback in language classrooms, and gesture-enhanced recasts seem to lead to better retention, enhancing learning outcomes. The neurolinguistic accounts seem to explain females' superiority in L2 ability. However, the nature of the previous literature was

highly controlled and experimental especially in the case of the neurological studies. What is still unknown is whether the aforementioned gender-driven differences become apparent in L2 development in a daily language classroom. It is logical to hypothesize that female learners will benefit better from verbal recasts due to their cognitive advantages as they can potentially better remember the error in the original production using their short-term memory and better retain their learning using their long-term memory. As gesture-enhanced recasts seem to facilitate better retention of learning via a more effective use of long-term memory in general, it is possible that gestures further emphasize females' long-term memory.

A second possibility, of course, is that this addition of gestures may facilitate better long-term memory for males, perhaps even to the point of being similar to females. Based on the existing studies, it is not possible to hypothesize which gender better benefits from seeing gestures during the recasts, but it is a question worthy of empirical investigation.

### The present study

The following research questions were formulated to address these topics as reviewed in the previous section:

1. Does the effectiveness of verbal recasts vary between males and females?
2. Does the effectiveness of gesture-enhanced recasts vary between males and females?

### Method

The data used in this study is taken from a larger data set (which was also used for Nakatsukasa, 2016), which investigated the effectiveness of gestural recasts and recasts on the acquisition of locative prepositions. The results from the 2016 study showed that learners were better able to retain their learning in the long-term when recasts were given with gestural cues. Based on the previous studies that have shown the difference in learners' psychology and memory system of males and females, the present study reanalyzes a part of the same data in order to investigate if gender plays a role in L2 development through the communicative tasks.

### Participants

All of the participants were enrolled in low-intermediate ESL classes at a large state university in the United States. Out of a total of 69 adult ESL learners who signed up for the study, 40 participants passed the initial screening test and completed all of the sessions. The oral production pretest, which is described later in this section, served as a screening test and only learners who achieved less than 80% were included for this study in order to avoid a potential ceiling effect. Among those, 13 (six females, seven males) received verbal recasts, and the remaining 27 (13 females, 14 males) received gesture-enhanced recasts. All of the participants were new arrivals to the United States. It was the first semester for most of the students ( $n = 37$ ) or the second semester for a few ( $n = 3$ ). The participants' L1s were Arabic, Chinese, Korean, and Thai, and their average age was 20.4 years ( $SD = 1.88$ ).

### Target structure

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, & Williams, 1983) and English learners (e.g., Lindstromberg, 2010). For example, Jarvis and Odlin (2000) examined how L1 Swedish and Finnish speakers use English locative prepositions and argued that their omission and overgeneralization of prepositions are due to 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 referent (e.g., referent is seen as a container for *in* but a flat surface for *on*), direction of the axis (e.g., vertical for *above* but horizontal for *next to*), transitivity (e.g., *crashed into* is transitive, but *go in* can be transitive and intransitive). Such complex dimensions of prepositions also seem to result in difficulty acquiring L2 prepositions.

Among many meanings of each preposition, locative prepositions are “depictable,” although the degree of “depictability” varies. For example, it is easier to depict “A cup is on the table” because it follows the prototype of schema of *on* – that the object (a cup) is placed on top of the referent (the table). However, “A spider is on the ceiling” is more difficult because the schema needs to be flipped horizontally to understand the situation (Lindstromberg, 2010). 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 stages: (1) pretest, (2) 60-min treatment sessions, (3) immediate posttest within 24 hours of the treatment sessions, and (4) delayed posttest seven days after the treatment sessions. The detailed procedure is presented in the next section. The materials used in this study are identical to Nakatsukasa (2016).<sup>1</sup>

### Oral production test

Three versions of PowerPoint slides were developed, each of which was composed of two practice questions, eight questions, and four distractors to be used for the oral production tests. In the slides, the participants viewed an image of a room, and they were asked to focus on one specific item (e.g., "Please find the birds"). Then, they were asked to verbally respond to a prompt asking for a location of a specific object (e.g., "Where are the birds?"). Each question was created to elicit the prepositions, such as *above*, *under*, *in*, *on*, and *next to*. These responses were audio recorded using a voice recorder. The distractor questions were also included to deter learners' attention to locations, and they were asked questions that were not related to the locations of the items (e.g., "Please find a poster," then, "What is the title of the movie?"). 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.

### Communicative tasks

A total of two communicative tasks were designed for this study to elicit prepositions in class. The tasks were constructed based on Ellis's (2003) definition of focused tasks – that is, a task that includes a gap and a clearly defined outcome. In this study, the focus was on the use of locative prepositions, as well as meaningful interaction.

In the first task, the participants were divided into two groups. Group A was asked to pretend to be a group of burglars and hide a diamond in the least likely place to be found. Group B was asked to pretend to be a detective team and find the diamond. A floor plan of a house was attached to the blackboard, and each team was also given a letter-sized print of the same floor plan.

The researcher told participants that they would switch roles in a second run and that the winning team would be the one that could locate the diamond with fewer questions when they were assigned the role of detectives. As for the questions,

the participants asked yes/no questions – such as "Is the diamond in the bedroom?" – until they found it. In addition, the group that played the role of a detective team was asked to point to the locations using a laser pointer. The second task was identical to the first one, except that the participants were told to find a key, instead of a diamond. The two activities lasted about 30 minutes each.

### Procedure

Each condition took about two weeks, and a total of two semesters were used to complete the two sets of data collections. On the first day, participants were asked to complete the background questionnaire and the oral production test as pretest. About a week later, participants completed the treatment session consisting of the two communicative activities during which the participants received recasts. The immediate posttest, which was the oral production test in a different version from the pretest, was administered a day after the treatment session.<sup>2</sup> This posttest was repeated a week after the treatment session as a delayed posttest.

### Verbal recasts

In this condition, the participants received oral recasts. When the participants did not use prepositions or used nontargetlike prepositions during the tasks, the researcher provided verbal recasts without gestures. To keep the consistency of recast quality, the researcher positioned her hands next to her body and tried not to emphasize any words as much as possible. Overall, a total of 38 recasts (an average of 9.50 [SD = 3.90] feedback moves per class of 15) were provided verbally.

### Gesture-enhanced recasts

This condition is identical to the previous condition except that gesture-enhanced recasts were used instead of verbal recasts. Specifically, during the information-gap activities, which were designed to elicit locative prepositions (*on*, *in*, *under*, *next to*, and *above*), the researcher used the same gestures that Nakatsukasa (2016) used (See Figure 1). Those gestures correspond to a native English speaker's gestures when describing the locations for *under*, *above*, *on*, and *next to* as shown in Tutton (2011). Overall, a total of 35 recasts (an average of 11.75 [SD = 1.59] feedback moves per class of 15) were provided verbally.

2. In order to measure learners' awareness, the author also administered stimulated recall. However, the learners' comments mostly concerned their active participation in activities and none were about linguistic targets or corrections. See Nakatsukasa (2016) for the detailed contents of the stimulated recall comments.

1. As in Nakatsukasa (2016), the present study also incorporated a grammar test. However, a majority of the students scored higher than 80% at the time of pretest and it did not allow further statistical analysis.

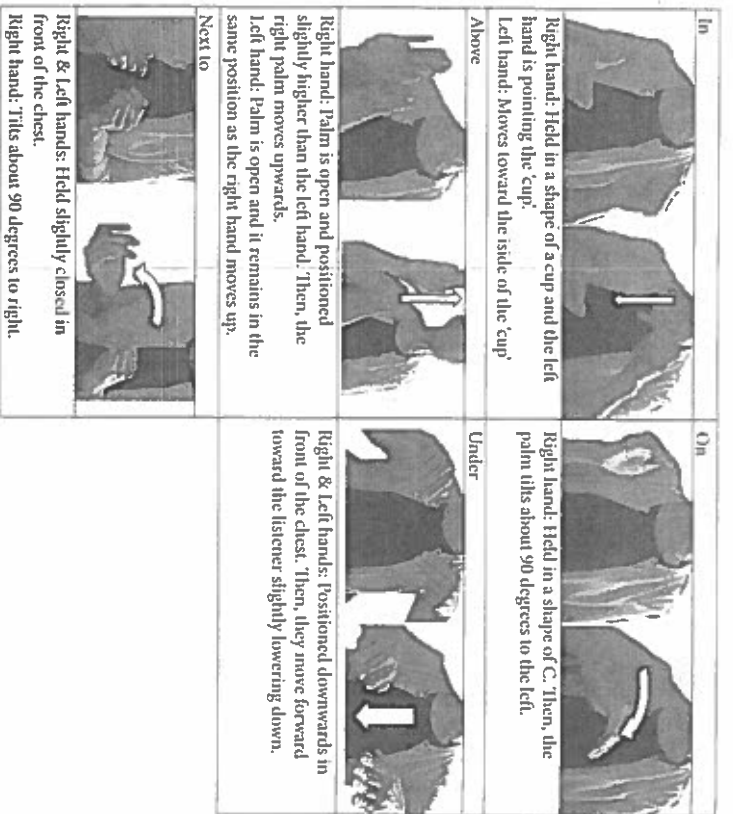


Figure 1. Illustrations of gestures used in preposition communicative tasks (Taken from Nakatsukasa (2016))

Note: "Right" and "Left" are from the speaker (gesturer)'s vantage point

Analysis of oral production tests

In order to account for the change in the learners' use of prepositions, I compared the scores obtained from the oral production tests across the pretest, immediate posttest, and delayed posttest using a repeated-measures ANOVA. In both conditions, the test scores were compared to identify if either males or females performed significantly differently from each other. Before calculating the level of significance, Mauchly's sphericity test was used in order to verify whether the assumption of sphericity was violated. Sphericity indicates that the variances of all the conditions are equal and a repeated-measures ANOVA is susceptible if the assumption of sphericity is violated. In cases when this assumption was violated, Greenhouse-Geisser adjusted scores were used for further analysis in order to minimize the effects of violation in the assumption of sphericity. In addition, the effect size was

calculated by  $\eta^2$ . I interpreted the effect size following the guideline by Cohen (1998): When  $\eta^2$  was around .01, it was interpreted as a small effect size, whereas .06 was interpreted as a medium effect size, and .14 and above was interpreted as a large effect size. Calculating the effect size was thought to be especially important for this study, as the sample size for verbal recasts was particularly small. Finally, one-way ANOVAs were used to understand at what point the males and females performed differently, such as at the time of the immediate or delayed posttest. When significant, the effect size was calculated using Cohen's  $d$ . Following Cohen (1998), the effect size was interpreted as follows: .20 to .30 as a small effect, around .50 as a medium effect, and above .80 as a large effect.

## Results

### Males vs. females in verbal recasts

As for the learners in the verbal recasts condition, first, I compared learners' oral production test scores using ratios between males and females. The descriptive statistics showed that the pretest scores were 51.10 ( $SD = 21.00$ ) for females and 56.00 ( $SD = 18.10$ ) for males. In the immediate posttest, females scored 71.70 ( $SD = 20.10$ ), and males scored 86.50 ( $SD = 11.10$ ). In the delayed posttest, females scored 76.30 ( $SD = 21.00$ ), and males scored 64.60 ( $SD = 19.70$ ). The visual representation of test scores is shown in Figure 2.

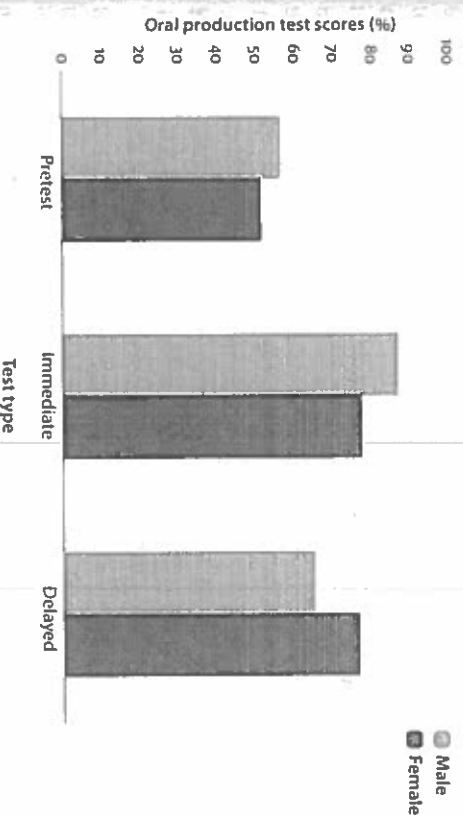


Figure 2. Oral production test scores from verbal recasts condition

According to Mauchly's test of sphericity, the sphericity of the data was confirmed ( $p < .01$ ). A repeated-measures ANOVA showed that there was a significant time effect  $F(2, 18) = 9.12, p < .01$ , a nonsignificant time and gender interaction effect  $F(2, 18) = 2.43, p = .12$ , and a nonsignificant gender effect  $F(1, 9) = 235.10, p = .77$ . Effect sizes, however, were large for time and the interaction between time and gender, but small for gender:  $\eta^2 = .50, \eta^2 = .21$ , and  $\eta^2 = .01$ , respectively.

Although the significant level was not reached in the repeated-measures ANOVA, it is also crucial to administer a post-hoc analysis in order to investigate whether or not there were any differences in the two groups' gain scores. The descriptive statistics showed that the average gain score from the pretest to immediate posttest for males was  $-7.70$  ( $SD = 17.36$ ) and  $25.56$  ( $SD = 12.24$ ) for females. A one-way ANOVA showed no significant differences in these gain scores,  $F(1, 11) = 3.90, p = .07$  ( $d = 1.56$ ), although the effect size was large. The gain score from the immediate to delayed posttest for males was  $-5.10$  ( $SD = 39.06$ ) and  $-6.10$  ( $SD = 29.06$ ) for females. A one-way ANOVA showed that there was again no significant difference in these gain scores  $F(1, 11) = .00, p = .96$  ( $d = .03$ ). In short, the results showed no significant differences in either gain scores for learners who received verbal recasts.

#### Males vs. females in gesture-enhanced recasts

In the second condition, I compared the males and females who received the gesture-enhanced recasts. The descriptive statistics showed that the pretest scores were  $61.30$  ( $SD = 16.20$ ) for females and  $56.30$  ( $SD = 11.90$ ) for males. In the immediate posttest, females scored  $79.80$  ( $SD = 14.80$ ), and males scored  $83.60$  ( $SD = 15.00$ ). In the delayed posttest, females scored  $85.20$  ( $SD = 10.10$ ), and males scored  $86.60$  ( $SD = 6.57$ ). The visual representation of test scores is shown in Figure 3.

As for the oral production test, Mauchly's Test of Sphericity did not confirm the sphericity of the data ( $p = .97$ ); therefore, Greenhouse-Geisser adjusted scores were used. A repeated-measures ANOVA showed that there was a significant time effect  $F(1.99, 33.89) = 38.30, p < .001$ , nonsignificant time and gender interaction  $F(0.96, 33.89) = 0.96, p = .39$ , and nonsignificant gender effect  $F(1, 33) = .00, p = .99$ . Effect sizes, however, were large for time but small for time and gender interaction and for a gender effect:  $\eta^2 = .69, \eta^2 = .05$ , and  $\eta^2 = .00$ , respectively.

Again, the post-hoc analysis involved the gain scores from pretest to immediate posttest and from immediate posttest to delayed posttest. As for the comparison between pretest to immediate posttest, males' mean gain score was  $17.36$  ( $SD = 30.79$ ) and females' gain score was  $8.38$  ( $SD = 28.58$ ). A one-way ANOVA confirmed that these differences were not significant  $F(1, 29) = .68, p = .42$  ( $d = .302$ ). As for the comparison between immediate to delayed posttest, males' gain score was  $-27.91$

( $SD = 41.54$ ) and females' gain score was  $-.14$  ( $SD = 19.83$ ). A one-way ANOVA confirmed that the difference was significant with a strong effect size  $F(1, 29) = 4.96, p = .03$  ( $d = .85$ ).

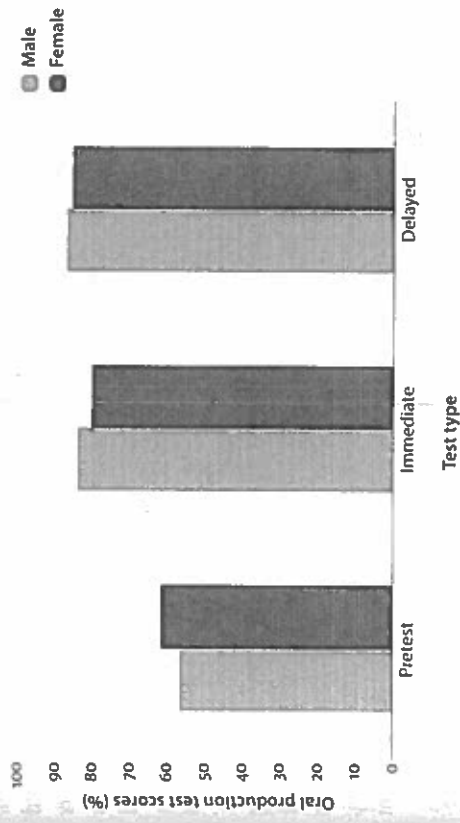


Figure 3. Oral production test scores from gesture-enhanced recasts condition

#### Summary of results

Overall, the results showed that there were no significant differences in the overall test scores and in the gain scores between males and females following verbal recasts. As for gesture-enhanced recasts, there was also no significant difference in the overall results, but there was a significant difference in the gain score from the immediate to delayed posttest indicating that the females significantly maintained their learning better than males.

#### Discussion and conclusion

The present study examined whether or not the effectiveness of L2 instruction, notably recasts, varies depending on learners' gender. The results have indicated that the effectiveness may vary depending on how recasts were provided. Specifically, when recasts were provided only verbally, both males and females exhibited similar patterns in L2 development. However, when recasts were provided with gestures, females were better able to retain their learning long-term than males.

As for the first finding that showed no significant difference between males and females in either pretest to immediate posttest or immediate posttest to delayed posttest following verbal recasts, it contradicts the studies which reported females' better L2 performance in general (e.g., Matsuda & Goebel, 2004; Payne & Lynn, 2011; Wuchterer & Reiterer, 2016) and a series of studies which collectively reported the females' advantage in short-term memory (Blecker, Bolla-Wilson, Agnew, & Meyers, 1988; Jensen & Reynolds, 1983; Kail & Siegel, 1978; Kramer, Delis, Kaplan, O'Donnell, & Prifitera, 1997, as cited in Kausanskaya, Marian, & Yoo, 2011) and long-term memory (e.g., Hartshorne & Ullman, 2006; Ullman, Miranda, & Travers, 2007). The first difference may be attributed to the differences in the design of the present and prior studies. Specifically, the previous studies used the L2 learners' performance at one point, whereas this study examined L2 development resulting from a brief classroom instruction. L2 learning is a highly complex system because the entire process is affected by various cognitive and psychological factors such as motivation, working memory, aptitude, and personality, to name a few (Dörnyei, 2014). Thus, it is important not to ignore other factors, which affect L2 development when relating the findings of empirical studies to long-term L2 development.

The second difference in these results can be explained by the relationship between the implicit nature of recasts and L2 processing. The verbal recasts used in this study were rather implicit in nature because no other cues, such as intonation or visual aids that highlighted the linguistic targets, were included. This, presumably, may have resulted in procedural processing, which is an implicit form of processing, upon which males tend to rely better compared to females. In addition, this finding did not support the females' advantages in their declarative memory system, which is part of long-term memory and plays an integral role in explicit learning. This contradiction may again be explained by the characteristics of the implicit nature of verbal recasts. Receiving the implicit form of recasts did not seem to have resulted in explicit learning, and thus, resulted in an almost equal L2 developmental pattern between males and females.

The second analysis revealed that when receiving verbal recasts with gestures, females retained their learning better than males from the immediate to delayed posttest, even though there was no significant difference across gender from pretest to immediate posttest. This coincides with the existing studies in L2 performance in which females outperformed males (e.g., Matsuda & Goebel, 2004; Payne & Lynn, 2011; Wuchterer & Reiterer, 2016). Taking the results from the verbal recast condition, that there were no significant differences between males and females following verbal recasts, it is important to explore the possible impact of gender-specific cognitive systems to account for this finding. First, according to the Dual Coding Theory, incorporating verbal and nonverbal modalities such as gestures reinforces long-term learning – thus, it is possible that gesture-enhanced recasts reinforced long-term learning,

which is closely related to the active use of long-term memory (Clark & Paivio, 1991). Meanwhile, the existing studies have reported females' advantage with respect to long-term memory (e.g., Hartshorne & Ullman, 2006; Ullman, Miranda, & Travers, 2007). Combining these two notions, it is possible that females were better able to take advantage of gesture-enhanced recasts due to their superior long-term memory compared to males. Secondly, in comparison to the verbal recasts, gesture-enhanced recasts are possibly more explicit because the gestures depicted the erroneous uses of prepositions. This, arguably, may have led to the active use of declarative or explicit learning, in which females are reported to have advantages (e.g., Hartshorne & Ullman, 2006; Ullman, Miranda, & Travers, 2007). These two explanations, which account for the difference in the cognitive systems between males and females, could provide reasons why the females in this study were better able to retain their learning long-term following gesture-enhanced recasts compared to males.

#### Limitations and future directions

As is the case for many studies, this study is not without limitations. First and foremost, it must be noted that the number of participants in the verbal recast condition became extremely limited because of logistic reasons. To better take advantage of the inferential statistics, it would have been ideal to include a larger number of participants.

Secondly, even though this study illustrated gender-based differences in L2 learning, it must be mentioned that there are various intervening variables related to gender such as learners' task preferences and topic interests as briefly mentioned in the Discussion section. For example, Lee and Pulido (2017) conducted a study on L2 vocabulary acquisition through reading with Korean English-as-a-foreign-language students. Their results indicated that there was a significant interaction between topic interest and gender on vocabulary acquisition. If so, it is possible that task preference can be influenced by gender. This study employed only one type of a communicative task. Additional studies with various types and topics of common communicative tasks are needed to further understand how gender may play a role in daily classroom learning.

Thirdly, although the long-term benefits of gesture-enhanced recasts were found for the females and the author speculated that this result is due to the gender-specific cognitive system such as long-term memory, it is impossible to single out other factors that may have influenced the results of the gesture-enhanced recast condition. For example, a few studies in a field of developmental psychology have shown that the presence of speakers' gestures promoted children's attention by examining their eye gaze (Deak, Flom, & Pick, 2000; Goldin-Meadow, 2004).



Therefore, in order to fully understand the phenomenon of facilitative roles of pedagogical gestures for females and to truly argue that if effectiveness of exposure to gesture depends on gender type, future studies need to examine the level of learners' attention using an eye-tracker. It will also be informative to assess the possible interaction between learners' level of attention and their long-term learning.

Fourth, a few methodological issues need to be addressed. To begin with, this study did not control for the instructor's gender. The female researcher served as an instructor for all the data collection, therefore, it is impossible to eliminate the possibility that the results might be different if the study is replicated with a male instructor. It would have been informative to have a male instructor to counter-balance the instructor's gender, as well. Next, it is also important to note that the researcher visited each classroom only for the purpose of data collection. Thus, the results could follow a different pattern if the study is to be replicated with an instructor who has been assigned to teach the course over a semester, as educational studies have shown that an instructor seems to interact with males and females differently. Jones and Dindia (2004), for example, conducted a meta-analysis of 32 studies about the interaction between gender and instruction and their results show that both male and female instructors tend to interact with male students more frequently and more negatively (i.e., blaming) than with female students. Indeed, the accumulation of those gender-specific interactions might naturally influence L2 learning, as well as learners' gender-specific cognitive abilities. Thus, to fully understand gender influence on L2 learning longitudinally, it is also important to have a non-researcher as an instructor and observe how the instructor interacts with males and females in a daily classroom.

#### Pedagogical implications

As the field of SLA has matured and researchers have identified various individual differences that impact L2 learning, it is important to take a further step to raise awareness for language educators to understand *how* some learners tend to benefit more from corrective feedback over others. In this study, I presented how a gesture-enhanced recast, an explicit type of recast, can potentially be more beneficial for females than males. This, however, does not mean that I suggest avoid using gestures while providing verbal corrective feedback. As Nakatsukasa (2016) showed, the gesture-enhanced recasts were more effective than verbal recasts in general. What I suggest, instead, is that language instructors carefully observe what learners learn from corrective feedback, and tailor the explicitness of feedback depending on learners' individual differences, including their gender.

With all the limitations mentioned above, I hope this study serves as one of the initial steps into examining the effectiveness of recasts between males and females, and contributes to the increasing research investigating relationships between learners' individual differences and corrective feedback. This study showed that males and females developed their L2 similarly following oral recasts; however, when the gesture-enhanced recasts were used, females significantly outperformed males in terms of long-term retention of learning. A growing number of studies have addressed how individual differences affect the effectiveness of corrective feedback; however, the empirical studies which investigated the impact of gender on L2 development in a classroom setting is still fairly limited. I hope that this study contributed to the existing studies on recasts to further understand the relationship between learners' individual differences and the effectiveness of recasts.

#### Acknowledgements

The data presented in this study were collected as a part of my doctoral dissertation. I am truly grateful for the *Language Learning Dissertation Grant* that funded this project.

#### References

- Allen, L. Q. (1995). The effects of emblematic gestures on the development and access of mental representations of French expressions. *Modern Language Journal*, 79(4), 521–529. doi:10.1111/j.1540-4781.1995.tb05454.x
- Ammar, A., & Spada, N. (2006). One size fits all?: Recasts, prompts, and L2 learning. *Studies in Second Language Acquisition*, 28(4), 543–574. doi:10.1017/S0272263106060268
- Bleecker, M. L., Bolla-Wilson, K., Agnew, J., & Meyers, D. A. (1988). Age-related sex differences in verbal memory. *Journal of Clinical Psychology*, 44(3), 403–411. doi:10.1002/1097-4679(198805)44:3<403::AID-JCLP2270440315>3.0.CO;2-o
- Bowden, H. W., Sanz, C., & Stafford, C. A. (2005). Individual differences: Age, sex, working memory, and prior knowledge. In C. Sanz (Ed.), *Mind and context in adult second language acquisition: Methods, theory, and practice* (pp. 105–140). Washington, DC: Georgetown University Press.
- Celce-Murcia, M., Larsen-Freeman, D., & Williams, H. A. (1983). *The grammar book: An ESL/EFL teacher's course* (2nd ed). Boston, MA: Heinle & Heinle Publishers.
- Clark, J. M., & Paivio, A. (1991). Dual coding theory and education. *Educational Psychology Review*, 3(3), 149–170. doi:10.1007/BF01320076
- Cohen, J. (1998). *Statistical power analysis for the behavioral sciences*. Mahwah, NJ: Lawrence Erlbaum Associates.
- 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. doi:10.1073/pnas.081884106

- Davies, M. (2006). Paralinguistic focus on form. *TESOL Quarterly*, 40(4), 841–855. doi:10.2307/40264316
- Deák, G. O., Plonk, R. A., & Pick, A. D. (2000). Effects of gesture and target on L2- and 18-month-olds' joint visual attention to objects in front of or behind them. *Developmental Psychology*, 36(4), 511–523. doi:10.1037/0012-1649.36.4.511
- Dönyei, Z. (2014). Researching complex dynamic systems: 'Retrodictive qualitative modelling' in the language classroom. *Language Teaching*, 47(01), 80–91.
- Doughy, C. (1994). Finetuning of feedback by competent speakers to language learners. In J. E. Alatis (Ed.), *Georgetown University Round Table on Languages and Linguistics: Educational Linguistics, Crosscultural Communication, and Global Interdependence* (pp. 96–108). Washington, DC: Georgetown University Press.
- Ellis, R. (2003). *Task-based language learning and teaching*. Oxford: Oxford University Press.
- Ellis, R., & Sheen, Y. H. (2006). Reexamining the role of recasts in second language acquisition. *Studies in Second Language Acquisition*, 28(4), 575–600. doi:10.1017/S027226310606027X
- Faraco, M., & Kida, T. (2008). Gesture and the negotiation of meaning in a second language classroom. In S. McCafferty & G. Stann (Eds.), *Gesture: Second language acquisition and classroom research* (pp. 280–297). London: Routledge.
- Goldin-Meadow, S. (2004). Gesture's role in the learning process. *Theory into Practice*, 43(4), 314–321. doi:10.1207/s15430421tp4304\_10
- Go, J. (2012). Corrective feedback and working memory capacity in interaction-driven L2 learning. *Studies in Second Language Acquisition*, 34(03), 445–474. doi:10.1017/S0272263112000149
- Green, J., & Oxford, R. (1995). A closer look at learner strategies, L2 proficiency, and gender. *TESOL Quarterly*, 29(2), 36–72. doi:10.2307/3587655
- Han, Z. (2002). A study of the impact of recasts on tense consistency in L2 output. *Teachers of English to Speakers of Other Languages*, 36(4), 543–572.
- Hartshorne, J. K., & Ullman, M. T. (2006). Why girls say "holded" more than boys. *Developmental Science*, 9(1), 21–32. doi:10.1111/j.1467-7687.2005.00459.x
- Havranek, G. (1999). The effectiveness of corrective feedback: Preliminary results of an empirical study. *Acquisition et Interaction en Langue Étrangère*, 2, 189–206.
- Hudson, N. (2011). *Teacher gesture in a post-secondary English as a second language classroom: A sociocultural approach*. (Unpublished doctoral dissertation). University of Nevada, Las Vegas.
- Jarvis, S., & Odlin, T. (2000). Morphological type, spatial reference, and language transfer. *Studies in Second Language Acquisition*, 22(4), 535–556. doi:10.1017/S0272263100004034
- Jensen, A., & Reynolds, C. (1983). Sex differences on the WISC-R. *Personality and Individual Differences*, 4, 223–226. doi:10.1016/0191-8869(83)90029-6
- Jones, S. M., & Dindia, K. (2004). A meta-analytic perspective on sex equity in the classroom. *Review of Educational Research*, 74(4), 443–471. doi:10.3102/00346543074004443
- Kail, R. V., & Siegel, A. W. (1978). Sex and hemisphere differences in the recall of verbal and spatial information. *Cortex: A Journal Devoted to the Study of the Nervous System and Behavior*, 14(4), 557–563. doi:10.1016/S0010-9453(78)80030-6
- Kanmura, N. (2015). The effectiveness of intensive and extensive recasts on L2 acquisition for implicit and explicit knowledge. *Linguistics and Education*, 29, 59–72. doi:10.1016/j.linged.2014.11.002
- Kao, R. K. (2001). Where have the prepositions gone? A study of English prepositional verbs and input enhancement in instructed SLA. *International Review of Applied Linguistics in Language Teaching (IRAL)*, 39(3), 195–216. doi:10.1515/iral.2001.002
- Kausanskaya, M., Nartian, V., & You, J. (2011). Gender differences in adult word learning. *Acta Psychol (Amst)*, 137(1), 24–35. doi:10.1016/j.actpsy.2011.02.002
- Kelly, S. D., & Lee, A. L. (2012). When actions speak too much louder than words: Hand gestures disrupt word learning when phonetic demands are high. *Language and Cognitive Processes*, 27, 793–807. doi:10.1080/01690965.2011.581252
- Kelly, S. D., McDevitt, T., & Isch, M. (2009). Brief training with co-speech gesture lends a hand to word learning in a foreign language. *Language and Cognitive Processes*, 24(2), 313–334. doi:10.1080/0169096080365567
- Kissau, S. (2006). Gender differences in motivation to learn French. *Canadian Modern Language Review/La Revue Canadienne Des Langues Vivantes*, 62(3), 401–422. doi:10.3138/cmlr.62-3-401
- Kissau, S., Kolano, L. Q., & Wang, C. (2010). Perceptions of gender differences in high school students' motivation to learn Spanish. *Foreign Language Annals*, 43(4), 703–721. doi:10.1111/j.1944-9720.2010.01101.x
- Kissau, S., & Turnbull, M. (2008). Boys and French as a second language: A research agenda for greater understanding. *Canadian Journal of Applied Linguistics*, 11, 151–170.
- Koul, R., Roy, L., Kewckuekol, S., & Ploisawasachia, S. (2009). Multiple goal orientations and foreign language anxiety. *System*, 37(4), 676–688. doi:10.1016/j.system.2009.09.011
- Kramer, J. H., Delis, D. C., Kaplan, E., O'Donnell, L., & Prifitera, A. (1997). Developmental sex differences in verbal learning. *Neuropsychology*, 11(4), 577–584. doi:10.1037/0894-4105.11.4.577
- Lee, S., & Pulido, D. (2017). The impact of topic interest, L2 proficiency, and gender on EFL incidental vocabulary acquisition through reading. *Language Teaching Research*, 20(1), 118–135. doi:10.1177/1362168816637381
- Lechman, J. (2003). Recasts and second language development. *Studies in Second Language Acquisition*, 25(1), 37–63. doi:10.1017/S0272263103000020
- Li, S. (2010). The effectiveness of corrective feedback in SLA: A meta-analysis. *Language Learning*, 60(2), 309–365. doi:10.1111/j.1467-9922.2010.00961.x
- Lindstromberg, S. (2010). *English prepositions explained*. Amsterdam: John Benjamins. doi:10.1017/978-1-107-24512-57
- Lim, T. K. (1994). Gender-related differences in intelligence: Application of confirmatory factor analysis. *Intelligence*, 19(2), 179–192. doi:10.1016/0160-2896(94)90012-4
- Loewen, S., & Nabei, T. (2007). Measuring the effectiveness of oral corrective feedback on L2 knowledge. In A. Mackey (Ed.), *Conversational interaction in second language acquisition* (pp. 339–360). Oxford: Oxford University Press.
- Loewen, S., & Philp, J. (2006). Recasts in the adult English L2 classroom: Characteristics, explicitness, and effectiveness. *Modern Language Journal*, 90, 536–556. doi:10.1111/j.1540-4781.2006.00465.x
- Long, M. H. (1996). The role of the linguistic environment in second language acquisition. In W. C. Ritchie & T. K. Bhatia (Eds.), *Handbook of research on second language acquisition* (Vol. 2, pp. 413–468). New York, NY: Academic Press.
- Long, M. H. (2007). *Problems in SLA*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Long, M. H., Inagaki, S., & Ortega, L. (1998). The role of implicit negative feedback in SLA: Models and recasts in Japanese and Spanish. *Modern Language Journal*, 82(3), 357–371. doi:10.1111/j.1540-4781.1998.tb02133.x
- 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. doi:10.1111/j.1467-1770.2001.tb00019.x
- Lyster, R. (2004). Differential effects of prompts and recasts in form-focused instruction. *Studies in Second Language Acquisition*, 26, 399–432. doi:10.1017/S0272263104263021

- Lyster, R., & Ranta, L. (1997). Corrective feedback and learner uptake. *Studies in Second Language Acquisition*, 19(1), 37–66. doi:10.1017/S0272263197001034
- Macedonia, M., & Klimesch, W. (2014). Long-term effects of gestures on memory for foreign language words trained in the classroom. *Mind, Brain, and Education*, 8(2), 74–88. doi:10.1111/mbe.12047
- Mackey, A., Gass, S., & McDonough, K. (2000). How do learners perceive interactional feedback? *Studies in Second Language Acquisition*, 22, 471–497. doi:10.1017/S0272263100004022
- 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. 407–453). Oxford: Oxford University Press.
- Mackey, A., & Philp, J. (1998). Conversational interaction and second language development: Recasts, responses, and red herrings? *Modern Language Journal*, 82(3), 338–356. doi:10.1111/j.1540-4781.1998.tb01211.x
- Matsuda, S., & Gobel, P. (2004). Anxiety and predictors of performance in the foreign language classroom. *System*, 32, 21–36. doi:10.1016/j.system.2003.08.002
- Nakatsukasa, K. (2016). Efficacy of recasts and gestures on the acquisition of locative prepositions. *Studies in Second Language Acquisition*, 38, 771–799. doi:10.1017/S0272263115000467
- Nassaji, H. (2009). Effects of recasts and elicitations in dyadic interaction and the role of feedback explicitness. *Language Learning*, 59, 411–452. doi:10.1111/j.1467-9922.2009.00511.x
- Nicholas, H., Lightbown, P. M., & Spada, N. (2001). Recasts as feedback to language learners. *Language Learning*, 51(4), 719–758. doi:10.1111/0023-8333.00172
- Oliver, R. (1995). Negative feedback in child NS-NNS conversation. *Studies in Second Language Acquisition*, 17, 459–481. doi:10.1017/S0272263100014418
- Oxford, R., Nyikos, M., & Ehrman, M. (1988). Vive la différence? Reflections on sex differences in use of language learning strategies. *Foreign Language Annals*, 21(4), 321–329. doi:10.1111/j.1944-9720.1988.tb01076.x
- Pajares, F. (2003). Self-efficacy beliefs, motivation, and achievement in writing: A review of the literature. *Reading and Writing Quarterly*, 19, 139–159. doi:10.1080/10573560308222
- Pajares, F., & Valiante, G. (2001). Gender differences in writing motivation and achievement of middle school students: A function of gender orientation? *Contemporary Educational Psychology*, 26(3), 366–381. doi:10.1006/ceps.2000.1069
- Pajares, F., Valiante, G., & Chicoung, Y. F. (2007). Writing self-efficacy and its relation to gender, writing motivation, and writing competence: A developmental perspective. In S. Hidi & P. Boscolo (Eds.), *Writing and motivation* (pp. 141–159). Bingley: Emerald.
- Payne, T. W., & Lynn, R. (2011). Sex differences in second language comprehension. *Personality and Individual Differences*, 50(3), 434–436. doi:10.1016/j.paid.2010.10.026
- 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.
- Reiterer, S. M., Hu, X., Erb, M., Rota, G., Nardo, D., Grould, W., Winkler, S., & Ackermann, H. (2011). Individual differences in audio-vocal speech imitation aptitude in late bilinguals: Functional neuro-imaging and brain morphology. *Frontiers in Psychology*, 2, 1–12. doi:10.3389/fpsyg.2011.00271
- Robinson, P. (2010). Aptitude and second language acquisition. *Annual Review of Applied Linguistics*, 25(3), 46–73.
- Sheen, Y. H. (2004). Corrective feedback and learner uptake in communicative classrooms across instructional settings. *Language Teaching Research*, 8, 263–300. doi:10.1191/1362168804lr1460a
- Sheen, Y. (2008). Recasts, language anxiety, modified output, and L2 learning. *Language Learning*, 58(4), 835–874. doi:10.1111/j.1467-9922.2008.00480.x
- Sime, D. (2006). What do learners make of teachers' gestures in the language classroom? *International Review of Applied Linguistics in Language Teaching (IRAL)*, 44, 211–230.
- 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. doi:10.1017/CBO9781139524575.009
- Smotrova, T. (2015). Making pronunciation visible: Gesture in teaching pronunciation. *TESOL Quarterly*, 51(1), 59–89. doi:10.1002/tesq.276
- Smotrova, T., & Lantolf, J. P. (2013). The function of gesture in lexically focused L2 instructional conversations. *Modern Language Journal*, 97(2), 397–416. doi:10.1111/j.1540-4781.2013.12008.x
- Sueyoshi, A., & Hardison, D. M. (2005). The role of gestures and facial cues in second language listening comprehension. *Language Learning*, 55(4), 661–699. doi:10.1111/0023-8333.2005.00320.x
- Tellier, M. (2008). The effect of gestures on second language memorisation by young children. *Gesture*, 8, 219–235. doi:10.1075/gest.8.2.061el
- Tellier, M. (2010). Manuscrit auteur, publié dans "L'enseignement des langues vivantes étrangères à l'école. In C. Corblin & J. Sauvage (Eds.), *L'apprentissage et l'enseignement des langues vivantes à l'école. Impacts sur le développement de la langue maternelle* (pp. 31–54). Paris: L'Harmattan.
- Tutton, M. (2011). How speakers gesture when encoding location with English on and French sur. *Journal of Pragmatics*, 43(14), 3431–3454. doi:10.1016/j.pragma.2011.07.012
- Ullman, M. T., Miranda, R. A., & Travers, M. L. (2007). Sex differences in the neurocognition of language. In J. B. Becker, K. J. Berkley, N. Geary, E. Hampson, & J. Herman (Eds.), *Sex differences in the brain: From genes to behavior* (pp. 291–310). Oxford: Oxford University Press. doi:10.1093/acprofoso/9780195311587.003.0015
- Wang, W., & Loewen, S. (2015). Nonverbal behavior and corrective feedback in nine ESL university-level classrooms. *Language Teaching Research*, 20(4), 459–478. doi:10.1177/1362168815572739
- Wucherer, B. V., & Reiterer, S. M. (2016). Language is a girly thing, isn't it? A psycholinguistic exploration of the L2 gender gap. *International Journal of Bilingual Education and Bilingualism*, 1, 1–17. doi:10.1080/13670050.2016.1142499
- 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(10), 235–263. doi:10.1017/S0272263109990519
- 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. doi:10.1093/applin/ams044

