

Studies in Second Language Acquisition, 2013, **35**, 127–165.
doi:10.1017/S0272263112000708

THE CASE AGAINST THE CASE AGAINST RECASTS

Jaemyung Goo

Gwangju National University of Education

Alison Mackey

Georgetown University

In the previous 20 years, more than 60 studies have been carried out within the input and interaction approach to SLA (Long, 2007; Mackey 2012), many of which have found positive associations between different types of recasts and the learning of a range of linguistic forms for a number of different second languages (L2s), in different learning contexts, with adults and with children. However, the following claims also appear: (a) recasts are not effective, (b) recasts are effective only in laboratories and not in classrooms, and (c) other types of feedback are more effective when compared with recasts. We demonstrate important methodological and interpretative problems in the small number of studies on which these negative claims are based, including issues with (a) modified output opportunities, (b) single-versus-multiple comparisons, (c) form-focused instruction, (d) prior knowledge, and (e) out-of-experiment exposure. We conclude by suggesting that making a case against recasts is neither convincing nor useful for advancing the field and that more triangulated approaches to research on all types of corrective feedback, employing varied and rigorous methodological designs, are necessary to further our understanding of the role of corrective feedback in L2 learning.

According to the interactionist perspective on second language (L2) learning (Gass, 1997, 2003; Gass & Mackey, 2006, 2007; Long, 1996, 2007;

Correspondence concerning this article should be addressed to Jaemyung Goo, Department of English Education, Gwangju National University of Education, 55 Pilmundaero Buk-gu, Gwangju, 500-703, South Korea; e-mail: jg349@gnue.ac.kr.

Mackey, 2007, 2012; Mackey, Abbuhl, & Gass, 2012; Pica, 1994, 1996), negotiation for meaning during interaction provides favorable grounds for L2 learning to occur because it offers L2 learners potentially beneficial opportunities to receive comprehensible input (Krashen, 1982, 1985), to produce output through which they can test their hypotheses about L2 forms and process the L2 syntactically (Swain, 1985, 1995, 2005), and to become aware of the gap between their interlanguage and the target language (Robinson, Mackey, Gass, & Schmidt, 2012; Schmidt, 1990, 1993, 1995, 2001; Schmidt & Frota, 1986). Together, these elements can lead learners to make important efforts to refine and restructure their interlanguage. Beneficial effects of interaction—and of corrective feedback provided during interaction—on the acquisition of L2 lexical and grammatical features are clearly evidenced in recent meta-analyses (e.g., Keck, Iberri-Shea, Tracy-Ventura, & Wa-Mbaleka, 2006; Mackey & Goo, 2007; but see also Li, 2010; Lyster & Saito, 2010; Russell & Spada, 2006, for other meta-analyses involving corrective feedback in L2 research). However, as interaction researchers generally agree, the relationship between conversational interaction involving negotiation for meaning and L2 acquisition is not a simple causal one but is complex and likely to be impacted by various factors such as developmental readiness, the type of corrective feedback, internal learner capacities (e.g., working memory [WM]), social factors, task characteristics, context, and the type of target structure, just to name a few (Goo, 2012; Long, 1996, 2007; Mackey, 2007, 2012; Mackey et al., 2012; Philp & Mackey, 2010).

Given the complex nature of the interaction-learning relationship, researchers have become particularly interested in the specific features of interaction that may influence the extent to which negotiation for meaning during interaction benefits L2 learners (for a review and summary of interaction research, see Gass & Mackey, 2006, 2007; Long, 2007; Mackey, 2012; Mackey et al., 2012; Mackey & Goo, in press; Spada & Lightbown, 2009). In particular, considerable attention has been paid to recasts, leading to discussions of their efficacy on L2 learning in interactional settings, often in comparison with other instructional methods or other types of corrective feedback (e.g., Ammar, 2008; Ammar & Spada, 2006; Ayoun, 2001; Braidi, 2002; Carpenter, Jeon, MacGregor, & Mackey, 2006; Dilans, 2010; Doughty & Varela, 1998; Egi, 2007a, 2007b, 2010; Ellis, 2007; Ellis, Basturkmen, & Loewen, 2001; Ellis, Loewen, & Erlam, 2006; Ellis & Sheen, 2006; Erlam & Loewen, 2010; Goo, 2012; Han, 2002; Ishida, 2004; Iwashita, 2003; Lai, Fei, & Roots, 2008; Loewen, 2009; Loewen & Erlam, 2006; Loewen & Nabei, 2007; Loewen & Philp, 2006; Long, 1996, 2007; Long, Inagaki, & Ortega, 1998; Lyster, 1998a, 1998b, 2004; Lyster & Izquierdo, 2009; Lyster & Ranta, 1997; Mackey, Gass, & McDonough, 2000; Mackey & Philp, 1998; McDonough, 2007; McDonough & Mackey, 2006; Nassaji, 2009; Nicholas, Lightbown, & Spada, 2001; Ohta, 2000; Oliver, 1995, 1998, 2000; Panova & Lyster, 2002; Philp, 2003; Révész, 2009;

Révész & Han, 2006; Romanova, 2010; Sachs & Suh, 2007; Sagarra, 2007; Saito & Lyster, 2012; Sato & Lyster, 2007; Sauro, 2009; Sheen, 2004, 2006, 2007a, 2007b, 2008, 2010; Shintani & Ellis, 2010; Smith, 2010; Trofimovich, Ammar, & Gatbonton, 2007; Varnosfadrani & Basturkmen, 2009; Yang & Lyster, 2010).

Even if the study of recasts (and other forms of negative evidence) in first language (L1) research has been somewhat downplayed (e.g., Gordon, 1990; Grimshaw & Pinker, 1989; Morgan, Bonamo, & Travis, 1995; Pinker, 1989)—despite some empirical evidence in favor of the potential value of recasts in L1 development (e.g., Baker & Nelson, 1984; Bohannon & Stanowicz, 1988; Farrar, 1992; Saxton, 1997; Saxton, Backley, & Gallaway, 2005; Saxton, Kulcsar, Marshall, & Rupra, 1998)—L2 research on recasts has gained momentum, particularly in light of communicative and task-based approaches to language teaching and the interaction approach to L2 learning. This has yielded a barrage of descriptive and experimental studies (see Braidí, 2002; Ellis & Sheen, 2006; Long, 2007; Mackey, 2012; Nicholas et al., 2001, for reviews of L2 research on recasts). Space constraints preclude a historical review of all the research on recasts here, but in general, the amount of work done suggests it is a prominent and promising area of interest in L2 interaction research, and reviews can be found in other places, including the plethora of recent handbooks and encyclopedias (see, for example, Chapelle, 2012; Gass & Mackey, 2012; Hinkel, 2011).

WHAT ARE RECASTS AND WHY SHOULD WE STUDY THEM?

Drawing from the L1 literature, Long (1996) defined recasts as “utterances that rephrase a child’s utterance by changing one or more sentence components (subject, verb, or object) while still referring to its central meanings” (p. 434). Long (2007) later defined a corrective recast as “a reformulation of all or part of a learner’s immediately preceding utterance in which one or more nontargetlike (lexical, grammatical, etc.) items is/are replaced by the corresponding target language form(s), and where, throughout the exchange, the focus of the interlocutors is on meaning, not language as object” (p. 77; but see also Ellis & Sheen, 2006; Nicholas et al., 2001, for similar definitions). In short, recasts are more targetlike versions of learners’ nontargetlike utterances.

There is general agreement that recasts, although classified as implicit negative feedback in usual taxonomies of various types of feedback (e.g., Long, 2007; Long & Robinson, 1998), constitute positive evidence (i.e., the provision of targetlike input) as well as negative evidence (i.e., the provision of a slightly different alternative to learners’ original output to signal that an error has occurred). The immediate juxtaposition of the learner’s erroneous utterance and the corrective recast (a) results

in semantic transparency—namely, the temporal immediacy of the recast along with the fact that the learner is likely to understand all or part of the recast as it is a reformulated version of what she or he just said allows L2 learners to utilize more attentional resources to focus on form and form-function mapping; (b) enhances the saliency of target forms and precipitates the noticing of the changes made to the learner's original utterance; and, as a consequence, (c) leads L2 learners to compare the target form included in a recast with the erroneous form that they just produced (some or all of these issues are pointed out in a number of early papers, including Doughty, 2001; Doughty & Varela, 1998; Leeman, 2003; Long, 1996, 2007; Long et al., 1998; Long & Robinson, 1998). As Long (2007) explains, “recasts convey needed information about the target language *in context*, when interlocutors share a *joint attentional focus*, and when the learner already has *prior comprehension* of at least part of the message, thereby facilitating form-function mapping” (p. 77). He further notes that “learners are *vested* in the exchange, as it is their message that is at stake, and so will probably be *motivated* and *attending*, conditions likely to facilitate *noticing* of any new linguistic information in the input” (pp. 77–78).

Researchers have found that a number of factors mediate the degree to which recasts facilitate language learning. For example, Mackey and Philp (1998) provided evidence that developmental readiness may be a mediating factor, with developmentally more advanced learners benefiting more from recasts than developmentally less advanced learners (see also Ammar, 2008; Ammar & Spada, 2006; Trofimovich et al., 2007). Similarly, Ammar and Spada (2006) suggested that high-proficiency learners tend to be more sensitive to corrective feedback and more likely to notice the corrective nature of recasts than lower-proficiency learners.

The efficacy of recasts has also been argued to depend, to a great extent, on factors such as intonation, length, and number of changes. These factors may impact the saliency or noticeability of recasts and, consequently, L2 learning outcomes (e.g., Egi, 2007a, 2007b, 2010; Loewen & Philp, 2006; Nassaji, 2009; Philp, 2003; Sheen, 2006). In general, researchers have found that short or partial recasts with one or two corrective changes are more effective at facilitating noticing than long or full recasts with more changes, most likely for reasons such as limitations in working memory or cognitive capacity. The saliency of recasts may also be affected by setting and discourse contexts—for example, the nature of the language activity in which the learners are engaged.

The type of target language feature is also likely to mediate the extent to which learners benefit from different types of recasts. As is often pointed out, the effectiveness of L2 instruction can depend on the complexity or saliency of the various target structures and the perceived

difficulty of those structures (e.g., de Graaff & Housen, 2009; Dörnyei, 2009; Spada, 2011; Spada & Lightbown, 2008; see DeKeyser, 2005, for discussions of salience and difficulty). Thus, recasts may be more effective with some linguistic areas or some grammatical structures than with others (e.g., Ellis, 2007; Iwashita, 2003; Leeman, 2003; Mackey et al., 2000; Ortega & Long, 1997). As Long (2007) explains, “recasts or other delicate, unobtrusive forms of corrective feedback work satisfactorily for some linguistic targets (e.g., meaning-bearing items) better than others, but more explicit, more intrusive intervention is required for communicatively redundant, acoustically nonsalient forms” (p. 112). Long also notes that recasts may be more effective for linguistic structures or forms that are difficult to learn and thus require long-term treatments, whereas explicit feedback may be sufficient for relatively easy structures or forms requiring short-term treatments.

Recast researchers have also asked whether the effectiveness of recasts may be affected to varying degrees by individual differences, such as language aptitude, WM capacity, intelligence, personality, motivation, learning styles, and learning strategies (see, for instance, Dörnyei, 2005, 2009; Dörnyei & Skehan, 2003; Ellis, 2004; Goo, 2012; Mackey, 2012; Mackey, Philp, Egi, Fujii, & Tatsumi, 2002; Robinson, 2002, 2005, 2007; Sagarra, 2007; Sawyer & Ranta, 2001; Sheen, 2007b, 2008; Skehan, 2002; Trofimovich et al., 2007). Other potential—but underexplored—factors that may mediate the effects of recasts on L2 learning include age, gender, interlocutors, settings, and task characteristics. For example, Mackey, Oliver, and Leeman (2003) investigated the impact of age and interlocutors on negative feedback and negotiation patterns; Oliver (2000) examined the effect of age and settings on the provision and use of negative feedback; Ross-Feldman (2007) looked at the relationship between gender and learning opportunities in interaction; and Révész (2009), Révész and Han (2006), and Révész, Sachs, and Mackey (2011) discuss the impact of task characteristics on the effectiveness of recasts. Additionally, researchers have questioned whether the role of recasts may vary depending on the type of outcome measure that is used, and in particular, whether those measures trigger mental processing identical to that employed during learning. As Spada and Lightbown (2008) note, “according to TAP [transfer appropriate processing], learners retrieve knowledge best if the processes for retrieval are similar to those that were used in the learning condition” (p. 190). Thus, outcome measures need to be taken into account when assessing the potential of recasts for L2 learning (Mackey, 2012).

The recast literature is replete with calls for more studies to be carried out to obtain further insights into how these myriad factors interact to impact—or not—recast-driven L2 learning. However, what is clear is that regardless of the existence of factors that mediate the extent to which learners can benefit from recasts, empirical evidence

for a beneficial role of recasts in SLA has been observed in a great deal of interaction research.

It is important to note at this point that we are not advocating a recasts-only approach to the provision of feedback. Instead, we are pointing out that many recast researchers are now focusing on why, when, and how recasts are effective in and of themselves, as opposed to comparing recasts with other forms of feedback to see if they are more, less, or equally effective. This is because most recast researchers recognize a number of different routes to learning associated with different types of corrective feedback. In other words, because recasts and other forms of feedback are believed to be inherently different, such studies typically focus on in-depth examinations of one form of feedback or another, as opposed to comparisons of multiple forms.

THE CASE FOR RECASTS

As previously noted, L2 research to date has provided evidence that recasts facilitate L2 development across research settings, learning contexts, and modes of interaction (e.g., Ayoun, 2001; Doughty & Varela, 1998; Goo, 2012; Han, 2002; Ishida, 2004; Iwashita, 2003; Leeman, 2003; Loewen & Nabei, 2007; Long et al., 1998; Mackey & Philp, 1998; McDonough, 2007; McDonough & Mackey, 2006; Nassaji, 2009; Petersen, 2010; Sachs & Suh, 2007; Sagarra, 2007; Saito & Lyster, 2012; Trofimovich et al., 2007; see also Long, 2007, and Mackey, 2012, for a review of research on recasts, and Mackey & Goo, 2007, for a meta-analysis showing large mean effect sizes for recasts). Most of these studies are laboratory-based experiments conducted in either foreign or second language learning contexts; however, some classroom-based studies have also reported a facilitative role for recasts (e.g., Doughty & Varela, 1998; Goo, 2012; Loewen & Nabei, 2007).

In a second language context, for instance, Mackey and Philp (1998) conducted an experiment in which English as a second language (ESL) learners participated in interactional activities with native speaker (NS) interlocutors. Their findings showed that recasts provided during dyadic activities between NSs and nonnative speakers (NNSs) facilitated the development of English question formation, as measured by Pienemann and Johnston's (1987) ESL question developmental sequence. They also found that students who were developmentally ready benefited more from recasts than those who were not. Han's (2002) experimental study also confirmed beneficial effects for recasts on L2 learning in a second language context. Han examined the effects of recasts on tense consistency in ESL learners' output and found that recasts had a positive impact on students' ability to maintain tense consistency in their oral and written productions. Moreover, learner awareness of tense

consistency (as measured by the frequency of self-correction in the tasks) in the recast group increased to a greater degree than in the non-recast group. In Nassaji's (2009) study, recasts led to a greater increase in learner accuracy than did elicitations. More recently, Saito and Lyster (2012) showed that Japanese ESL learners, who received recasts in addition to form-focused instruction (FFI), significantly improved their pronunciation of English /ɪ/ in spontaneous speech. No improvement was witnessed for those who received FFI only. Saito and Lyster's finding is in line with Mackey et al.'s (2000) observation that phonological recasts were likely to be perceived as being about phonology, which increases the likelihood of learners benefiting from recasts.

Evidence in favor of the effectiveness of recasts has also been found in experimental studies, at least for certain structures. Long et al. (1998) reported two experiments in which recasts were compared with models (i.e., positive evidence) in terms of their effects on L2 learning, one for Japanese (i.e., adjective ordering and a locative construction including the morphological target *-kute*, which is the combination of the morphemes *-ku* and *-te* used to link two adjectives by making the first adjective gerundive) and the other for Spanish (i.e., topicalization of direct objects and adverb placement). Recasts proved to be beneficial for the learning of Spanish adverb placement but not the other targets (see Ortega & Long, 1997, regarding the Spanish experiment). Iwashita (2003) reported that recasts were predictive of learner performance on the Japanese *te*-form verb (i.e., a verb morpheme used to indicate commands and requests in Japanese) on the immediate posttest—regardless of learner performance on the pretest—but not on the other target structures (i.e., Japanese locative word order and locative particle use). Using a time-series design, Ishida (2004) observed a significant increase in accuracy in her Japanese as a foreign language learners' use of the Japanese aspectual form *-te i-(ru)* after intensive recasting. Comparing four groups—that is, recasts, enhanced salience, negative evidence, and a control—Leeman (2003) found that the recast group was more accurate than the control group on both Spanish gender and number agreement on the immediate posttest (i.e., oral production) and on number agreement on the delayed posttest.

In a foreign language context, McDonough and Mackey (2006) replicated Mackey and Philp's (1998) finding that recasts promoted ESL question development by showing that recasts provided during task-based interaction were significantly predictive of Thai English as a foreign language (EFL) learners' question development (i.e., advancing from Stage 4 to Stage 5 in their production of questions based on Pienemann and Johnston's [1987] English question developmental sequence). In another study conducted in a Thai EFL context, McDonough (2007) compared recasts with clarification requests in terms of their relative effects on the emergence of English simple past activity verbs. Recasts and clarification

requests were both found to be facilitative of learners' production of the target structure, with no significant differences between the two. Similarly, Lyster and Izquierdo (2009) found that both recasts and prompts led to gains with grammatical gender markers in French. As in McDonough (2007), no significant differences were found between the recast and nonrecast (i.e., prompt) conditions. Studies that compared recasts and prompts and found no differences are included in the Case for Recasts section because of the fact that recasts were found to be better than no feedback or control conditions; to us, this is evidence that recasts work. Whether or not they work equally well, better, or worse than other types of feedback is less the issue because we believe that direct comparisons between recasts and other types of feedback can be problematic.

Turning to classroom-based studies, Doughty and Varela (1998) reported that their ESL learners who received recasts showed significant pre- to posttest gains in accuracy on the simple past and the past conditional on both oral and written measures. It should be noted, however, that those in the recast condition were provided with an explicit type of recast, which Doughty and Varela called *corrective recasting*; that is, they received a recast preceded by a repetition with stress and rising intonation in some instances, which is argued by some researchers to be a double feedback move (e.g., Lyster, 1998b; Lyster & Izquierdo, 2009). Loewen and Nabei (2007) also found some evidence for recasts as their recast group outperformed the no-feedback and control groups in performance on English question formation on the timed grammaticality judgment test (GJT). Loewen and Nabei, however, did not find any significant differences among the three treatment conditions: recasts, clarification requests, and metalinguistic feedback. A recent classroom-based study conducted by Goo (2012) also demonstrated a facilitative role of recasts for L2 learning. The study was designed to compare the effects of recasts and metalinguistic feedback on the acquisition of the English *that*-trace filter. Fifty-four Korean EFL learners participated in one-way information gap activities in a classroom setting and received either recasts or metalinguistic feedback depending on their group affiliation. Results showed that the two experimental groups significantly outperformed the control group in both grammaticality judgment and written production, and that there was no significant difference between the two experimental groups.

Recasts delivered via computer-based technology have also been found to be beneficial—for example, in the learning of the French *passé composé* and *imparfait* (Ayoun, 2001), backshifting of English verbs from the past to the past perfect in indirect reported speech (Sachs & Suh, 2007), Spanish noun-adjective agreement (Sagarra, 2007), English possessive determiners and transitive and intransitive verbs (Trofimovich et al., 2007), and ESL question forms (Petersen, 2010). Petersen's study

is of particular value in that he compared oral, face-to-face recasts with written, computer-mediated recasts and found that the modality of the recasts (i.e., oral vs. written) did not affect the overall results, strengthening claims for the across-the-board utility of recasts. On the noticing of recasts, Lai et al. (2008) found contingent recasts were noticed significantly more often than noncontingent recasts. Using eye-tracking technology, Smith (2010) found evidence of noticing of more than 60% of intensive recasts provided during a synchronous computer-mediated communication activity.

Additional evidence for the facilitative role of recasts (previously illustrated) was reported in our 2007 meta-analysis (Mackey & Goo, 2007), in which findings of interaction studies published between 1990 and 2006 were analyzed. We found large mean effect sizes for recasts on all three posttests ($M = 0.96$, $SD = 1.04$, for immediate posttests; $M = 1.69$, $SD = 1.13$, for short-term delayed posttests; and $M = 1.22$, $SD = 0.85$, for longer-term delayed posttests). Our meta-analysis provided convincing evidence for the positive effects of recasts on language development in the short and longer term, and on the development of a range of grammatical and lexical features (see also Long, 2007, for a review).

In sum, empirical studies to date have demonstrated that recasts play a facilitative role in L2 learning. However, the argument that the jury is still out on recasts began to emerge with Lyster and Ranta's (1997) observation of French immersion classrooms along with Lyster's (1998b) data interpretation, which suggested that recasts may not be effective at promoting L2 development in classroom contexts because of relatively fewer instances of learner responses to recasts (as compared to other forms of feedback). This argument was made on the basis of a relatively small number of repairs following recasts evidenced in Lyster's observations of a specific context—namely, French immersion classrooms. As Lyster put it, “the analysis of recasts as used by the four immersion teachers in the present study leads to the suggestion that, in studies investigating the effects of negative evidence on classroom SLA, recasts themselves may be red herrings” (Lyster, 1998b, p. 74). It is to this case against recasts that we now turn.

THE CASE AGAINST RECASTS

The case against recasts began primarily with classroom-based descriptive research that focused on the extent to which recasts lead to uptake and that casted doubt on the effectiveness of recasts in L2 learning due to reports of relatively fewer instances of learner uptake following recasts (e.g., Lyster, 1998a, 1998b; Lyster & Ranta, 1997; Panova & Lyster, 2002). For example, in Lyster and Ranta's (1997) descriptive study of corrective feedback and learner uptake, they defined uptake as

“a student’s utterance that immediately follows the teacher’s feedback and that constitutes a reaction in some way to the teacher’s intention to draw attention to some aspect of the student’s initial utterance” (p. 49) and observed that even though teachers frequently used recasts to provide corrective feedback, recasts did not lead to as much uptake as other feedback types such as elicitation and metalinguistic feedback; a finding that was reiterated in Lyster’s (1998a, 1998b) subsequent papers (see also Panova & Lyster, 2002, for similar results). Lyster suggested that (a) learners tend to perceive recasts as responding to the content—rather than the incorrectness—of their utterances or as alternative ways of saying the same thing and that (b) the absence of uptake could mean that recasts were noticed much less than other forms of feedback and were therefore developmentally less effective. For example, on the basis of lower rates of uptake or repair following recasts and translations, compared to other feedback moves, Panova and Lyster (2002) suggested that “thus, if recasts and translations are essentially corrective in purpose, there is little evidence that L2 learners in the present study processed them as such” (p. 591). In other words, although Lyster (1998b) acknowledged Mackey and Philp’s (1998) point that “it is unwarranted to equate learner uptake with L2 learning” (p. 74), he asserted that, in his data, less uptake followed recasts than other feedback types and related this to evidence of less noticing of the corrective nature of recasts.

Triggered predominantly by Lyster and his colleagues’ skepticism over the role of recasts in L2 development, a few L2 interaction researchers continued comparing recasts with other types of feedback, such as metalinguistic feedback, clarification requests, elicitations, and repetitions (e.g., Ammar, 2008; Ammar & Spada, 2006; Ellis, 2007; Ellis et al., 2006; Goo, 2012; Loewen & Nabei, 2007; Lyster, 2004; Lyster & Izquierdo, 2009; McDonough, 2007; Nassaji, 2009; Sheen, 2007b; Yang & Lyster, 2010). Overall findings in this line of interaction research are mixed (see Table 1 for a brief summary). Seven studies were interpreted by their authors as providing evidence that recasts may not be as effective as other feedback types, and these consisted of two or more forms of feedback in most cases (Ammar, 2008; Ammar & Spada, 2006; Ellis, 2007; Ellis et al., 2006; Lyster, 2004; Sheen, 2007b; Yang & Lyster, 2010). Four studies showed nonsignificant differences between recasts and other feedback types and were interpreted as such (Goo, 2012; Loewen & Nabei, 2007; Lyster & Izquierdo, 2009; McDonough, 2007); one study showed evidence for the relative efficacy of recasts over other forms of feedback (Nassaji, 2009); and in only a couple of the studies did a recast condition fail to outperform a control condition (e.g., Ellis et al., 2006; Sheen, 2007b). Furthermore and somewhat counterintuitively, Ellis et al. (2006) found that a control group that received no treatment at all (i.e., a testing group with no interactional task) outperformed the recast group on the

delayed posttest for the grammatical items (i.e., on the English regular past tense morpheme *-ed*) on the untimed GJT.

THE CASE AGAINST THE CASE AGAINST RECASTS

Lyster and Ranta (1997) stated that “the recast, the most popular feedback technique, is the least likely to lead to uptake of any kind” (p. 54). However, contrary to such claims, recasts have been shown in some studies to lead to a fair amount of uptake. For example, Oliver (1995) observed that child L2 learners correctly incorporated 35% of the recasts that they received when given the opportunity and when appropriate. Similarly, Braidí (2002) observed that 34.21% of the total recasts provided by NS interlocutors were incorporated when Oliver’s (1995) appropriateness criterion was taken into account. When compared to 18% of grammatical repairs following recasts observed in Lyster and Ranta (1997), Oliver’s and Braidí’s findings, in terms of the number of correct repairs and the coding scheme that they employed, are of particular importance as counterevidence to Lyster and Ranta’s early argument against the use and utility of recasts. Additionally, in her longitudinal study of Japanese as a foreign language learners, Ohta (2000) found that learners produced private speech in response to recasts directed at the whole class or at other students, which implies that L2 learners do in fact produce uptake. In their analysis of form-focused episodes, Ellis et al. (2001) found that 71.6% ($n = 131$) of the total recasts ($N = 183$) provided during interaction led to learner uptake. Of those instances of learner uptake, 76.3% were successful repairs. Additionally, substantial learner uptake has been observed in some instructional settings that involve explicit language-focused exchanges (Oliver & Mackey, 2003; Sheen, 2004).

In other words, as Lyster (see Lyster & Mori, 2006) subsequently recognized if—and to what extent—recasts lead to uptake seems to be a context-dependent question, in which the language and classroom setting, task, and instructor framing are important considerations. Learners’ familiarity with the instructor’s feedback techniques may also be an issue (see Carpenter et al., 2006, for further discussion of this). Thus, we dispute the claim that recasts do not lead to uptake as well as the implication that uptake is a measure of learning. We see these as empirical questions.

Lyster’s emphasis on uptake has generated criticism from a number of SLA researchers (e.g., Gass, 2003; Long, 2007; Mackey & Philp, 1998). As Mackey and Philp (1998) pointed out, uptake is unlikely to be an appropriate direct measure of the effectiveness of recasts, whereas pretest-to-posttest effects serve as the benchmark for the effectiveness of recasts in terms of development or acquisition. Similarly, Ohta’s (2000) analysis of learners’ private speech data led her to argue that “the

Table 1. Recent interaction studies comparing different types of feedback

Study	Participants	Target structure(s)	Design	Dependent variable measures	Reported results
Lyster (2004)	179 fifth graders in French immersion school	French grammatical gender	<ol style="list-style-type: none"> 1. FFI + recasts 2. FFI + prompts 3. FFI only 4. Control Treatment: Noticing activities, awareness activities, practice activities, and various communicative activities	Binary-choice test, text-completion test, object-identification test, picture-description test	<ol style="list-style-type: none"> 1. FFI-prompt group performed significantly better than FFI-recast group on written posttests (binary-choice and text-completion) 2. Evidence in favor of FFI-prompts on oral tests
Ammar and Spada (2006)	64 sixth grade intensive ESL learners in Montreal area	Third-person possessive determiners (his and her)	<ol style="list-style-type: none"> 1. Recast group 2. Prompt group 3. Control Treatment: One instructional session and 11 practice sessions (reported only as communicative activities) over 4 weeks (30–45 min for each session)	Passage correction, picture description	<ol style="list-style-type: none"> 1. Overall, prompt group performed better than recast group (differences were significant between the two experimental groups except on immediate posttest in the oral picture-description task) 2. Low-proficiency learners benefited more from prompts than from recasts 3. High-proficiency learners benefited equally from both types of feedback

<p>Ellis, Loewen, and Erlam (2006)</p>	<p>34 ESL students in private language school in New Zealand</p>	<p>English past tense morpheme <i>-ed</i></p> <p>1. Recast group 2. Metalinguistic explanation 3. Control</p> <p>Treatment: Two half-hour storytelling activities over 2 days</p>	<p>Oral imitation test, untimed GJT, metalinguistic test</p>	<p>1. No between-group difference on immediate posttests 2. On delayed posttests, metalinguistic group performed significantly better than recast and control groups on oral imitation and untimed GJTs 3. Metalinguistic group was significantly better at generalizing past tense morpheme to new items than recast group (only in delayed posttests)</p>
<p>Ellis (2007)</p>	<p>34 adult ESL learners in private language institute in New Zealand</p>	<p>Past tense <i>-ed</i> and comparative <i>-er</i></p> <p>1. Recast group 2. Metalinguistic group 3. Control</p> <p>Treatment: Two half-hour storytelling activities over 2 days for the past tense <i>-ed</i> and description activities for the comparative</p>	<p>Oral imitation test, untimed GJT, metalinguistic knowledge test</p>	<p>1. Recasts did not lead to differential effects on acquisition of the two structures 2. Metalinguistic feedback had differential effect on acquisition of the two target structures.</p>

Continued

Table 1. Continued

Study	Participants	Target structure(s)	Design	Dependent variable measures	Reported results
Loewen and Nabei (2007)	66 Japanese university EFL learners	English question formation	<ol style="list-style-type: none"> 1. Metalinguistic group 2. Recast group 3. Elicitation group 4. Interaction-with-no-feedback group 5. Control Treatment: One spot-the-difference task and one guess-the-storyline task (30 min for the entire treatment session)	GJT (timed and untimed), oral production test	<ol style="list-style-type: none"> 1. No significant between-group difference in the untimed GJT and the oral production test 2. Significant difference between the feedback groups and no-feedback group (and control group) in the timed one, but no significant differences among the feedback groups
McDonough (2007)	74 Thai first-year university EFL learners	Simple past activity verbs	<ol style="list-style-type: none"> 1. Recast group 2. Clarification request group 3. Interaction with no feedback group Treatment: Two-way information exchanges and information gap tasks	Oral production	<ol style="list-style-type: none"> 1. Both clarification request and recast groups facilitated emergence of past tense activity verbs 2. No advantage for clarification requests over recasts

Sheen (2007b)	80 adult ESL learners in community college American language program	English articles	<ol style="list-style-type: none"> 1. Recast group 2. Metalinguistic correction (recast followed by metalinguistic information) 3. Control <p>Treatment: Two narrative tasks in two sessions (30–40 min for each task)</p>	Speeded dictation, writing test, error correction	<ol style="list-style-type: none"> 1. Metalinguistic group outperformed both the recast and control groups 2. For metalinguistic group, significant positive relationship between immediate and delayed gain scores on the one hand and their language analytical ability on the other 3. Immediate gains in metalinguistic group were also positively correlated with learners' attitude toward error/error correction. 4. For recast group, no significant relationship between the gain scores and learners' ability for language analyses, nor their reported feelings about error correction
---------------	--	------------------	---	---	---

Continued

Table 1. Continued

Study	Participants	Target structure(s)	Design	Dependent variable measures	Reported results
Ammar (2008)	64 sixth grade intensive ESL learners in Montreal area	Third-person possessive determiners (his and her)	<p>1. Recast group</p> <p>2. Prompt group</p> <p>3. Control</p> <p>Treatment: One instructional session and 11 practice sessions (reported only as communicative activities) over 4 weeks. (30–45 min for each session)</p>	<p>Oral picture description (developmental stages),</p> <p>computerized fill-in-the-blank (reaction time)</p>	<p>1. All learners in prompt group moved up to post-emergence stages 5–7, whereas 65% of recast group learners did</p> <p>2. All prompt group learners were able to stay at the post-emergence stage, but one-third of learners in recast and control groups failed to maintain their stage increase achieved on the immediate posttest</p> <p>3. Prompt group was significantly faster than recast group in performing a computerized fill-in-the-blank task</p>

Lyster and Izquierdo (2009)	25 university French learners	French grammatical gender	<ol style="list-style-type: none"> 1. Recast group 2. Prompt group Treatment: A 3-hr form-focused instruction over a 2-week period and two 30-min feedback sessions	Object-identification task, picture-description task, binary-choice and reaction time measures	<ol style="list-style-type: none"> 1. Significant pre-to-post changes in all measures 2. No significant difference between recasts and prompts in any measures
Nassaji (2009) ^a	42 adult ESL learners	Multiple targets	<ol style="list-style-type: none"> 1. Recast group 2. Elicitation group Treatment: One picture-sequencing task for 10–15 min	Identifications and corrections of ill-formed utterances	<ol style="list-style-type: none"> 1. Recasts led to more successful identifications and corrections of ill-formed utterances 2. Explicit forms in both types of feedback led to more successful identifications and corrections of erroneous utterances

Continued

Table 1. Continued

Study	Participants	Target structure(s)	Design	Dependent variable measures	Reported results
Yang and Lyster (2010)	72 university English learners in China	English past tense	<p>1. Recast group</p> <p>2. Prompt group</p> <p>3. Control group</p> <p>Treatment: Two dictogloss tasks, one question-and-answer activity, one picture-cued narrative activity (approximately 2 hr over 2-week period)</p>	<p>Oral production (story-retelling task), written production (story-composition task)</p>	<p>1. No significant group effect on the oral production of irregular past tense</p> <p>2. Significant pre-to-post changes in accurate use of irregular past tense for all three groups on oral production (significant pre-to-delayed-post gains for prompt and control groups, but not recast group)</p> <p>3. Only prompt group made significant pre-to-post and pre-to-delayed-post gains in accurate use of regular past tense</p> <p>4. No significant between-group difference on written production of irregular past tense</p> <p>5. Significant pre-to-post and pre-to-delayed-post improvements for both experimental groups (irregular past tense)</p> <p>6. No significant group effect on written production of regular past tense</p>

- 7. Significant pre-to-post gains on written production of regular past tense for all three groups
- 8. Significant pre-to-delayed-post gains on written production of regular past tense only for prompt group

Goo (2012)	54 Korean EFL learners attending Korean university	English <i>that</i> -trace filter	<ul style="list-style-type: none"> 1. Recast group 2. Metalinguistic group 3. Control group Treatment: Two one-way information gap activities	GJT and written-production test	<ul style="list-style-type: none"> 1. Both experimental groups performed significantly better than control group on both outcome measures 2. No significant differences between the recast and metalinguistic groups in either measure
------------	--	-----------------------------------	--	---------------------------------	--

^a Nassaji's (2009) results were based on a small set of descriptive data (percentage), and no inferential statistical analyses were performed.

efficacy of recasts should not be doubted based on the presence or absence of an overt oral response” (p. 66). In their study on the relationship between responses to recasts and ESL question development, McDonough and Mackey (2006) found that primed production—operationalized as “a learner’s use of the question form provided in the recast to ask a new question” (p. 705)—was a statistically significant predictor of learner performance on English question formation, but immediate repetition, one type of uptake in Lyster and Ranta’s sense, was not. Loewen and Philp (2006) also cautioned against the inappropriateness of considering uptake as a measure of L2 noticing or learning. Their data showed no statistically significant difference among three feedback types (i.e., recasts, inform, and elicitations) in learner accuracy on the tailor-made immediate and delayed posttests, despite the fact that elicitations led to a much higher rate of successful uptake (83.1%) than did recasts (59.6%) during classroom interaction. Lyster and Izquierdo’s (2009) study also showed that recasts appeared to be as effective as prompts (i.e., clarification requests and repetitions) at promoting the development of French gender marking, even though opportunities to produce immediate uptake were provided for the prompt condition but not allowed for the recast condition. This finding also suggests that, even without uptake or repairs, recasts can facilitate L2 learning. Long (2007) argued that Lyster’s claims and concerns regarding the utility of recasts were simply unwarranted and pointed to supportive evidence for the positive value and potential utility of recasts (p. 97). Long also questioned the validity of immediate uptake as an indication of acquisition in relation to the issue of deployment versus acquisition, critiqued the coarsely grained definition of uptake, and explicated some questionable interpretations of classroom data made in Lyster’s and Lyster and Ranta’s studies. Finally, Long pointed to the benefits of the unobtrusiveness of recasts when compared with the obtrusiveness of explicit feedback (Long, 2007).

METHODOLOGICAL ISSUES

Having considered the issue of uptake and its meaning, we next turn to methodological issues and question many experimental and quasi-experimental studies (summarized in Table 1) that investigated the effectiveness of different types of corrective feedback (i.e., recasts vs. other types of feedback). In what follows, we describe five issues in detail—(a) modified output opportunities, (b) single-versus-multiple comparisons, (c) form-focused instruction, (d) prior knowledge, and (e) out-of-experiment exposure—as we examine the case against recasts. We have also provided a checklist of these issues for all the studies referred to in the current study (case-for and case-against papers) in Table 2.

Table 2. Cited studies: Methodological checklist

Study	Modified output	Simple vs. multiple comparisons	Form-focused instruction	Prior knowledge ^{c,d}
Ammar (2008)	√	√	√	√
Ammar and Spada (2006)	√	√	√	√
Ayoun (2001)		N/A		√
Doughty and Varela (1998)	√	N/A		(√)
Ellis (2007)	√	√		√
Ellis, Loewen, and Erlam (2006)	√	√		√
Goo (2012)				
Han (2002)	√	N/A		
Ishida (2004)	√	N/A		
Iwashita (2003)	√	N/A		
Leeman (2003)		N/A		
Loewen and Nabei (2007)	√			√
Long, Inagaki, and Ortega (1998)		N/A		
Lyster (2004)	√	√	√	√
Lyster and Izquierdo (2009) ^a	√	√	√	√
Mackey and Philp (1998)	√	N/A		
McDonough (2007)	√			
McDonough and Mackey (2006) ^b	(√)	N/A		
Nassaji (2009)	√	√		
Sagarra (2007)		N/A		
Sheen (2007b)	√	√		
Yang and Lyster (2010)	√	√		√

Note. A check mark indicates a given study (row) involves a corresponding methodological issue (column). N/A means “not applicable.”

^a Lyster and Izquierdo (2009) controlled for modified output opportunities in such a way that the recast group was not allowed to produce modified output, but the prompt group was.

^b McDonough and Mackey (2006) investigated the impact of learner responses to recasts on L2 development. Modified output opportunities can be seen as partially controlled in their study.

^c If learners obtained higher than 60% accuracy scores on a pretest, it can be argued that prior knowledge was under controlled. Doughty and Varela (1998) used medians, and their study was conducted in classrooms; thus, prior knowledge may not have been strictly controlled.

^d Our final criterion, out-of-experiment exposure, is difficult to confirm one way or the other unless indicated in research reports. Thus, we cannot be sure that out-of-experiment exposure did or did not occur in any of the other studies because, unlike in Ellis (2007) and Ellis et al. (2006), whose reports indicated it occurred, it was not discussed.

Modified Output Opportunities

As previously discussed, modified output is optional following recasts, but other forms of feedback, such as clarification requests, metalinguistic feedback, repetitions, and elicitations either indirectly or directly push learners to produce modified output. This modified output has been argued to lead learners to focus on the linguistic form in which the feedback was provided. More specifically, as Swain (1985, 1993, 1995, 2005) argued, modified output provides potential opportunities for learners to notice their linguistic problems and to test hypotheses about the target forms, which may in turn contribute to the development of L2 accuracy. Empirical evidence for these claims has been found in a number of studies (e.g., McDonough, 2005; McDonough & Mackey, 2006; Nobuyoshi & Ellis, 1993; Takashima & Ellis, 1999).

To test the benefits of one type of feedback over another, it would be ideal to carefully control production of modified output following corrective feedback—regardless of the type of feedback—such that learners produce it or are prompted to move on without producing it (as was the case in Goo, 2012; in Leeman, 2003; and in Long et al.'s [1998] recasts vs. model study). This presents challenges, however, for the treatment task to prompt or constrain production in a way that is discourse appropriate. In short, if opportunities for modified output are not controlled, learners who receive corrective feedback with modified output opportunities (e.g., clarification requests, metalinguistic feedback, elicitations, and repetitions) could be seen to be in a relatively advantageous position, all other things being equal, compared to those learners who receive feedback that rarely leads to modified output. The discourse difficulty of prohibiting modified output coupled with the idea that the opportunity for modified output is an integral part of some types of corrective feedback are primary reasons why some researchers, ourselves included, have come to believe the recasts-versus-other-types-of-feedback comparison is one of apples and oranges. Thus, it is less helpful in the quest to understand interaction and its effects on learning than in-depth examinations of one particular feedback type or another (but not in competition with one another) because a wide range of feedback has been shown to be developmentally effective.

In other words, as alluded to earlier, many recast researchers take the perspective that interaction provides different routes to learning, and recasts, which make relatively few participatory demands on the learner, provide just one of the routes. Recasts are one of the tools that NSs, teachers, and other interlocutors have at their disposal, and they are interesting to study because a preponderance of evidence has shown that they work, which leads us to question how, when, and why they work. Various other types of interaction also lead to learning,

including prompts with output opportunities. We believe that the interaction research agenda would be well served by (a) investigating the conditions and processes by which the different feedback types work and (b) focusing on helpful constructs such as linguistic form, interlocutor effects, and so on.

Returning to methodology, modified output opportunities were not controlled for in most of the studies comparing recasts with other forms of feedback (Ammar, 2008; Ammar & Spada, 2006; Ellis, 2007; Ellis et al., 2006; Loewen & Nabei, 2007; Lyster, 2004; McDonough, 2007; Nassaji, 2009; Sheen, 2007b; Yang & Lyster, 2010). In Lyster and Izquierdo's (2009) study, for example, modified output production was encouraged for those receiving prompts but not allowed for those receiving recasts. Researchers who have conducted classroom-based, quasi-experimental studies on prompts versus recasts (e.g., Ammar, 2008; Ammar & Spada, 2006; Lyster, 2004; Yang & Lyster, 2010) have argued that because controlling for modified output production makes the treatment somewhat artificial, it thus compromises ecological validity. However, this is clearly a trade-off. Maintaining ecological validity at the cost of introducing a potentially confounding variable compromises internal validity. As many methodologists have pointed out, there can be no external validity without internal validity. We believe that in (quasi-) experimental studies, internal validity should precede ecological validity because claims can still be made with respect to the independent variables examined in an experimental study even without ecological validity if internal validity has been maintained, whereas the converse is not true. In other words, researchers should question the legitimacy of research lacking internal validity no matter how ecologically valid a given study context may be. Of course, it would be ideal for research to be designed in such a way that satisfies both internal and ecological validity, but when working in authentic, real-world classrooms, we do not always have that luxury. This is one reason why much of the recast research has been carried out in laboratories. Ultimately, though, we do not believe that anyone—case-against and case-for researchers included—is arguing that teachers should drop or exclusively use one feedback type over another. If the goal is to advance our understanding of L2 cognitive processes, again, we believe comparing unique feedback types, given their inherent differences, is an apples-versus-oranges comparison.¹

More on Apples and Oranges: Single-versus-Multiple Comparisons

In studies providing evidence against the utility of recasts, recasts have been compared to another treatment condition that often involves

more than one type of feedback, which, in essence, compares one variable with multiple variables (e.g., Ammar, 2008; Ammar & Spada, 2006; Ellis, 2007; Ellis et al., 2006; Lyster, 2004; Lyster & Izquierdo, 2009; Sheen, 2007b; Yang & Lyster, 2010). Three versions of this single-versus-multiple comparison exist. First, a recast condition operationalized as only recasts has been compared with a prompt condition, where this prompt contained multiple types of feedback (e.g., metalinguistic feedback, elicitations, repetitions, and clarification requests). For example, in both Ammar (2008) and Ammar and Spada (2006), learners in the prompt condition received metalinguistic feedback, elicitations, or repetitions at their NS teacher's discretion. Prompts in Lyster's (2004) and Yang and Lyster's (2010) studies additionally included clarification requests. One argument against this methodology is that learners receiving multiple types of feedback have more opportunities to benefit from contextually appropriate feedback than those exposed to only one type of feedback during the entire task.

A second single-versus-multiple comparison involves comparing a recast condition to another condition involving one type of feedback move first followed by another if the learner's response to the first feedback move was still problematic. For instance, Lyster and Izquierdo (2009) utilized repetitions as well as clarification requests in the nonrecast condition (i.e., prompt condition). A repetition was used when modified output immediately following a clarification request was still nontargetlike, which, in effect, made it a double feedback move. It seems likely that receiving two—rather than one—forms of corrective feedback, and following up when targetlike modified output was not produced by making another attempt at eliciting it, would render the feedback more salient and successful than simply using one form of feedback to indicate the ungrammaticality of learner utterances. In other words, it would seem that recasts were at a distinct disadvantage in this study.

Lastly, in a number of studies, two different types of feedback on the identical nontargetlike utterance were provided in a single feedback turn under a prompt condition, which enhanced focal attention to information delivered through the feedback, whereas only a single recast was provided in each feedback attempt under a recast condition, thus hindering those who received recasts (Ellis, 2007; Ellis et al., 2006; Sheen, 2007b). For example, in Ellis et al.'s (2006) study, also reported in Ellis (2007), the recast condition was operationalized as only recasts, whereas a metalinguistic feedback condition was operationalized as a combination of a repetition and metalinguistic information. Sheen (2007b) also operationalized her metalinguistic condition as a recast immediately followed by metalinguistic information in a single turn. Despite these imbalances in the groups, Ellis (2007), Ellis et al. (2006), and Sheen (2007b) all interpreted their results as suggestive of the fact that

metalinguistic feedback is more beneficial to L2 learning than recasts but ignored the fact that (a) recasts were operationalized monolithically, whereas the comparison was not; (b) two types of corrective feedback on one erroneous utterance were provided in a single turn, whereas only a single recast was provided for the recast group; and (c) the recast was neither preceded nor followed by any other type of feedback. However, it may be the case that the improved performance of the metalinguistic group was an outcome of an enhanced combined effect (or a synergy) for repetitions and metalinguistic feedback in Ellis (2007) and Ellis et al. (2006), and for recasts and metalinguistic feedback in Sheen (2007b), but not metalinguistic feedback or information alone. As Sheen (2007b) acknowledged:

simply providing learners with metalinguistic comments may “prime” the learners but they need time to use the explicit information they obtain from the feedback to acquire the feature. On the other hand, providing learners with the correct form together with metalinguistic information affords both positive and negative evidence, which together appear to be sufficient to produce an immediate effect. (p. 319)

There is no evidence that metalinguistic feedback alone was more effective than recasts at facilitating the learning of the English articles in Sheen’s (2007b) study because there was no metalinguistic-feedback-only group. As such, there appears to be no clear-cut reason to believe that Sheen’s study finds recasts are ineffective, and hence it should not be interpreted as such.

Form-Focused Instruction

In a number of recast-versus-prompt studies, an instructional session was also included regardless of the treatment condition as an integral part of the experimental treatments (e.g., Ammar, 2008; Ammar & Spada, 2006; Lyster, 2004; Lyster & Izquierdo, 2009). That is, all learners in the treatment conditions were provided with instruction on the linguistic targets at an early stage of the experiment because, to respond to teacher prompts (e.g., clarification requests), learners needed some prior knowledge of the structure (see Takashima & Ellis, 1999, for such an argument). Because the form is provided in a recast, a recast does not necessarily require prior knowledge of the target form; however, because recasts are contingent on error, the form must be obligatory but missing for the recast to provide it. Thus, the task has to create contexts for the form to occur. It could be argued that a research design with form-focused instruction included may be biased to some extent

toward the prompt condition because not only is the starting point likely to influence overall results but it also makes the interpretation of results regarding the exact cause of learning somewhat difficult. For instance, Lyster (2004) reported that form-focused instruction worked better when combined with prompts than when it was combined with recasts. Because both feedback conditions (i.e., recasts and prompts) involved form-focused instruction, it seems questionable to claim that the results are due to the feedback treatments only. Any differential impact of prompts and recasts may not be due wholly, or at all, to the difference between prompts and recasts but may be due, to some extent, to the moderating role of the form-focused instruction. In this sense, we do not believe that Lyster's (2004) findings can be viewed as unequivocally suggesting that prompts are more efficacious than recasts because his study provides no (direct) evidence for this claim.

For the same reason, it is by no means clear whether any improved performance witnessed in Ammar (2008) and Ammar and Spada (2006) is attributable to structured form-focused instruction administered prior to the actual feedback treatment, to corrective feedback provided during the treatment, or to the combination of the two. If learners already know the correct form of a target prior to feedback treatments through form-focused instruction, acquisition may not be easily distinguishable from deployment (see Long, 2007; Long et al., 1998, for discussion). This issue is quite complex when we take into account the fact that different types of corrective feedback, including recasts, may contribute to increased control of a form that is partially known, which some researchers would consider to be part of development. Space precludes a full discussion of these arguments, but they are interesting ones that lead to compelling empirical questions for future research to pursue. Testing the impact of recasts or other forms of feedback on unknown and partially known forms would be worthwhile endeavors if robust designs were used.

It should also be noted that recast researchers do not always select only learners who scored zero but often those who scored low on the pretest, and this is routinely done in research on recasts (e.g., Han, 2002; Iwashita, 2003; Leeman, 2003; Long et al., 1998, among others). In contrast, the provision of form-focused instruction to learners with zero knowledge has the advantage of controlling for the type and amount of prior knowledge. In some cases, this is an acceptable compromise as long as it is appropriately pointed out in the interpretation of the results.

As previously noted, form-focused instruction is provided on the basis of the assumption that learners need to have some basic knowledge of a given target if they are to respond to prompts (Ammar, 2008; Ammar & Spada, 2006; Lyster, 2004; Lyster & Izquierdo, 2009). However, metalinguistic feedback, which is used along with other types of prompts in the

prompt condition (e.g., Ammar, 2008; Ammar & Spada, 2006; Lyster, 2004; Yang & Lyster, 2010), can actually provide relevant grammatical information on new target(s) that have not been learned. Given this, the rationale provided by the researchers for the necessity of administering a form-focused instructional session is not particularly convincing (for example, Ammar's [2008] rationale that "given that prompts cannot be used to elicit forms students do not know already [Lyster, 2004], an instruction component in which the PD [possessive determiner] rule was explained and practiced was deemed necessary" [p. 191]). Put another way, it is not unreasonable to compare the efficacy of recasts and prompts—without the provision of formal instruction—as long as the prompt condition contains metalinguistic feedback because grammatical information can be provided through that metalinguistic feedback. This point seems not to have been taken into account in most of the studies conducted by Lyster, Ammar, and their colleagues (e.g., Ammar, 2008; Ammar & Spada, 2006; Lyster, 2004; Yang & Lyster, 2010). Future research could carry out studies of this nature (without providing instruction), could examine L2 targets to which learners have never been exposed, or could incorporate tailor-made posttests. Again, we believe that this sort of study, which takes prior knowledge into consideration and balances it across conditions, would be a worthwhile undertaking. It is to this topic that we turn next.

Prior Knowledge

Prior knowledge has not always been controlled for in some previous studies (e.g., Ammar & Spada, 2006; Ellis, 2007; Ellis et al., 2006; Loewen & Nabei, 2007; Yang & Lyster, 2010). For example, in Ellis's (2007) report of his findings, we see that there was room for the metalinguistic feedback group (pretest scores on an untimed GJT, $M = 0.689$, $SD = 0.265$) to further develop explicit knowledge of the comparative, but the group already had considerable explicit knowledge of the past tense morpheme (pretest scores: $M = 0.844$, $SD = 0.108$). In short, any significant pre-to-post development of explicit knowledge of the past tense morpheme was unlikely for the metalinguistic feedback group. Additionally, the recast group already possessed considerable explicit knowledge of both target morphemes prior to the treatment, as evidenced in their performance on the pretest ($M = 0.854$, $SD = 0.129$, and $M = 0.855$, $SD = 0.159$, for the past tense *-ed* and the comparative *-er*, respectively, in an untimed GJT). In other words, looking at these scores, it is unsurprising that recasts did not lead to any differential effects on the acquisition of the two target forms because there was no room for improvement (i.e., a ceiling effect). Thus, the results of Ellis and his colleagues' studies (e.g., Ellis, 2007; Ellis

et al., 2006) must be interpreted accordingly: Prior knowledge was an intervening variable.

The same issue is found in Loewen and Nabei (2007), in which it was reported that learner accuracy on the untimed grammaticality judgment pretest on English question formation ranged from 70% to 84%. Because they conducted an ANCOVA due to the significant between-group difference in learner performance on the pretest, no results were reported on pre- to posttest gains on the untimed GJTs. However, as most learners already had a fair amount of target language knowledge before the treatments, it is unsurprising that no significant between-groups differences were found in the ANCOVA.

An additional issue is that some researchers employed ANCOVAs even though there was no covariate that should have been taken into account. Ammar and Spada (2006) used ANCOVAs for all their statistical analyses with learner performance on the pretest as a covariate. However, they found no significant between-group differences and did not report whether the results of the pretest correlated with or predicted in a regression analysis those of the posttests, as either is a prerequisite for running an ANCOVA. If there are no between-groups differences on the pretest, and if there is little correlation between pre- and posttest scores, then mixed ANOVAs are likely to be a standard statistical choice with time as a within-subjects variable and group as a between-subjects variable. The selection of a separate ANCOVA on learner performance on each separate test disburdened Ammar and Spada's obligation to expound or comment on the improved performance of the control group on the oral tests ($M = 47.5$, $SD = 27.4$, on the pretest; $M = 62.9$, $SD = 19.1$, on the immediate posttest; and $M = 60.9$, $SD = 16.9$, on the delayed posttest; all unexpectedly high increases). Likewise, their choice of test meant they did not have to comment on the almost nonexistent effect of recasts observed on the written tests ($M = 11.1$, $SD = 3.5$, on the pretest; $M = 11.8$, $SD = 3.8$, on the immediate posttest; and $M = 12.3$, $SD = 4.0$, on the delayed posttest), which seems to warrant some explanation given the students' strong pre- to posttest (and pre- to delayed posttest) improvement on the oral tests.

In addition to the issue with their choice of statistical tests, Ammar and Spada's claim that learners with some prior knowledge, categorized as high-proficiency learners in the study, benefited more or less equally from recasts and prompts seems to us to be problematic. They note that they did not find any significant differences among the three groups (in the case of the oral production task) or between the two experimental groups (in the case of the passage correction task). What the results of the oral production measure actually indicate is not that learners with prior knowledge can benefit equally from both types of feedback but that corrective feedback for those with some prior

knowledge may not be as effective as suggested in previous research. The nonsignificant between-group differences on the oral production measure also imply that for those with prior knowledge, as long as they participate in interactional activities, learning can take place. Their study, however, did not provide any evidence for this possibility due to the absence of a pure control group—that is, a group with no communicative activities. It may be that learners in the control group had enough prior knowledge to take advantage of interactional treatments as practice opportunities, even without corrective feedback (as illustrated in their Figure 4, p. 560). Thus, prior knowledge in their study can be argued to be a confounding factor that renders distinctions between the experimental groups and the control group problematic. Regarding the nonsignificant difference between the two experimental groups in their performance on the passage correction task, contrary to their claim, the results indicate that neither feedback move facilitated the learning of the target forms (i.e., English third-person possessive determiners, *his* and *her*). Not only do their high pretest scores (i.e., more than 65% in accuracy on the pretest for both groups) delimit the range of possible improvement but their pre- to posttest gains are also rather negligible (see Table B2, p. 574). Therefore, in our view, there is no statistically significant evidence for Ammar and Spada's claim that "prompts and recasts were equally effective for high-proficiency learners" (p. 562).

Yang and Lyster's (2010) study compared three groups (i.e., recast, prompt, and control), with the researchers reporting a large pre- to posttest effect size for the control group ($d = 0.94$) on the oral production of irregular past tense. A large pre- to posttest effect size ($d = 0.82$) was also found for the control group on the written production of regular past tense. We would argue that these high pre- to posttest developments by the control group may stem from the fact that prior knowledge was not controlled for ($M = 55.56$, $SD = 20.85$, in the oral production pretest on irregular past tense; and $M = 67.58$, $SD = 26.36$, in the written production pretest on regular past tense). As noted earlier in relation to Ammar and Spada's study, interactional activities, specifically designed to elicit targets for the experiment, might have evoked learners' prior knowledge of the targets and might have provided practice opportunities for learners in the control group even without corrective feedback. Therefore, it seems to us that, as discussed in Long (2007), successful deployment of existing knowledge on the targets, rather than the acquisition of new targets, might be a better indicator of what actually occurred in their study. It may also be speculated that out-of-experiment exposure to the target forms, which we discuss in more detail in the next section, might have contributed to the pre- to posttest developments by their control

group. Regardless of exactly what resulted in the unexpected outcomes, the large pre- to posttest developments by the control group appear to jeopardize the validity of their study and thus the reliability of their overall findings.

Out-of-Experiment Exposure

We now turn to a discussion of out-of-experiment exposure to the linguistic form targeted. With regard to the metalinguistic group's improved performance on the oral imitation posttest 2 with ungrammatical sentences, compared to the oral imitation posttest 1 on the past tense *-ed*, Ellis (2007) suggested that "the effect of the metalinguistic feedback, then, may have been simply that of 'freshening up' their explicit knowledge of this structure, enabling them to attend more closely to the instances of past tense *-ed* in the input they were exposed to between posttest 1 and posttest 2, which were likely to have been plentiful" (p. 359). What this implies is that the participants were exposed to input regarding the past tense *-ed* outside the treatment setting because no treatment was provided between the two posttests. This in turn suggests that the improved performance of the metalinguistic group might not have resulted from the metalinguistic feedback that they received during the treatment, and it thus calls into question the validity of the overall results. Although we understand the argument that metalinguistic feedback benefited L2 learning by becoming a useful reference when the participants encountered aural input or needed to produce output containing the target, this out-of-experiment exposure is still a threat to the internal validity of the study. From the fact that Ellis (2007) and Ellis et al. (2006) shared the same data set on the past tense morpheme *-ed* we can conclude that Ellis et al.'s somewhat unexpected findings—namely, that there were no significant differences between recasts and metalinguistic feedback on any immediate posttest measures but that learners receiving metalinguistic feedback significantly outperformed those receiving recasts on most of the delayed posttest measures—may mainly be due to this out-of-experiment exposure. This kind of history effect is a threat to the internal validity of empirical studies conducted over a long period of time involving multiple treatment sessions and two or more posttests, especially in an ESL (as opposed to EFL) setting. The effect is somewhat difficult to control for altogether; nevertheless, it should be carefully considered at the initial stage of designing a study so that it can be minimized as much as possible. One way to achieve this, for instance, is to select a target feature to which learners have limited exposure in their daily life. In this respect, foreign language contexts provide a more favorable environment for

experimental research in a methodological sense compared to second language contexts. It should be noted that, as in our note to Table 2, without explicit statements in the research report like those made by Ellis, extraexperimental exposure is a potential threat to a wide range of SLA research.

SUMMARY OF PROBLEMS

We have argued that existing empirical studies that have compared recasts and other types of feedback fall short of providing clear and convincing evidence of the relative inefficacy of recasts due to methodological limitations. In summary, (a) modified output opportunities were not controlled for in most of the studies designed to investigate the relative effect of one type of feedback over another on L2 development; (b) recasts have been compared to another treatment condition (e.g., prompts) involving two or more feedback types; (c) form-focused instruction has been included as a part of the experimental treatment in some recast-versus-prompt studies; (d) prior knowledge has not been controlled for in some studies, making any significant improvement unlikely; (e) out-of-experiment exposure to a linguistic form targeted in a given study can threaten validity.

Some of these limitations (e.g., out-of-experiment exposure and prior knowledge) may also be applied to studies of recasts. Ongoing work in the field—including our own—is beginning to address these thorny issues. We have provided some initial steps toward this in Table 2. In this article, we certainly do not intend to imply that the case for recast research has no room for methodological improvement. Rather, we would point out that the case against recasts seems to be based on a shaky foundation: a small number of studies that were mostly carried out in one or two specific contexts and that suffer from some serious methodological limitations. The case for recasts consists of a much larger sample of studies, in a wider range of contexts, and addresses a number of different variables. It is clear that the debate is a complex one.

Although space constraints preclude a discussion, it should be noted that the case against recasts researchers have been primarily motivated by the nuances of pedagogical concerns. Many of the recast researchers are often concerned with cognitive mechanisms as well as classroom applications and sometimes prioritize the former over the latter. Both, obviously, are legitimate directions for research. As such, a number of comments have been made in the literature on the topic of theoretical approaches, empirical testing, and the directions of the field (Hulstijn, 2012).

CONCLUSION

We believe that most of the findings of previous empirical studies that compare recasts with prompts or other feedback types, which express negative reservations about recasts and claim that the jury is out as to their usefulness, should be interpreted with care. We do not believe that any study to date has provided clear-cut, convincing evidence that recasts are significantly less effective than metalinguistic feedback or prompts due to the methodological limitations outlined in the previous sections. In other words, the jury should not be out on recasts in classroom contexts because the evidence is not in. We do not believe that it is an open question as to whether or not recasts work; we believe that question is settled. However, methodologically sound research is needed to explore the effectiveness of different types of feedback in the acquisition of various linguistic forms but not necessarily in comparison with one another, unless convincing arguments are made for why it makes sense to compare such different constructs. If sound arguments are made (and it is very likely that, from classroom perspectives, they can be), and, crucially, if they are followed up with sound designs, evidence can be gathered. However, the burden of proof is still in the hands of researchers (a) to devise more rigorous studies that meet solid methodological standards and expectations (as well as readers to review empirical studies critically), and (b) to not repeat claims without considering the methodologies of the studies on which these claims are based.

What seems to be problematic to us is that the case-against researchers have compared recasts and prompts under conditions that are more or less advantageous to one type of feedback but not to the other without a clear understanding of the conditions under which each feedback type would be found to be effective. We believe that for those who want to carry out comparisons, it is most constructive to identify the exact factors that enhance or reduce the effectiveness of recasts and prompts and then carry out tests that are theoretically and empirically sound.

Finally, it bears repeating that, in our view, recasts and prompts seem to function differently in their contribution to L2 development. Modified output naturally follows the prompts but not the recasts. Participatory demands are different. Prior knowledge requirements are likely to be different. It is fruitless to investigate comparative effects of recasts versus prompts (and metalinguistic feedback) when, in fact, recasts and other forms of feedback seem likely to play unique roles in L2 learning. Additionally, it is also possible that recasts and prompts may work synergistically to effect positive changes in L2 development. For example, prompts—like clarification requests, metalinguistic feedback, and elicitations—may help remind learners about information provided

earlier through recasts. For these reasons, rather than making simple comparisons, research time would be better spent exploring how the effectiveness of corrective feedback interacts with such mediating factors as noticing and attention (especially via recent eye-tracking techniques, see Smith & Renaud, in press), individual differences in cognitive capacities (e.g., short-term memory, working memory, see Goo, 2012), social factors (e.g., learning setting, social status of interlocutors, see Philp & Mackey, 2010), and the type of target (e.g., whether recasts and prompts facilitate the acquisition of L2 pragmatics and phonology, areas rarely examined in interaction research). In other words, using more triangulated approaches to research on corrective feedback that employ various and rigorous methodological designs is the way to move forward. However, for those who do find it worthwhile to do comparative investigations of feedback types, as well as those who do not, methodology is always going to be the heart of the matter.

Received 8 August 2011

Accepted 4 November 2011

Final Version Received 27 July 2012

NOTES

1. An anonymous reviewer pointed out that much of the comparative research compares not only the input-output dimension but also the implicit-explicit dimension of corrective feedback. For example, recasts and metalinguistic feedback differ in terms of not only their likelihood of eliciting modified output from the learner but also their explicitness. The confounding of the input-output and implicit-explicit variables also threatens the validity of comparative research. Space precludes a more complete discussion of this point, but we agree with the reviewer that this is more grist for the apples-oranges mill.

REFERENCES

- Ammar, A. (2008). Prompts and recasts: Differential effects on second language morphosyntax. *Language Teaching Research*, 12, 183–210.
- Ammar, A., & Spada, N. (2006). One size fits all? Recasts, prompts, and L2 learning. *Studies in Second Language Acquisition*, 28, 543–574.
- Ayoun, D. (2001). The role of negative and positive feedback in the second language acquisition of the *passé composé* and *imparfait*. *Modern Language Journal*, 85, 226–243.
- Baker, N. D., & Nelson, K. E. (1984). Recasting and related conversational techniques for triggering syntactic advances by young children. *First Language*, 5, 3–22.
- Bohannon, J. N., & Stanowicz, L. (1988). The issue of negative evidence: Adult responses to children's language errors. *Developmental Psychology*, 24, 684–689.
- Braidi, S. M. (2002). Reexamining the role of recasts in native-speaker/nonnative-speaker interactions. *Language Learning*, 52, 1–42.
- Carpenter, H., Jeon, K.-S., MacGregor, D., & Mackey, A. (2006). Learners' interpretations of recasts. *Studies in Second Language Acquisition*, 28, 209–236.
- Chapelle, C. (Ed.). (2012). *The encyclopedia of applied linguistics*. Oxford: Blackwell.
- de Graaff, R., & Housen, A. (2009). Investigating the effects and effectiveness of L2 instruction. In M. H. Long & C. J. Doughty (Eds.), *The handbook of language teaching* (pp. 726–755). Oxford: Blackwell.

- DeKeyser, R. M. (2005). What makes learning second-language grammar difficult? A review of issues. *Language Learning*, 55, 1–25.
- Dilans, G. (2010). Corrective feedback and L2 vocabulary development: Prompts and recasts in the adult ESL classroom. *Canadian Modern Language Review*, 66, 787–815.
- Dörnyei, Z. (2005). *The psychology of the language learner: Individual differences in second language acquisition*. Mahwah, NJ: Erlbaum.
- Dörnyei, Z. (2009). *Psychology of second language acquisition*. Oxford: Oxford University Press.
- Dörnyei, Z., & Skehan, P. (2003). Individual differences in second language learning. In C. J. Doughty & M. H. Long (Eds.), *The handbook of second language acquisition* (pp. 589–630). Malden, MA: Blackwell.
- Doughty, C. J. (2001). Cognitive underpinnings of focus on form. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 206–257). New York: Cambridge University Press.
- Doughty, C. J., & Varela, E. (1998). Communicative focus on form. In C. J. Doughty & J. Williams (Eds.), *Focus on form in classroom second language acquisition* (pp. 114–138). New York: Cambridge University Press.
- Egi, T. (2007a). Interpreting recasts as linguistic evidence: The roles of linguistic target, length, and degree of change. *Studies in Second Language Acquisition*, 29, 511–537.
- Egi, T. (2007b). Recasts, learners' interpretations, and L2 development. In A. Mackey (Ed.), *Conversational interaction in second language acquisition: A collection of empirical studies* (pp. 249–267). Oxford: Oxford University Press.
- Egi, T. (2010). Uptake, modified output, and learner perceptions of recasts: Learner responses as language awareness. *Modern Language Journal*, 94, 1–21.
- Ellis, R. (2004). Individual differences in second language learning. In A. Davies & C. Elder (Eds.), *The handbook of applied linguistics* (pp. 525–551). Oxford: Blackwell.
- Ellis, R. (2007). The differential effects of corrective feedback on two grammatical structures. In A. Mackey (Ed.), *Conversational interaction in second language acquisition: A collection of empirical studies* (pp. 339–360). Oxford: Oxford University Press.
- Ellis, R., Basturkmen, H., & Loewen, S. (2001). Learner uptake in communicative ESL lessons. *Language Learning*, 51, 281–318.
- Ellis, R., Loewen, S., & Erlam, R. (2006). Implicit and explicit corrective feedback and the acquisition of L2 grammar. *Studies in Second Language Acquisition*, 28, 339–368.
- Ellis, R., & Sheen, Y. (2006). Reexamining the role of recasts in second language acquisition. *Studies in Second Language Acquisition*, 28, 575–600.
- Erlam, R., & Loewen, S. (2010). Implicit and explicit recasts in L2 oral French interaction. *Canadian Modern Language Review*, 66, 877–905.
- Farrar, M. J. (1992). Negative evidence and grammatical morpheme acquisition. *Developmental Psychology*, 28, 90–98.
- Gass, S. M. (1997). *Input, interaction, and the second language learner*. Mahwah, NJ: Erlbaum.
- Gass, S. M. (2003). Input and interaction. In C. J. Doughty & M. H. Long (Eds.), *The handbook of second language acquisition* (pp. 224–255). Oxford: Blackwell.
- Gass, S. M., & Mackey, A. (2006). Input, interaction and output: An overview. *AILA Review*, 19, 3–17.
- Gass, S. M., & Mackey, A. (2007). Input, interaction, and output in second language acquisition. In B. VanPatten & J. Williams (Eds.), *Theories in second language acquisition: An Introduction* (pp. 175–199). Mahwah, NJ: Erlbaum.
- Gass, S. M., & Mackey, A. (2012). *The Routledge handbook of second language acquisition*. London: Routledge.
- Goo, J. (2012). Corrective feedback and working memory capacity in interaction-driven L2 learning. *Studies in Second Language Acquisition*, 34, 445–474.
- Gordon, P. (1990). Learnability and feedback. *Developmental Psychology*, 26, 217–220.
- Grimshaw, J., & Pinker, S. (1989). Positive and negative evidence in language acquisition. *Behavioral and Brain Sciences*, 12, 341–342.
- Han, Z. (2002). A study of the impact of recasts on tense consistency in L2 output. *TESOL Quarterly*, 36, 543–572.
- Hinkel, E. (Ed.). (2011). *Handbook of research in second language teaching and learning* (Vol. 2). London: Routledge.
- Hulstijn, J. H. (2012). Is the second language acquisition discipline disintegrating? *Language Teaching*. Advance online publication. doi:10.1017/S0261444811000620.

- Ishida, M. (2004). Effects of recasts on the acquisition of the aspectual form *-te i-(ru)* by learners of Japanese as a foreign language. *Language Learning*, 54, 311–394.
- 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.
- Keck, C. M., Iberri-Shea, G., Tracy-Ventura, N., & Wa-Mbaleka, S. (2006). Investigating the empirical link between task-based interaction and acquisition: A meta-analysis. In J. M. Norris & L. Ortega (Eds.), *Synthesizing research on language learning and teaching* (pp. 91–131). Amsterdam: Benjamins.
- Krashen, S. (1982). *Principles and practice in second language acquisition*. Oxford: Pergamon Press.
- Krashen, S. (1985). *The input hypothesis*. London: Longman.
- 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.
- Li, S. (2010). The effectiveness of corrective feedback in SLA: A meta-analysis. *Language Learning*, 60, 309–365.
- Loewen, S. (2009). Recasts in multiple response focus on form episodes. In A. Mackey & C. Polio (Eds.), *Multiple perspectives on interaction* (pp. 176–196). London: Routledge.
- Loewen, S., & Erlam, R. (2006). Corrective feedback in the chatroom: An experimental study. *Computer Assisted Language Learning*, 19, 1–14.
- Loewen, S., & Nabei, T. (2007). Measuring the effects of oral corrective feedback on L2 knowledge. In A. Mackey (Ed.), *Conversational interaction in second language acquisition: A collection of empirical studies* (pp. 361–377). Oxford: Oxford University Press.
- Loewen, S., & Philp, J. (2006). Recasts in the adult English L2 classroom: Characteristics, explicitness, and effectiveness. *Modern Language Journal*, 90, 536–556.
- Long, M. H. (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). San Diego, CA: Academic Press.
- Long, M. H. (2007). *Problems in SLA*. Mahwah, NJ: Erlbaum.
- 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, 357–371.
- Long, M. H., & Robinson, P. (1998). Focus on form: Theory, research, process. In C. J. Doughty & J. Williams (Eds.), *Focus on form in classroom second language acquisition* (pp. 15–41). New York: Cambridge University Press.
- Lyster, R. (1998a). Negotiation of form, recasts, and explicit correction in relation to error types and learner repair in immersion classrooms. *Language Learning*, 48, 183–218.
- Lyster, R. (1998b). Recasts, repetition, and ambiguity in L2 classroom discourse. *Studies in Second Language Acquisition*, 20, 51–81.
- Lyster, R. (2004). Differential effects of prompts and recasts in form-focused instruction. *Studies in Second Language Acquisition*, 26, 399–432.
- Lyster, R., & Izquierdo, J. (2009). Prompts versus recasts in dyadic interaction. *Language Learning*, 59, 453–498.
- Lyster, R., & Mori, H. (2006). Interaction feedback and instructional counterbalance. *Studies in Second Language Acquisition*, 28, 269–300.
- 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. (2007). Interaction as practice. In R. M. DeKeyser (Ed.), *Practice in a second language: Perspectives from applied linguistics and cognitive psychology* (pp. 85–110). New York: Cambridge University Press.
- Mackey, A. (2012). *Input, interaction and corrective feedback in L2 learning*. Oxford: Oxford University Press.
- Mackey, A., Abbuhl, R., & Gass, S. M. (2012). The interaction approach. In S. M. Gass & A. Mackey (Eds.), *The Routledge handbook of second language acquisition* (pp. 7–23). London: Routledge.
- 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., & Goo, J. (in press). Interaction approach in second language acquisition. In C. A. Chapelle (Ed.), *The encyclopedia of applied linguistics*. Oxford: Blackwell.
- Mackey, A., Oliver, R., & Leeman, J. (2003). Interactional input and incorporation of feedback: An exploration of NS-NNS and NNS-NNS adult and child dyads. *Language Learning*, 53, 35–66.
- Mackey, A., & Philp, J. (1998). Conversational interaction and second language development: Recasts, responses, and red herrings? *Modern Language Journal*, 82, 338–356.
- 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. 181–209). Amsterdam: Benjamins.
- McDonough, K. (2005). Identifying the impact of negative feedback and learners' responses on ESL question development. *Studies in Second Language Acquisition*, 27, 79–103.
- McDonough, K. (2007). Interactional feedback and the emergence of simple past activity verbs in L2 English. In A. Mackey (Ed.), *Conversational interaction in second language acquisition: A collection of empirical studies* (pp. 323–338). Oxford: Oxford University Press.
- McDonough, K., & Mackey, A. (2006). Responses to recasts: Repetitions, primed production, and linguistic development. *Language Learning*, 56, 693–720.
- Morgan, J. L., Bonamo, K. M., & Travis, L. L. (1995). Negative evidence on negative evidence. *Developmental Psychology*, 31, 180–197.
- Nassaji, H. (2009). Effects of recasts and elicitations in dyadic interaction and the role of feedback explicitness. *Language Learning*, 59, 411–452.
- Nicholas, H., Lightbown, P. M., & Spada, N. (2001). Recasts as feedback to language learners. *Language Learning*, 51, 719–758.
- Nobuyoshi, J., & Ellis, R. (1993). Focused communication tasks and second language acquisition. *ELT Journal*, 47, 203–210.
- Ohta, A. S. (2000). Rethinking recasts: A learner-centered examination of corrective feedback in the Japanese language classroom. In J. K. Hall & L. S. Verplaeste (Eds.), *The construction of second and foreign language learning through classroom instruction* (pp. 47–71). Mahwah, NJ: Erlbaum.
- Oliver, R. (1995). Negative feedback in child NS-NNS conversation. *Studies in Second Language Acquisition*, 17, 459–481.
- Oliver, R. (1998). Negotiation of meaning in child interactions. *Modern Language Journal*, 82, 372–386.
- Oliver, R. (2000). Age differences in negotiation and feedback in classroom and pairwork. *Language Learning*, 50, 119–151.
- Oliver, R., & Mackey, A. (2003). Interactional context and feedback in child ESL classrooms. *Modern Language Journal*, 87, 519–533.
- Ortega, L., & Long, M. H. (1997). The effects of models and recasts on the acquisition of object topicalization and adverb placement in L2 Spanish. *Spanish Applied Linguistics*, 1, 65–86.
- Panova, I., & Lyster, R. (2002). Patterns of corrective feedback and uptake in an adult ESL classroom. *TESOL Quarterly*, 36, 573–595.
- Petersen, K. A. (2010). *Implicit corrective feedback in computer-guided interaction: Does mode matter?* (Unpublished doctoral dissertation). Georgetown University, Washington, DC.
- 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.
- 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.
- Pica, T. (1994). Research on negotiation: What does it reveal about second-language learning conditions, processes, and outcomes? *Language Learning*, 44, 493–527.
- Pica, T. (1996). Do second language learners need negotiation? *International Review of Applied Linguistics in Language Teaching*, 34, 1–21.

- Pienemann, M., & Johnston, M. (1987). Factors influencing the development of language proficiency. In D. Nunan (Ed.), *Applying second language acquisition research* (pp. 45–141). Adelaide, Australia: National Curriculum Resource Centre, AMEP.
- Pinker, S. (1989). *Learnability and cognition: The acquisition of argument structure*. Cambridge, MA: MIT Press.
- Révész, A. (2009). Task complexity, focus on form, and second language development. *Studies in Second Language Acquisition*, 31, 437–470.
- Révész, A., & Han, Z. (2006). Task content familiarity, task type, and efficacy of recasts. *Language Awareness*, 15, 160–179.
- Révész, A., Sachs, R., & Mackey, A. (2011). Task complexity, uptake of recasts, and second language development. In P. Robinson (Ed.), *Second language task complexity: Researching the cognition hypothesis of language learning and performance* (pp. 203–235). Amsterdam: Benjamins.
- Robinson, P. (Ed.). (2002). *Individual differences and instructed language learning*. Amsterdam: Benjamins.
- Robinson, P. (2005). Aptitude and second language acquisition. *Annual Review of Applied Linguistics*, 25, 46–73.
- Robinson, P. (2007). Aptitudes, abilities, contexts, and practice. In R. M. DeKeyser (Ed.), *Practice in a second language: Perspectives from applied linguistics and cognitive psychology* (pp. 256–286). New York: Cambridge University Press.
- Robinson, P., Mackey, A., Gass, S. M., & Schmidt, R. (2012). Attention and awareness in second language acquisition. In S. M. Gass & A. Mackey (Eds.), *The Routledge handbook of second language acquisition* (pp. 247–267). London: Routledge.
- Romanova, N. (2010). Planning, recasts, and learning of L2 morphology. *Canadian Modern Language Review*, 66, 843–875.
- Ross-Feldman, L. (2007). Interaction in the L2 classroom: Does gender influence learning opportunities? In A. Mackey (Ed.), *Conversational interaction in second language acquisition: A collection of empirical studies* (pp. 53–77). Oxford: Oxford University Press.
- Russell, J., & Spada, N. (2006). The effectiveness of corrective feedback for the acquisition of L2 grammar: A meta-analysis of the research. In J. Norris & L. Ortega (Eds.), *Synthesizing research on language learning and teaching* (pp. 133–164). Amsterdam: Benjamins.
- Sachs, R., & Suh, B.-R. (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. (2007). From CALL to face-to-face interaction: The effect of computer-delivered recasts and working memory on L2 development. In A. Mackey (Ed.), *Conversational interaction in second language acquisition: A collection of empirical studies* (pp. 229–248). Oxford: Oxford University Press.
- Saito, K., & Lyster, R. (2012). Effects of form-focused instruction and corrective feedback on L2 pronunciation development of /ɪ/ by Japanese learners of English. *Language Learning*, 62, 595–633.
- Sato, M., & Lyster, R. (2007). Modified output of Japanese EFL learners: Variable effects of interlocutor vs. feedback types. In A. Mackey (Ed.), *Conversational interaction in second language acquisition: A collection of empirical studies* (pp. 123–142). Oxford: Oxford University Press.
- Sauro, S. (2009). Computer-mediated corrective feedback and the development of L2 grammar. *Language Learning and Technology*, 13, 96–120.
- Sawyer, M., & Ranta, L. (2001). Aptitude, individual differences, and instructional design. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 319–353). New York: Cambridge University Press.
- Saxton, M. (1997). The contrast theory of negative input. *Journal of Child Language*, 24, 139–161.
- Saxton, M., Backley, P., & Gallaway, C. (2005). Negative input for grammatical errors: Effects after a lag of 12 weeks. *Journal of Child Language*, 32, 643–672.
- Saxton, M., Kulcsar, B., Marshall, G., & Rupra, M. (1998). Longer-term effects of corrective input: An experimental approach. *Journal of Child Language*, 25, 701–721.
- Schmidt, R. (1990). The role of consciousness in second language learning. *Applied Linguistics*, 11, 129–158.

- Schmidt, R. (1993). Consciousness, learning, and interlanguage pragmatics. In G. Kasper & S. Blum-Kulka (Eds.), *Interlanguage pragmatics* (pp. 21–42). Oxford: Oxford University Press.
- Schmidt, R. (1995). Consciousness and foreign language learning: A tutorial on the role of attention and awareness in learning. In R. Schmidt (Ed.), *Attention and awareness in foreign language learning* (pp. 1–63). Honolulu: University of Hawai'i Press.
- Schmidt, R. (2001). Attention. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 3–32). New York: Cambridge University Press.
- Schmidt, R., & Frota, S. (1986). Developing basic conversational ability in a second language: A case study of an adult learner of Portuguese. In R. Day (Ed.), *Talking to learn: Conversation in second language acquisition* (pp. 237–322). Rowley, MA: Newbury House.
- Sheen, Y. (2004). Corrective feedback and learner uptake in communicative classrooms across instructional settings. *Language Teaching Research*, 8, 263–300.
- Sheen, Y. (2006). Exploring the relationship between characteristics of recasts and learner uptake. *Language Teaching Research*, 10, 361–392.
- Sheen, Y. (2007a). The effect of focused written corrective feedback and language aptitude on ESL learners' acquisition of articles. *TESOL Quarterly*, 41, 255–283.
- Sheen, Y. (2007b). 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. (2008). Recasts, language anxiety, modified output, and L2 learning. *Language Learning*, 58, 835–874.
- Sheen, Y. (2010). Differential effects of oral and written corrective feedback in the ESL classroom. *Studies in Second Language Acquisition*, 32, 203–234.
- Shintani, N., & Ellis, R. (2010). The incidental acquisition of English plural –s by Japanese children in comprehension-based and production-based lessons: A process-product study. *Studies in Second Language Acquisition*, 32, 607–637.
- Skehan, P. (2002). Theorising and updating aptitude. In P. Robinson (Ed.), *Individual differences and instructed language learning* (pp. 69–93). Amsterdam: Benjamins.
- Smith, B. (2010). Employing eye-tracking technology in researching the effectiveness of recasts in CMC. In F. M. Hult (Ed.), *Directions and prospects for educational linguistics* (pp. 79–97). New York: Springer Verlag.
- Smith, B., & Renaud, C. (in press). Eyetracking and classroom interaction. In K. McDonough & A. Mackey (Eds.), *New perspectives on classroom interaction in second language research*. Amsterdam: Benjamins.
- Spada, N. (2011). Beyond form-focused instruction: Reflections on past, present and future research. *Language Teaching*, 44, 225–236.
- Spada, N., & Lightbown, P. M. (2008). Form-focused instruction: Isolated or integrated? *TESOL Quarterly*, 42, 181–207.
- Spada, N., & Lightbown, P. M. (2009). Interaction research in second/foreign language classrooms. In A. Mackey & C. Polio (Eds.), *Multiple perspectives on interaction* (pp. 157–175). London: Routledge.
- Swain, M. (1985). Communicative competence: Some roles of comprehensible input and comprehensible output in its development. In S. M. Gass & C. Madden (Eds.), *Input in second language acquisition* (pp. 235–253). Rowley, MA: Newbury House.
- Swain, M. (1993). The output hypothesis: Just speaking and writing aren't enough. *Canadian Modern Language Review*, 50, 158–164.
- Swain, M. (1995). Three functions of output in second language learning. In G. Cook & B. Seidlhofer (Eds.), *Principle and practice in applied linguistics: Studies in honour of H. G. Widdowson* (pp. 125–144). Oxford: Oxford University Press.
- Swain, M. (2005). The output hypothesis: Theory and research. In E. Hinkel (Ed.), *Handbook of research in second language teaching and learning* (pp. 471–483). Mahwah, NJ: Erlbaum.
- Takashima, H., & Ellis, R. (1999). Output enhancement and the acquisition of the past tense. In R. Ellis (Ed.), *Learning a second language through interaction* (pp. 173–188). Amsterdam: Benjamins.

- Trofimovich, P., Ammar, A., & Gatbonton, E. (2007). How effective are recasts? The role of attention, memory, and analytical ability. In A. Mackey (Ed.), *Conversational interaction in second language acquisition: A collection of empirical studies* (pp. 171–195). Oxford: Oxford University Press.
- Varnosfadrani, A. D., & Basturkmen, H. (2009). The effectiveness of implicit and explicit error correction on learners' performance. *System*, 37, 82–98.
- Yang, Y., & Lyster, R. (2010). Effects of form-focused practice and feedback on Chinese EFL learners' acquisition of regular and irregular past tense forms. *Studies in Second Language Acquisition*, 32, 235–263.