AAUSC 2013 Volume – Issues in Language Program Direction

Individual Differences, L2 Development, and Language Program Administration: From Theory to Application

Cristina Sanz, Georgetown University
Beatriz Lado, Lehman College (CUNY)
Editors

Stacey Katz Bourns, Harvard University
Series Editor

CENGAGE Learning®

Australia • Brazil • Japan • Korea • Mexico • Singapore • Spain • United Kingdom • United States

AAUSC 2013 Volume - Issues in Language Program Direction: Individual Differences, L2 Development, and Language Program Administration: From Theory to Application
Cristina Sanz, Beatriz Lado, and Stacey Katz Bourns

Product Director: Beth Kramer

Product Assistant: Daniel Cruse

Associate Media Developer: Patrick Brand

Executive Market Development Manager: Ben Rivera

Rights Acquisitions Specialist: Jessica Elias

Manufacturing Planner: Betsy Donaghey

Art and Design Direction, Production Management, and Composition: PreMediaGlobal

Library of Congress Control Number: 2013948412

ISBN-13: 978-1-285-76058-2

ISBN-10: 1-285-76058-1

Cengage Learning
200 First Stamford Place, 4th Floor
Stamford, CT 06902
USA

Cengage Learning is a leading provider of customized learning solutions with office locations around the globe, including Singapore, the United Kingdom, Australia, Mexico, Brazil, and Japan. Locate your local office at **www.cengage.com/global**.

Cengage Learning products are represented in Canada by Nelson Education, Ltd.

To learn more about Cengage Learning Solutions, visit **www.cengage.com**

Purchase any of our products at your local college store or at our preferred online store **www.cengagebrain.com**

Printed in the United States of America
1 2 3 4 5 6 7 17 16 15 14 13

Chapter 4

Computer-delivered Feedback and L2 Development: The Role of Explicitness and Working Memory

Nuria Sagarra, Rutgers University

Rebekha Abbuhl, California State University, Long Beach

Researchers, instructors, and program administrators have long struggled with the question of how to improve second language (L2) learners' grammatical accuracy. Researchers have zeroed in on corrective feedback as a potential mechanism for helping learners recognize gaps in their knowledge of the target language, and recent years have witnessed a surge of studies on the relative effectiveness of different forms of feedback. Within this body of studies, recasts, or more targetlike reformulations of learners' initial non-targetlike utterances, have received particular attention, with some researchers arguing that this form of feedback is an ideal method of drawing learners' attention to discrepancies between interlanguage productions and targetlike norms (e.g., Long, 2007).

Broadly speaking, recasts are believed to accrue acquisitional benefits because they provide an unobtrusive signal to the learner that her or his utterance was not well formed (negative evidence), and, at the same time, provide a model or reformulation of that utterance (positive evidence). The close proximity of the original non-targetlike utterance and the targetlike reformulation may help the learner notice the discrepancy between the two, especially since the learner's meaning is held constant (Doughty, 2001; Doughty & Varela, 1998; Leeman, 2003; Long, 1996, 2007; Long & Robinson, 1998). A number of empirical studies have provided evidence to support these claims (e.g., Doughty & Varela, 1998; Han, 2002; Havranek, 2002; Iwashita, 2003; Leeman, 2003; Long, Inagaki & Ortega, 1998; Mackey & Oliver, 2002; Mackey & Philp, 1998; Oliver & Mackey, 2003), and recent meta-analyses have also supported the role of recasts in L2 development. For example, Mackey and Goo (2007), in their meta-analysis of classroom- and laboratory-based studies on various forms of corrective feedback, reported larger effect sizes for recasts than for either negotiation or metalinguistic feedback.

In meaning-focused classrooms, however, learners (and in particular lower proficiency learners) may misconstrue recasts as conversation continuations, thereby missing the corrective intent of this discourse move (Lyster, 1998, 2004; Lyster & Ranta, 1997; Panova & Lyster, 2002; see, however, Goo & Mackey, 2013). If the effectiveness of recasts is contingent on learners recognizing this corrective intent (as a number of researchers have suggested, e.g., Carroll, 2001; Ellis, Loewen, & Erlam, 2006), this tendency may limit the developmental effectiveness of recasts (Egi, 2007). A recent meta-analysis focusing solely on studies of corrective feedback in the classroom (Lyster & Saito, 2010) provides some evidence to support this claim, as the researchers found smaller effect sizes for recasts than

for prompts (discourse moves that prompt learners to self-correct, such as clarification requests and repetition). Recasts targeting certain linguistic targets may also lack saliency. For example, a number of studies (Al-Surmi, 2012; Carpenter, Jeon, MacGregor, & Mackey, 2006; Gass & Lewis, 2007; Kim & Han, 2007; Mackey, Gass, & McDonough, 2000; Sheen, 2006) indicate that L2 learners find recasts targeting morphosyntactic errors more difficult to notice than recasts targeting phonology and lexis.

To make the corrective nature of recasts more overt, researchers have suggested that recasts be presented in a computer context (e.g., Yilmaz & Yuksel, 2011), provided in the written mode (e.g., Sagarra & Abbuhl, 2013), and/or enhanced (e.g., through the use of prosodic highlighting or typographical markings such as bolding or underlining). However, the effectiveness of enhanced recasts seems to be linked to input modality: enhanced oral recasts are more effective than unenhanced oral recasts (e.g., Leeman, 2003; Nassaji, 2009), but enhanced written recasts are as effective as unenhanced written recasts (e.g., Sagarra & Abbuhl, 2013; Sachs & Suh, 2007). Because written enhancements do not appear to be effective, in this study we explore an alternative way to make written recasts more salient: by combining them with a preemptive metalinguistic explanation. To compare this feedback with Sagarra and Abbuhl's (2013) unenhanced written recasts, enhanced written recasts, and no feedback groups, we used the same type of participants (beginning L1 English learners of L2 Spanish enrolled in the same course) and identical pretest, treatment, and post test materials. In accordance with Sagarra and Abbuhl, we examine a possible correlation between working memory (WM) capacity and learners' ability to benefit from recasts, following studies indicating that higher WM span learners are better equipped to benefit from recasts than lower WM span learners (Sagarra & Abbuhl, 2013; Mackey, Philp, Egi, Fujii, & Tatsumi, 2002; Mackey, Adams, Stafford, & Winke, 2010; Mackey & Sachs, 2011; Révész, 2012).

Literature Review

This literature review is divided into two sections. The first section presents studies on three ways of heightening the saliency of recasts as feedback on grammatical errors: using written instead of oral recasts, enhancing recasts, and combining them with proactive metalinguistic explanations. The second section summarizes the literature on the relationship between WM and recasts.

Making Recasts More Salient

A number of studies have investigated ways of making the corrective intent of recasts more apparent to learners. One suggestion has been to provide written recasts instead of oral recasts. As noted by a number of researchers, the relative permanence of the written word (e.g., in text-based chats) means that learners can read and reread the input. Furthermore, without the time pressure present in oral face-to-face interaction, learners may have greater cognitive

resources available to process (and produce) longer stretches of language, focus on form, and notice instances of corrective feedback (Hummel & French, 2010; Payne & Ross, 2005; Payne & Whitney, 2002; Sauro, 2009; Warschauer, 1996; Yilmaz & Yuksel, 2011).

In terms of empirical research, however, findings are inconclusive. Although not directly comparing written and oral feedback types, a number of studies conducted in online chatrooms (which employ written feedback) have indirectly touched upon the issue of modality. For example, Yilmaz and Yuksel (2011) reported greater noticing of form in chatrooms than in oral face-to-face interactions. Others, however, have argued that certain aspects of the chatroom context may work against learners' attempts to focus on form, including the potential for off-task behavior in this medium (Loewen & Erlam, 2006), the common "split-turns" in text-based communication (where, e.g., an error might appear many lines from its reformulation) (e.g., Lai & Zhao, 2006; Lai, Fei, & Roots, 2008; Loewen & Erlam, 2006; Sauro, 2009), and individuals' tendency to prioritize meaning over form in this context (especially when the text-chat is not moderated, Loewen & Reissner, 2009). Loewen and Erlam (2006) cited these possible factors after finding no effects for either recasts or metalinguistic prompts (as measured by timed and untimed grammaticality judgment tests) in their chatroom-based study with beginning English language learners. In her comparison of recasts and metalinguistic feedback provided in the chatroom, Sauro (2009) similarly concluded that the corrective intent of recasts might not have been sufficiently apparent to learners in the chatroom, as measured by their responses on a grammaticality judgment test.

Studies *directly* comparing written and oral feedback are rare. In one classroom-based study to investigate this issue, Sheen (2010) reported that direct, written corrections were more effective than oral recasts (as measured by a speeded dictation test, a written narrative test, and an error correction test), that there were no differences between written and oral metalinguistic explanations, and that all feedback types were better than no feedback for intermediate L2 learners of English. Sheen speculated that students did not recognize the corrective intent of the oral recasts and that more explicit forms of feedback (direct written correction and metalinguistic feedback in any modality) may have been necessary for her students to make gains with the target structure. Sagarra and Abbuhl (2013) conducted a similar study comparing written and oral feedback in a computer-delivered context. They compared the effects of utterance rejections, recasts, and enhanced recasts delivered in the written or the oral mode for beginning L2 Spanish learners (L1 English) learning noun-adjective gender and number agreement. The findings revealed that all recasts (written and oral) yielded higher linguistic accuracy and more modified output on written and oral posttests up to two months after treatment than did utterance rejections, and utterance rejections were in turn superior to no feedback.

In addition to using the written mode instead of the oral mode to make the corrective intent of recasts more apparent, some scholars have proposed enhancing recasts. Recasts can be enhanced orally, via added stress and rising intonation, or typographically, marking inflections in bold, contrasting colors, or in capital

letters. Concerning orally enhanced recasts in the classroom, Leeman (2003) found that enhanced recasts yielded greater gains as compared to a control group. Most importantly, in the one-week delayed posttest, only the enhanced recast group significantly outperformed the control group on gender agreement scores. Similarly, Nassaji (2009) reported that orally enhanced recasts with and without prompts and explicit prompts led to greater post-interaction corrections among L2 English learners than did the more implicit forms of feedback, including unenhanced recasts.

Studies investigating enhanced written recasts, however, have not reported similar benefits. For example, Sachs and Suh (2007) found no differences between enhanced (with bolding and highlighting) and unenhanced written recasts in their study of English learners and backshifting of verbs in indirect reported speech. Similarly, Sagarra and Abbuhl (2013) did not obtain significant differences between typographically enhanced and unenhanced recasts, even though orally enhanced recasts did lead to more targetlike production and modified output than orally unenhanced recasts. In addition, orally enhanced recasts were more effective than typographically enhanced recasts on the oral interaction posttests. It may be speculated that typographical enhancements, while potentially drawing students' attention to the recast, are not sufficient in and of themselves for triggering the necessary cognitive comparison between the reformulation and the original utterance. An alternative explanation is that typographical enhancements may facilitate the development of receptive skills but not of productive skills. As Cho (2010) explains, "although input enhancement can draw learners' attention to form, it does not necessarily stimulate learners' further cognitive processing of that input, which may be necessary for the development of productive skills" (p. 82). Some support for this claim may be found in previous studies on visual input enhancement that have reported limited effects (Alanen, 1995; Cho, 2010) or no effects (Izumi, 2002; Kim, 2006; Leow, Egi, Nuevo, Tsai, 2003; Shook, 1999; Wong, 2003).

An alternative to both written presentation and enhancement is to combine recasts with negative evidence such as metalinguistic explanations. Metalinguistic explanations can be provided after the recast, as a double feedback move (reactively) or a priori (a form of "preemptive" negative evidence, Long & Robinson, 1998). In order to examine double feedback, Sheen (2007) compared the effects of recasts alone to recasts combined with metalinguistic explanation. The double-feedback group outperformed the recast and control groups on both the immediate and delayed posttests, with no significant difference between the recast and control groups. In one of the few studies to examine the effect of preemptive negative evidence, Lyster (2004) found that form-focused instruction (FFI) combined with prompts led to greater linguistic accuracy than FFI combined with recasts or with no feedback among L2 French learners.

Given the conflicting findings regarding feedback modality, enhancement, and the effect of combining recasts with preemptive negative evidence, more work is clearly needed in this area. This is particularly the case given the increasing evidence that recasts do not benefit all learners equally, but may be more effective for some learners than for others. One individual difference factor that has received particular attention in recent years is that of WM, to which we turn next.

Working Memory and the Learners' Ability to Benefit from Recasts

It has been noted that oral recasts place heavy demands upon a learner's WM capacity (the ability to simultaneously store and process incoming information during complex cognitive tasks) (Baddeley, 2007). According to Baddeley's domain-specific multiple-resource model, WM is comprised of a central executive and three slave systems with independent limited capacities: two devoted to the temporary maintenance of visual information (the visuospatial sketchpad) and verbal and acoustic information (the phonological loop), and one in charge of the flow of information between long-term memory and the other two slave systems. Maintenance of information in the phonological loop is thought to decay rapidly (generally, after about two seconds), unless the information is refreshed through a sub-vocal rehearsal process (as when you repeat a phone number to yourself in order to avoid forgetting it). Concerning recasts, in order to notice a "gap" between an initial non-targetlike utterance and a targetlike reformulation, the learner must hold onto both utterances long enough to compare the two and identify the location and nature of the discrepancy.

The question thus becomes how we can help learners with smaller WM capacities benefit from recasts. One suggestion has been to make recasts more salient and memorable to learners, for example, by employing shorter recasts with fewer changes and/or by using prosodic highlighting (e.g., Bao, Egi, & Han, 2011; Doughty, 1994; Egi, 2007; Kim & Han, 2007; Nassaji, 2009; Philp, 2003; Sheen, 2006). Along with the recommendation that teachers use more explicit forms of oral recasts (e.g., those employing prosodic highlighting), another suggestion for "leveling the playing field" for learners with diverse WM capacities is to employ *written* recasts. Researchers have suggested that the reduced cognitive burden typically associated with reading the written word may be particularly beneficial for learners with smaller WM capacities (e.g., Hummel & French, 2010).

Although a number of studies have provided evidence that learners with smaller WM capacities are less likely to experience linguistic gains and produce modified output after recasts than are learners with larger WM capacities (e.g., Sagarra & Abbuhl, 2013; Mackey et al., 2002, 2010; Mackey & Sachs, 2011; Révész, 2012), few studies have compared multiple forms of feedback with respect to the differential demands they place on learners' WM capacity (see also Li, this volume). For example, even though Mackey et al.'s (2010) study involved interlocutors providing various types of feedback (e.g., prompts and recasts), the researchers' primary focus was on the relationship between WM capacity and the production of modified output, not on the mediating role of the type of feedback. Sagarra and Abbuhl (2013) examined multiple forms of feedback and the mediating role of WM and found that WM was positively related to the performance of students receiving oral recasts (both enhanced and unenhanced), but not to students receiving written recasts or utterance rejections. Given the scarcity of studies in this area and the importance of examining learner difference variables in studies of feedback and second language acquisition, further research is warranted.

The Study

As a whole, the results of previous studies suggest that there is a complex relationship between feedback explicitness, modality, and WM in learners' linguistic development. Few studies, however, have examined feedback modality, and fewer still have investigated the mediating roles of WM and feedback explicitness (including the value of recast enhancements and combining recasts with negative evidence). If the field is to determine *under what conditions* recasts work and *for whom,* further research investigating these variables will be necessary. To this end, the present study examines the following research questions:

1. Are there differential effects on Spanish learners' accuracy with noun-adjective gender or number agreement after receiving automated computer-delivered written recasts, written recasts with typographical enhancement, or written recasts preceded by a metalinguistic explanation?
2. Is there a significant difference among the three conditions with respect to the production of targetlike modified output?
3. What is the role of WM in these results?

Gender and Number Agreement in Spanish

The target structure was gender (masculine or feminine) and number (singular or plural) agreement between inanimate nouns and descriptive adjectives with transparent gender in Spanish. Nouns and adjectives with transparent gender like the ones used in this study end with the suffix /*-o*/ for masculine nouns (*libro blanco* "white-masculine-singular book-masculine-singular") and the suffix /*-a*/ for feminine nouns (*silla blanca* "white-feminine chair-feminine"). While gender in inanimate nouns is assigned arbitrarily, number is normally assigned semantically (quantity). In count nouns with transparent gender like the ones used in this study, regular number is formed by adding the suffix /*-s*/ to singular nouns (*libros blancos* "white-masculine-plural books-masculine-plural," *sillas blancas* "white-feminine-plural chairs-feminine-plural") (Zagona, 2002). Noun-adjective gender agreement is difficult for nonnative speakers of Spanish because it is redundant (determiners and nouns also mark gender), unreliable (more than 600 nouns do not follow the regular endings of /*-o*/ and /*-a*/), and non-salient (gender is marked by bound unstressed suffixes). This difficulty increases when learners are native speakers of a genderless L1 like English (gender is only marked on a few nouns like actor/actress). Furthermore, in English, adjectives are not marked for gender or number.

Method

Participants

The participants were 127 English learners of Spanish enrolled in a first-semester Spanish course at a U.S. university; they received extra credit and monetary compensation for participating. There were five criteria for participation: (1) being

between 18 and 40 years old (WM and processing speed start decreasing at the age of 40: Park et al., 2003), (2) not knowing any morphologically rich language apart from their basic knowledge of Spanish, (3) attending a metalinguistic explanation session (for the metalinguistic explanation group) and the vocabulary and grammar screening tests, (4) scoring at or above 80% on the vocabulary test (to ensure that lack of vocabulary knowledge did not impact the results), and (5) scoring at or below 25% on the grammar pretest (to ensure that previous knowledge of Spanish grammar was not a confounding factor). Finally, participants were randomly assigned to one of four conditions that differed in terms of type of written feedback provided: (1) no feedback (control) (*Move on to the next sentence*), (2) recasts (correct sentence), (3) recasts typographically enhanced via blue, bolding and capitalization of the adjectival suffix (henceforth "enhanced recasts") based on studies indicating positive effects of these enhancements on L2 development (for a review, see Wong, 2005), and (4) recasts combined with a preemptive metalinguistic explanation (henceforth "metalinguistic recasts"). The data for the first three conditions come from Sagarra and Abbuhl (2013) and the data for metalinguistic recasts are new.

Materials and Procedure

Tests and tasks were carried out during a two-week period in the classroom (as a group), the computer lab (as a group), and an office (individually). For internal validity purposes, instructors did not provide any grammatical explanation or activity about the target structure during the experiment.

Classroom Tasks

In the classroom, participants attended a vocabulary presentation (50 minutes). The purpose of the vocabulary presentation was to introduce students to the meaning of the nouns and adjectives in the study through pictures and L2 words. The presentation included only masculine singular forms so as to avoid providing additional clues that could affect the results. A trained research assistant conducted both this presentation and the metalinguistic explanation session via PowerPoint to ensure homogeneity among the groups.

Computer Laboratory Tasks

Four days after the vocabulary presentation, participants went to the computer laboratory to complete a vocabulary and a grammar screening test that also served as a pretest (10 minutes) and a metalinguistic explanation session (metalinguistic recast group only) (40 minutes). One week later, participants completed the treatment (maximum 35 minutes) and the immediate written posttest (15 minutes), and one week after the treatment they completed delayed written posttest I (15 minutes). One month after the treatment, participants completed delayed written posttest II (15 minutes) and the WM test (15 minutes). The vocabulary screening test involved matching the Spanish nouns of the experiment to their English equivalents and translating the target adjectives from English to Spanish. The goal of the metalinguistic explanation was to teach the students the rules of noun-adjective gender and number agreement in Spanish.

The pretest, treatment, and written posttests consisted of a cloze activity that required students to supply a Spanish adjective to sentences such as *la silla es ____ (white)* ("the chair is ____ (white)") using a web-based course management system called "A New Global Environment for Learning." The treatment had 32 sentences with the 32 nouns and the 16 adjectives covered in the vocabulary presentation, and the pretest and written posttests had 32 additional new nouns to test whether the learners could apply what they had learned to new exemplars. Familiar and new nouns named concrete objects and appeared in the Spanish textbook of the course, and there was an equal number of masculine, feminine, singular, and plural nouns. In addition, new nouns were English cognates, based on Nash's (1997) *NTC's Dictionary of Spanish Cognates*. A given noun-adjective combination appeared only once to control for practice effects, and both noun-adjective combinations and adjectives were pseudo-randomized in the pretest and posttests. Finally, in the pretest, treatment, and written posttests, sentences were balanced for gender and number and presented one at a time, and 75% of the sentences had transparent gender nouns (the focus of the study) and 25% opaque gender nouns that served as fillers.

For the treatment, participants were randomly assigned to one of four types of written feedback: (1) no feedback, (2) recasts, (3) enhanced recasts, and (4) metalinguistic recasts. Thus, groups (2) and (4) received the same type of feedback, with the only difference being that group (4) had been exposed to an explicit metalinguistic explanation one week earlier. For all the groups except the control, feedback was provided immediately after each sentence before the participant was able to move on to the next trial.

Following the last written posttest, participants completed a WM test that was adapted from Waters and Caplan's (1996) reading span test. Based on research showing that WM is language independent (e.g., Osaka, Osaka, & Groner, 1993; Xue, Dong, Jin, & Chen, 2004), the test was administered in the participants' L1. For the test, they completed three practice sets and 20 randomized experimental sets (80 experimental sentences total, each between 8 to 14 words long), each containing two to six sentences. For each set, participants looked at a 500-ms fixation sign, read the sentence silently, and indicated whether it was semantically plausible (*it was the milk that the child spilled*) or not (*it was the house that ate the frog*) by pressing a button (half of the sentences were plausible and half implausible). Then the word *Recall* appeared and participants wrote down the final word of each sentence within that set.

Office Tasks

In the office, participants met with a research assistant for one hour to complete two interactional posttests two-and-a-half months after the treatment: an oral posttest and an immediate repair posttest. As with the written pretest and post tests, these oral posttests each contained 32 sentences balanced for gender and number. The first consisted of an interactional one-way task where the student pretended to call a department store employee in order to ask her what products needed to be ordered (e.g., the student saw a picture of seven black tables in the list of objects that needed to be purchased and asked whether those items were

in stock: *¿*tienes mesas negros?* "do you have black-masculine-plural tables-feminine-plural?"). The research assistant would consult her picture and tell the student how many she had (e.g., *sí, tengo cinco* "yes, I have five"). Then the student wrote "2" next to the picture of seven black tables, indicating that two more would need to be ordered. The research assistant did not provide any feedback on the grammaticality of the students' utterances; in addition, *uno/una* "one" was not a possible answer as it agrees in gender with the noun. When the learner asked for an item that was not in stock, the research assistant simply said *No*.

The second oral posttest (the immediate repair posttest) was a spot-the-difference picture task involving before-and-after pictures of a remodeled apartment. The student was provided with the "before" pictures and was asked to determine which objects had changed color after the remodeling. The research assistant provided recasts to all utterances containing errors with noun-adjective agreement, and all recasts were accompanied by a short pause in order to give the learner the opportunity to modify her or his utterance (e.g., student: **en mi foto, la silla es naranjo* "in my picture, the chair-feminine-singular is orange-masculine-singular"; researcher: *la silla es naranja* "the-feminine-singular chair-feminine-singular is orange-feminine-singular", pause for the opportunity to repair, then *en mi foto también* "in my picture too" or *en mi foto, no* "in my picture, no"). The immediate repair posttest was included in order to assess the students' production of modified output.

These two oral posttests were preceded by a warm-up task, other oral activities (to make the goal of the study less obvious and assess L2 oral proficiency), and a wind-down task that was also used to obtain language background information. The oral activities and posttests had a model dialogue with masculine singular nouns at the beginning, and the two interlocutors shared their answers at the end of each task to mirror the communicative activities generally performed in class.

Scoring

The screening, pretest, and posttests received one point for correct answers and 0 for incorrect ones. Responses involving masculine singular nouns were not included in the statistical analyses, as learners tend to use this as the default form (e.g., White, Valenzuela, Kozlowska-Macgregor, & Yan-Kit, 2004). On the immediate repair posttest, the percentage of times the students produced modified output (a complete or partial repetition of the corrected utterance) after receiving feedback was calculated. The statistical analyses for the familiar and new nouns were combined as multiple dependent sample *t*-tests revealed no significant difference between the two types of nouns. With regard to the WM test, participants received one point per sentence (total 80 points) when they recalled the final word and made the correct plausibility judgment between 300 and 5,000 ms (college students need between 225 and 400 ms to process one word in their L1, Rayner & Pollatsek, 1989). We used a composite score following studies indicating a relationship between processing time and word recall (e.g., Barrouillet & Camos, 2007).

Results

Means and standard deviations can be found in Tables 4-1 and 4-2.

Inferential statistics for the written and the oral posttests were carried out independently because there were fewer participants on the oral posttests (the written posttests were conducted as part of the curriculum during regular class time but the oral posttests required learners to make an individual appointment with the researcher), and because the written posttests were conducted in a computer lab whereas the oral posttests were completed in an office.

Table 4-1 Descriptive Statistics for Linguistic Accuracy on Adjectives

	Control		Metalinguistic recasts		Written recasts		Enhanced recasts	
Test	**M**	**SD**	**M**	**SD**	**M**	**SD**	**M**	**SD**
Pretest								
Gender	.11	.32	.22	.42	.26	.16	.11	.31
Number	.20	.48	.41	.64	.36	.39	.15	.36
Immediate written posttest								
Gender	.08	.23	11.86	.66	9.00	5.20	8.95	4.31
Number	.85	.80	11.78	.64	9.76	4.02	9.29	4.21
Delayed written posttest I								
Gender	.15	.38	11.57	1.21	9.34	3.96	8.67	4.10
Number	.92	.95	11.13	2.73	9.97	3.91	8.67	4.10
Delayed written posttest II								
Gender	.77	1.48	11.14	1.32	8.62	4.98	7.52	5.36
Number	1.54	1.66	11.68	1.10	9.10	4.42	7.71	4.82
Delayed oral face-to-face interaction posttest I								
Gender	.15	.56	9.21	2.78	6.17	2.83	7.25	3.25
Number	.46	.66	10.05	1.99	8.67	3.11	6.25	3.17
Delayed oral face-to-face interaction posttest II (immediate repair)								
Gender	.54	.78	9.74	2.54	7.28	1.93	6.75	2.73
Number	2.00	2.45	10.13	2.03	9.06	2.56	8.17	1.90

k = 12 for each score. Also, control: *n* = 30 pretest, written posttests, *n* = 13 interactional posttests; metalinguistic recasts: *n* = 27 pretest, written posttests, *n* = 19 interactional posttests; recasts: *n* = 36 pretest, written posttests; *n* = 18 interactional posttests; enhanced recasts: *n* = 34 pretest, written posttests, *n* = 12 interactional posttests.

Table 4-2 Descriptive Statistics for Percentage of Production of Targetlike Modified Output on the Immediate Repair Test

	Control		Metalinguistic recasts		Recasts		Enhanced recasts	
	M%	**SD**	**M%**	**SD**	**M%**	**SD**	**M%**	**SD**
Gender	46.85	7.07	83.00	10.64	69.44	23.25	67.00	8.93
Number	61.92	7.09	93.11	7.32	95.39	4.37	97.42	2.90

To rule out potential differences among the groups that could bias the results, we conducted three one-way ANOVAs: one for WM scores, one for gender scores in the pretest, and one for number scores in the pretest. The three ANOVAs revealed no significant between-groups difference in WM ($F[3, 83] = 2.702, p > .05$) and pretest scores (gender, $F[3, 126] = 1.558, p > .05$; number agreement, $F(3, 126) = 1.773, p > .05$), suggesting that the groups were comparable before the treatment. The means (with standard deviations in parentheses) for the WM test were: control = 42.43 (11.74), recasts = 48.19 (12.62), enhanced recasts = 42.15 (10.33), and metalinguistic recasts = 40.33 (11.55). The means and standard deviations for the pretest and the posttests are shown in Table 4-1.

For linguistic accuracy, we carried out two ANCOVAs with WM as a covariate: two with a 2 (Structure: gender, number) × 3 (Posttest) × 4 (Group) factorial design for the written posttests, and two with a 2 (Structure) × 2 (Post test) × 4 (Group) factorial design for the oral posttests. The results of the two ANCOVAs revealed a significant main effect for Structure on the written posttests ($F[1, 158] = 5.983, p < .05$), because the groups tended to be more accurate with number than gender agreement. Although the groups also tended to produce number agreement more accurately than gender agreement on the oral posttests, the difference did not lead to a significant main effect ($F[1, 57] = .281, p > .05$). Most importantly, there was a significant main effect for Group, both on the written posttests ($F[3, 79] = 39.952, p < .01$) and the oral post tests ($F[3, 57] = 53.231, p < .01$). Bonferroni post hoc tests indicated that the metalinguistic recast group was more accurate than the other groups, and that the enhanced recast and the recast groups were in turn better than the control group for all the written posttests (all, $p < .01$). However, there was no significant main effect for Posttest (written posttests: $F[2, 158] = .480, p > .05$; oral posttests: gender: $F[1, 57] = 426, p > .05$) or WM (written posttests: $F[1, 79] = 5.664, p > .05$; oral posttests: $F[1, 57] = 1.135, p > .05$).

With regard to interactions, the interaction of Structure x Posttest × Group on the oral posttest was the only significant interaction ($F[3, 57] = 5.137, p < .01$). This interaction was caused by the general tendency for all groups to score higher on number than on gender agreement (English has number marking, but not gender marking on inanimate nouns) and to score higher on the second than on the first oral posttest (the immediate repair on the second posttest increased learners' focus on the target structure).

To examine the production of modified output on the immediate repair post test, we conducted an ANCOVA with a 2 (Structure) × 4 (Group) factorial design. In line with the results obtained for accuracy in the written and oral posttests, the findings indicated a significant main effect for Structure ($F[1, 57] = 20.453, p < .01$), and Bonferroni post hoc tests revealed that learners scored higher with number than gender agreement (see above for parallel findings with the written posttests). There was also a significant main effect for Group ($F[3, 57] = 53.070, p < .01$). Bonferroni post hoc tests showed that the metalinguistic recast group outperformed the rest with respect to gender agreement scores and the three recast groups were in turn better than the control group on gender and number

agreement scores. The lack of significant difference among the three recast groups on number agreement scores was due to ceiling effects. Finally, there was no significant effect for WM ($F[1, 57] = 4.677, p > .05$). As for interactions, there was a significant interaction of Structure x Group ($F[3, 57] = 7.340, p < .01$) because all groups tended to score higher on number than gender agreement. However, the interaction of Structure × WM was non-significant ($F[1, 57] = 5.130, p > .05$).

Discussion

The first research question asked whether Spanish learners' accuracy with noun-adjective gender or number agreement on written and oral posttests depended on whether the learner received unenhanced written recasts, typographically enhanced written recasts, or written recasts combined with preemptive negative evidence in the form of metalinguistic explanations. We found that on both the written posttests and oral face-to-face interactional posttests, the learners who received the combination of metalinguistic explanation and recasts were significantly more accurate than the learners who received typographically enhanced or unenhanced recasts. In addition, there was no significant difference between the enhanced and unenhanced recast groups, and all the recast groups outperformed the control group. The results of this study and a follow-up study, including a group that received metalinguistic explanations without recasts (in progress), which performed worse than the recasts + metalinguistic group, suggest that the provision of metalinguistic information prior to the task may have primed the learners and attuned them to the corrective feedback that was later provided. With noun-adjective agreement "on their radar", so to speak, and having had time to digest the information that was provided earlier, the learners in the combination group may have found that the recasts helped reactivate recently learned information.

Concerning the lack of a significant difference between the typographically enhanced and unenhanced written recast groups, our results are in line with previous studies reporting no difference between typographically enhanced and unenhanced input (e.g., Alanen, 1995; Leow, 1997; 2001; Izumi, 2002; Wong, 2003; Rott, 2007), as well as those comparing typographically enhanced and unenhanced written recasts in computer-based contexts (e.g., Sachs & Suh, 2007). As speculated by Cho (2010), written input enhancement may benefit receptive skills more than productive skills; as our measures targeted productive skills, any benefit of the visually enhanced recasts may not have been apparent. Future studies employing both receptive and productive measures of language acquisition will be needed to determine what effect, if any, visually enhanced recasts have on L2 development. The results of previous studies indicating a superiority of orally enhanced recasts over orally unenhanced recasts (e.g., Leeman, 2003; Nassaji, 2009) and even enhanced and unenhanced written recasts (Sagarra & Abbuhl, 2013) suggest that L2 processing in the oral mode

is more cognitively taxing than in the written mode because the oral mode does not allow learners to process the input in a self-paced manner or access previously processed information, something common in reading comprehension. Because oral input is cognitively more demanding, learners may benefit more from oral enhancements than they do from written ones. This may also explain why Sagarra and Abbuhl (2013) found WM effects in the learners' ability to benefit from oral, but not written, enhanced recasts on the oral, but not written, posttests.

Our second research question asked if the three conditions differentially impacted the production of modified output. We found that the metalinguistic recasts led to the greatest amount of targetlike modified output, with no significant difference between the enhanced and unenhanced recast groups (both recast groups, however, outperformed the control group). These results lend additional support to previous arguments that even if the corrective intent of a simple (bare) recast is detected by the learner, this may not be sufficient to help the learner *understand* the exact nature of the error (Sheen, 2010). Supplying metalinguistic explanations prior to the task may provide the necessary scaffolding that beginning learners need to make progress with certain non-salient and redundant aspects of L2 morphosyntax.

Our third and final research question asked what role WM played in how much learners are able to benefit from the three written feedback types under investigation. We did not find WM effects for any of the groups; that is, all learners, regardless of their WM capacity, benefited equally from the written feedback that was provided. This held true for both written and oral interactional posttests, linguistic accuracy and the production of targetlike modified output, and gender and number agreement. The lack of WM effects on number concord scores could be due to ceiling effects, but we believe that the reduced cognitive demands of processing written feedback may be responsible for the absence of a WM effect in the gender agreement scores. As learners were able to read and reread both their original utterance and the feedback that was provided, a comparison of the two forms was not cognitively demanding. Previous studies reporting a relationship between WM and recasts focused on *oral* recasts (e.g., Mackey et al., 2002, 2010; Mackey & Sachs, 2011). As oral recasts do place a heavy cognitive burden on the learner (both the original non-targetlike utterance and the reformulation have to be retained in memory long enough for a comparison of the two forms to be made), it is not surprising that learners with higher WM capacities tend to benefit more from oral recasts than lower WM span learners.

Future research, of course, will be needed to examine other linguistic targets with varying degrees of regularity and cross-linguistic similarity. In addition, investigating learners of differing levels of L2 proficiency will help determine the generalizability of the results presented here. Finer-grained measures of WM, including those that assess visuospatial WM along with phonological short-term memory, will also help determine if the effects of feedback are mediated by learners' capacities in those two areas.

Implications for Language Program Directors and Teachers

To recap, we found that learners who received metalinguistic explanations combined with recasts were significantly more accurate (and produced more targetlike modified output) than the learners who received either typographically enhanced or unenhanced recast groups. WM effects were also not found for any of the groups under investigation, suggesting that written recasts may be easier to process than the more ephemeral oral recasts. These results suggest that teachers should first of all consider taking advantage of current technologies that allow automated recasts to be provided in the written modality. For example, teachers could require students to attend laboratory sessions outside of the classroom where they would receive computer-generated feedback while completing tasks in the L2. In this context, learners may be attuned to form and faced with fewer distractions than are typically present in the classroom, which may help learners with diverse WM capacities make progress with certain redundant and non-salient aspects of L2 morphology. The results also suggest that teachers should consider providing metalinguistic explanations of the target morphosyntactic structures before sending students to the computer laboratory. These explanations could be provided in the traditional lecture format, but they could also be provided through interactive grammar tutorials that have accompanied many textbooks in recent years. If these are not available, students could be asked to consult websites that give explanations of the targets (e.g., http://www.studyspanish.com/lessons/adj1.htm for an explanation of noun adjective agreement in Spanish) or even videos on the target structure (e.g., http://www.youtube.com/watch?v=8v4xr6CES0c for the same target). These latter two options would be particularly helpful for hybrid or online courses.

In addition to automated feedback provided outside of the classroom, teachers could also consider providing individualized written recasts to certain morphosyntactic errors in their students' written work. In large classes where this is not feasible, teachers could take anonymous examples from students' written work, present them to the class, and provide the written recast (e.g., using an overhead projector) to the class as a whole. Alternatively, tutors from higher-level courses could be trained to provide written recasts to lower-level students. Language coordinators could help arrange tutor-tutee pairings, and could also help facilitate the needed training for the tutors. These tutoring sessions could take place face-to-face, but they could also be arranged for distance learners (e.g., by using software platforms such as Elluminate, which allow teachers and students to chat and use an interactive whiteboard during lessons).

It is a common observation in second and foreign language classrooms that certain learners benefit more from interactional feedback than others. Being aware of the potential effects of modality and explicitness, as well as interindividual differences in WM capacity, can help teachers better understand the differences they see in learners' processing and use of interactional feedback.

References

Alanen, R. (1995). Input enhancement and rule presentation in second language acquisition. In R. Schmidt (Ed.), *Attention and awareness in foreign language learning* (pp. 259–302). Honolulu, HI: University of Hawaii Press.

Al-Surmi, M. (2012). Learners' noticing of recasts of morpho-syntactic errors: Recast types and delayed recognition. *System, 40,* 226–236.

Baddeley, A. (2007). *Working memory, thought and action*. Oxford: Oxford University Press.

Bao, M., Egi, T., & Han, Y. (2011). Classroom study on noticing and recast features: Capturing learners' noticing with uptake and stimulated recall. *System, 39,* 215–228.

Barrouillet, P., & Camos, V. (2007). The time-based resource-sharing model of working memory. In N. Osaka, R. Logie, & M. D' Esposito (Eds.), *The cognitive neuroscience of working memory* (pp. 59–80). Oxford: Oxford University Press.

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

Carroll, S. (2001). *Input and evidence: The raw material of second language acquisition*. Amsterdam: John Benjamins.

Cho, M. (2010). The effects of input enhancement and written recall on noticing and acquisition. *Innovation in Language Learning and Teaching, 4,* 71–87.

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

Doughty, C. (2001). Cognitive underpinnings of focus on form. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 206–257). Cambridge: Cambridge University Press.

Doughty, C., & Varela, E. (1998). Communicative focus on form. In C. Doughty & J. Williams (Eds.), *Focus on form in classroom second language acquisition* (pp. 114–138). Cambridge: Cambridge University Press.

Egi, T. (2007). Interpreting recasts as linguistic evidence: The roles of linguistic target, length, and degree of change. *Studies in Second Language Acquisition, 29,* 511–537.

Ellis, R., Loewen, S., & Erlam, R. (2006). Implicit and explicit corrective feedback and the acquisition of L2 grammar. *Studies in Second Language Acquisition, 28,* 339–368.

Gass, S., & Lewis, K. (2007). Perceptions about interactional feedback: Differences between heritage language learners and non-heritage language learners. In A. Mackey (Ed.), *Conversational interaction in second language acquisition: A collection of empirical studies* (pp. 79–99). Oxford: Oxford University Press.

Goo, J., & Mackey, A. (2013). The case against the case against recasts. *Studies in Second Language Acquisition, 35*, 127–165.

Han, Z. (2002). A study of the impact of recasts on tense consistency in L2 output. *TESOL Quarterly, 36,* 543–572.

Havranek, G. (2002). When is corrective feedback most likely to succeed? *International Journal of Educational Research, 37,* 255–270.

Hummel, K., & French, L. (2010). Phonological memory and implications for the second language classroom. *Canadian Modern Language Review, 66,* 371–291.

Iwashita, N. (2003). Negative feedback and positive evidence in task-based interaction: Differential effects on L2 development. *Studies in Second Language Acquisition, 25,* 1–36.

Izumi, S. (2002). Output, input enhancement, and the noticing hypothesis. *Studies in Second Language Acquisition, 24*, 541–577.

Kim, Y. (2006). Effects of Input elaboration on vocabulary acquisition through reading by Korean learners of English as a foreign language. *TESOL Quarterly, 40,* 341–373.

Kim, J., & Han, Z. (2007). Recasts in communicative EFL classes: Do teacher intent and learner interpretation overlap? In A. Mackey (Ed.), *Conversational interaction in second language acquisition: A collection of empirical studies* (pp. 269–297). Oxford: Oxford University Press.

Lai, C., & Zhao, Y. (2006). Noticing and text-based chat. *Language Learning & Technology, 10,* 102–120.

Lai, C., Fei, F., & Roots, R. (2008). The contingency of recasts and noticing. *CALICO Journal, 26,* 70–90.

Leeman, J. (2003). Recasts and second language development: Beyond negative evidence. *Studies in Second Language Acquisition, 25,* 37–63.

Leow, R. (1997). The effects of input enhancement and text length on adult L2 readers' comprehension and intake in second language acquisition. *Applied Language Learning, 8*, 151–182.

Leow, R. (2001). Do learners notice enhanced forms while interacting with the L2? An online and offline study of the role of written input enhancement in L2 reading. *Hispania, 84*, 496–509.

Leow, R., Egi, T., Nuevo, A., & Tsai, Y. (2003). The roles of textual enhancement and type of linguistic item in adult L2 learners' comprehension and intake. *Applied Language Learning, 13,* 1–16.

Loewen, S., & Erlam, R. (2006). Corrective feedback in the chatroom: An experimental study. *Computer Assisted Language Learning, 19,* 1–14.

Loewen, S., & Reissner, S. (2009). A comparison of incidental focus on form in the second language classroom and chatroom. *Computer Assisted Kang Learning, 22,* 101–114.

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, NY: Academic Press.

Long, M. (2007). *Problems in SLA*. Mahwah, NJ: Lawrence Erlbaum Associates.

Long, M., Inagaki, S., & Ortega, L. (1998). The role of implicit negative feedback in SLA: Models and recasts in Japanese and Spanish. *The Modern Language Journal, 82,* 357–371.

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

Lyster, R. (1998). Recasts, repetition and ambiguity in L2 classroom discourse. *Studies in Second Language Acquisition, 20,* 51–80.

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

Lyster, R., & Ranta, L. (1997). Corrective feedback and learner uptake. *Studies in Second Language Acquisition, 19,* 37–66.

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

Mackey, A., Adams, R., Stafford, C., & Winke, P. (2010). Exploring the relationship between modified output and working memory capacity. *Language Learning, 60,* 501–533.

Mackey, A., Gass, S. M., & McDonough, K. (2000). How do learners perceive interactional feedback? *Studies in Second Language Acquisition, 22,* 471–497.

Mackey, A., & Goo, J. (2007). Interaction research in SLA: A meta-analysis and research synthesis. In A. Mackey (Ed.), *Conversational interaction in second language acquisition: A collection of empirical studies* (pp. 407–452). Oxford: Oxford University Press.

Mackey, A., & Oliver, R. (2002). Interactional feedback and children's L2 development. *System, 30,* 459–477.

Mackey, A., & Philp, J. (1998). Conversational interaction and second language development: Recasts, responses and red herrings? *Modern Language Journal, 82,* 338–356.

Mackey, A., Philp, J., Egi, T., Fujii, A., & Tatsumi, T. (2002). Individual differences in working memory, noticing of interactional feedback, and L2 development. In P. Robinson (Ed.), *Individual differences and instructed language learning* (pp. 181–210). Amsterdam: John Benjamins.

Mackey, A., & Sachs, R. (2011). Older learners in SLA research: A first look at working memory, feedback, and L2 development. *Language Learning, 62*(3), 704–740.

Nash, R. (Ed.) (1997). *NTC's dictionary of Spanish cognates.* Lincolnwood, IL: NTC Publishing group.

Nassaji, H. (2009). Effects of recasts and elicitations in dyadic interaction and the role of feedback explicitness. *Language Learning, 59*, 411–452.

Oliver, R., & Mackey, A. (2003). Interactional context and feedback in child ESL classrooms. *The Modern Language Journal, 87,* 519–533.

Osaka, M., Osaka, N., & Groner, R. (1993). Language independent working memory: Evidence from German and French reading span tests. *Bulletin of the Psychonomic Society, 31*, 117–118.

Panova, I., & Lyster, R. (2002). Patterns of corrective feedback and uptake in an adult ESL classroom. *TESOL Quarterly, 36,* 573–595.

Park, D., Welsh, R., Marschuetz, C., Gutchess, A., Mikels, J., Polk, T., Noll, D., & Taylor, S. (2003). Working memory for complex scenes: Age differences in frontal and hippocampal activations. *Journal of Cognitive Neuroscience, 15*(8), 1122–1134.

Payne, J., & Ross, B. (2005). Synchronous CMC, working memory, and L2 oral proficiency development. *Language Learning & Technology, 9,* 35–54.

Payne, J., & Whitney, P. (2002). Developing L2 oral proficiency through synchronous CMC: Output, working memory, and interlanguage development. *CALICO Journal, 20,* 7–32.

Philp, J. (2003). Constraints on "noticing the gap": Nonnative Speakers' noticing of recasts in NS-NNS interaction. *Studies in Second Language Acquisition, 25,* 99–126.

Rayner, K., & Pollatsek, A. (1989). *Psychology of reading.* Englewood Cliffs, NJ: Prentice Hall.

Révész, A. (2012). Working memory and the observed effectiveness of recasts on different L2 outcome measures. *Language Learning, 62,* 93–132.

Rott, S. (2007). The effect of frequency of input-enhancements on word learning and text comprehension. *Language Learning, 57*, 165–199.

Sachs, R., & Suh, B. (2007). Textually enhanced recasts, learner awareness, and L2 outcomes in synchronous computer-mediated interaction. In A. Mackey (Ed.), *Conversational interaction in second language acquisition: A collection of empirical studies* (pp. 197–227). Oxford: Oxford University Press.

Sagarra, N., & Abbuhl, R. (2013). Optimizing the noticing of recasts via computer-delivered feedback: Evidence that oral input enhancement and working memory help L2 learning. *Modern Language Journal, 97,* 196–216.

Sauro, S. (2009). Computer-mediated corrective feedback and the development of L2 grammar. *Language Learning & Technology, 13,* 96–120.

Sheen, Y. (2006). Exploring the Relationship between characteristics of recasts and learner uptake. *Language Teaching Research, 10,* 361–392.

Sheen, Y. (2007). The effects of corrective feedback, language aptitude, and learner attitudes on the acquisition of English articles. In A. Mackey (Ed.), *Conversational interaction in second language acquisition: A collection of empirical studies* (pp. 301–322). Oxford: Oxford University Press.

Sheen, Y. (2010). Differential effects of oral and written corrective feedback in the ESL classroom. *Studies in Second Language Acquisition, 32*, 203-234.

Shook, D. (1999). What foreign language reading recalls reveal about the input-to-intake phenomenon. *Applied Language Learning*, 10, 39–76.

Warschauer, M. (1996). Comparing face-to-face and electronic discussion in the second language classroom. *CALICO Journal, 13,* 7–25.
Waters, G., & Caplan, D. (1996). The measurement of verbal working memory capacity and its relation to reading comprehension. *The Quarterly Journal of Experimental Psychology, 49A,* 51–79.
White, L., Valenzuela, E., Kozlowska-Macgregor, M., & Yan-Kit, L. (2004). Gender and number agreement in nonnative Spanish. *Applied Psycholinguistics, 25*, 105–133.
Wong, W. (2003). Textual enhancement and simplified input: Effects on L2 comprehension and acquisition of non-meaningful grammatical form. *Applied Language Learning, 13,* 17–45.
Wong, W. (2005). *Input enhancement: From theory and research to the classroom*. Boston, MA: McGraw-Hill.
Xue, G., Dong, Q., Jin, Z., & Chen, C. (2004). Mapping of verbal working memory in nonfluent Chinese-English bilinguals with functional MRI. *Neuroimage, 22,* 1–10.
Yilmaz, Y., & Yuksel, D. (2011). Effects of communication mode and salience on recasts: A first exposure study. *Language Teaching Research, 15,* 457–477.
Zagona, K. (2002). *The syntax of Spanish*. Cambridge: Cambridge University Press.