How Useful are Recasts?: Factors Influencing Their Success and Problems in Testing

Kate Burrill[1]

INTRODUCTION

As the popular focus of instruction in language classrooms becomes increasingly meaning-oriented, teachers must nevertheless ensure that their students are also learning the correct form of the language they are studying. One of the ways to achieve this task is to provide negative feedback—correcting the student, either implicitly or explicitly. The most common type of negative feedback used in the classroom is recasting (Sheen, 2006; Panova & Lyster, 2002; Loewen & Philp, 2006). Recasting is defined as the reformulation of a non-native speaker's incorrect utterance by a native speaker in order to correct it (Gass & Selinker, 2008). Recasts are a common type of feedback for many possible reasons; one of the main reasons may be that they allow the teacher to maintain a focus on meaning while still giving the non-native speaker implicit correction on form (Han, 2002). The prevalence of recasts in the classroom has led to many studies on the topic, but results from the research have generally not provided strong evidence of its effectiveness.

In the following sections, I will first describe studies that used pretests and posttests in evaluating students' grammatical accuracy after receiving recasts. I will then describe studies

[1] Kate Burrill graduated from Teachers College in 2012 with an MA in Applied Linguistics. As a Fellow in the English Language Fellowship program, she is currently teaching English at the State Islamic School of Sunan Ampel in Surabaya, Indonesia.

that investigated student attention to recasts, followed by a discussion of the role of uptake in determining recast efficacy. Finally, I will address the factors that may influence the efficacy of recasting, as well as some key problems in testing that have been brought to light by the studies attempting to measure the effects of corrective feedback in the classroom.

STUDENT ACCURACY

The first studies to be reviewed all used a variety of formal tests to evaluate learners' accuracy in using a specific grammar form, and most included a pretest, posttest, and delayed posttest. The majority of these studies compared the efficacy of recasts to other forms of negative feedback. Loewen and Philp (2006) had a six-part research question, the first two parts of which dealt with the characteristics of recasts and the frequency of recasts in relation to other types of feedback in the L2 classroom. The researchers chose to compare recasts with elicitations, in which the learner is pushed to correct his or her errors without explicit correction or metalinguistic feedback (e.g. "Do we say that in French?") and with informs, in which the learner receives explicit metalinguistic feedback on the incorrect form (e.g., "You need a noun here"). Their second two questions addressed whether recasts resulted in higher accuracy on posttests and more instances of successful uptake than other forms of feedback. Finally, the last two questions asked which characteristics of recasts are associated with higher posttest scores and successful uptake.

Loewen and Philp (2006) carried out their study on classes composed of 118 adult L2 English learners. The researchers observed 32 hours of instruction, but analyzed only the 17 hours that were clearly meaning-focused, as they wanted the recasts to be implicit to the learner, which would probably be less likely during a form-oriented lesson. They then coded the data for

the types of corrective feedback (elicitation, informs, and recasts) used during class, and specifically tested only the students who had received these types of feedback. These students did not take a pretest, as the forms to be tested on were based on the errors they made that received corrective feedback. The learners were tested twice: 1 to 3 days after the feedback occurred, and again 13 to 15 days after it occurred. The researchers tested the students orally on the same grammatical or phonetic forms that they had been corrected on during classroom observation, as a way to judge whether the student could now correct his own error.

The results of Loewen and Philp's (2006) study showed no significant difference in student learning when comparing recasts to other forms of corrective feedback. They found that, as previous research has shown, recasts are used much more in the classroom than other types of corrective feedback. However, results on both the first and second posttest did not show significant differences between the three categories of recasts, informs, and elicitations. Although they did find the most instances of successful uptake following elicitations (83.1% vs. 59.6%), they discovered that this uptake did not necessarily lead to significantly higher scores on the posttest (50% vs. 59.4%). Furthermore, because the researchers did not administer a pretest, it was difficult to judge how much the students' accuracy improved overall.

Lyster and Izquierdo's (2009) study tested for the effect of recasts on accuracy compared with prompts. Prompts are a type of negative feedback in which the NNS is given signals by the NS, prompting him to self-repair, in contrast to recasts, which supply the correct reformulation immediately. Lyster and Izquierdo hypothesized that the students who received prompts as feedback would have higher accuracy scores on posttests than those who had received recasts as feedback. They posited that "learners receiving prompts will show more improvement as a result

of the opportunities that prompts provide for a deeper level of processing as learners are pushed to retrieve target forms and to produce modified output" (p. 465).

In Lyster and Izquierdo's (2009) study, the researchers tested 25 adult intermediate L2 learners of French. The students were placed in a 13-week university French class that consisted of three hours of instruction per week. The students were randomly divided into two groups, a recast and a prompt group. Each group attended two extra treatment sessions outside of regular class time. In these sessions, the teacher gave each group form-focused feedback using only prompts or only recasts, depending on the group. There were also three testing sessions during the course of the study. The participants were given a pretest on French grammatical gender the week before classes started, a posttest in the last week of the class (the ninth week), and a delayed posttest in the 14th week. Each testing session consisted of two oral sections and one computerized binary choice test.

Lyster and Izquierdo's (2009) results also showed no significant difference between recasting and prompts. The subjects *did* improve from the pretest to the delayed posttest; however, all parts of the test (oral and computerized binary testing) showed minimal difference when comparing prompts and recasts. For example, in the second oral picture-description posttest, students from the prompt group had a mean score of 3.11, while the recast group had a mean score of 3.16. In this case, the results from the recast group were slightly higher than the prompt group, contradicting the researchers' hypothesis.

Ellis, Loewen, and Erlam (2006) tried to find out whether implicit or explicit feedback was more successful when learning the regular past tense *-ed*. To do this, they divided 34 lower-intermediate adult ESL students into three experimental groups and asked them to narrate a picture story while the instructor provided different types of feedback. One group of 12 students

received only recasts, 12 received only metalinguistic information, and the third group of 10 received no corrective feedback.

The authors tested explicit L2 knowledge of the regular past tense by administering a pretest, posttest, and delayed posttest consisting of an untimed grammaticality judgment test, a metalinguistic knowledge test, and an oral imitation test, in which learners listened to a sentence that was either ungrammatical or grammatical, and then wrote down whether they thought the sentence was correct or incorrect. Then, learners repeated the statement orally, correcting the statement if it was wrong. The oral aspect of this test was designed to pressure the students to speak in real time and "operate by feel," thus aiming to test implicit knowledge of the grammar structure (p. 354).

While there were no significant differences between the three groups on the immediate posttest, the authors reported a significant advantage of the metalinguistic feedback group on the delayed oral imitation posttest and the delayed grammaticality judgment posttest. Interestingly, these groups showed improvement on old vocabulary as well as new vocabulary that the learners had never encountered before. The fact that students were able to extend their knowledge of the regular past tense form to new vocabulary items suggested that learners were applying grammatical rules to the language system, rather than simply memorizing items of vocabulary. Additionally, on the delayed grammaticality judgment posttest, students who had received metalinguistic feedback showed improvement in identifying grammatical sentences successfully, which Ellis (2005) cited as a sign of implicit language knowledge. Thus, the researchers found that metalinguistic feedback appeared to be more successful than recasts in the longer term towards improving students' implicit and explicit grammar knowledge.

Studies by Nassaji (2009) and Rezaei and Derakshan (2011) were similar to that of Ellis et al. (2006), because these researchers also found that the explicit metalinguistic feedback groups scored higher on posttests than the groups receiving recasts. Nassaji's study was especially similar, because she also found no significant difference between implicit (recasts) and explicit (elicitations) feedback groups on the immediate posttest, but delayed posttests from the explicit feedback group showed significant improvement in students' ability to identify and correct a variety of lexical and grammatical errors in their writing. Rezaei and Derakshan found that students' immediate posttest scores on the conditional form improved more in the metalinguistic feedback group than the recast group (a delayed posttest was not given).

A final study did not compare student accuracy after recasts to other corrective feedback, but focused more on testing recasts in contrast to no corrective feedback at all. Han (2002) tested whether learners who received recasts showed greater tense consistency, and greater awareness of tense consistency, than learners who did not receive corrective feedback. She defined "awareness" as the learner showing a systematic change in behavior over time. Han hypothesized that the learners she had selected as subjects would be receptive to recasts and would exhibit both greater tense consistency and greater awareness because they were developmentally ready. In other words, recasts would help those learners who had knowledge of when and how to use English tenses, but who lacked control of these forms.

The subjects in Han's (2002) study were eight upper-intermediate adult L2 English learners, who were divided into two groups of four subjects each. Over a period of eleven weeks consisting of a pretest, eight weeks of English instruction, a posttest, and a delayed posttest one month later, one group received recasts, while the contrast group did not receive recasts. In each session, the subjects in both groups were given the task of relating a story in both written and

oral forms. During classroom sessions, the instructor in the recast group gave recasts focusing on tense inconsistency and also gave random recasts of other incorrect forms. The instructor in the contrast group did not provide corrective feedback.

To determine the results, the oral and written narratives were transcribed and analyzed. Han (2002) found that recasts appeared to be successful in promoting tense consistency and awareness in both written and oral forms. On the pretest, both the recast and non-recast groups showed minimal tense consistency, because the learners showed a strong preference for present tense. For example, 0.28 was the mean proportion disparity between present and past tenses for both groups in the oral component. However, on the posttest and delayed posttest, learners from the recast group had much more control over their use of present or past tense. The disparity between tenses on the oral and written delayed posttest was 0.66 and 0.78 for the recast group, while the non-recast group disparity was 0.22 for oral and 0.1 for written, indicating that the non-recast group still relied too much on present tense.

As an additional qualitative component to her study, and to examine the learners' awareness of tense consistency, Han (2002) looked at the scores of two individuals: Jee-Young, from the recast group, and Blanka, from the contrast group. Jee-Young's ability to control tenses improved greatly during the course of the session, and the oral recasts she was given appeared to make her aware that both oral narratives as well as her written narratives required past tense. She also self-corrected tense more often as the classes continued. In contrast, Blanka, who was part of the non-recast group, did not show any improvement in tense consistency throughout the class session; her use of present and past tenses continued to fluctuate even on the posttest. Han cited the differences in these two learners' control and self-correction attempts as evidence of the

instructor's recasts making the learners not only more consistent in tense, but also more aware of their output.

STUDENT ATTENTION AND RESPONSES TO RECASTING

Not all studies on the effects of recasts are centered on data from pretests and posttests. Many relevant studies also investigated learner attention to recasts, and attempted to determine whether student interpretation of recasts aligns with teacher intent.

Carpenter, Jeon, MacGregor, and Mackey (2006) investigated the role of ambiguity and extralinguistic signals in recasts. Many researchers, including Lyster (1997) and Panova and Lyster (2002), have argued that recasts are not as effective as other forms of corrective feedback because learners encounter them in the same forms of discourse as non-corrective repetitions, and therefore learners may not realize the corrective function of recasts. Other researchers have also suggested that learners may rely more on paralinguistic cues in order to distinguish corrective recasts from non-corrective repetition for confirmation. In their research questions, Carpenter et al. asked first whether learners could recognize recasts when removed from their immediate discourse context, and second, if learners used nonlinguistic cues such as facial expressions or gestures when determining whether the teacher's comment was a recast or a repetition.

To create their stimulus, Carpenter et al. (2006) videotaped 26 adult high intermediate and advanced level ESL students carrying out task-based activities in dyadic interactions with native English-speaking teachers. The teachers were asked to use recasts and repetitions as appropriate, but that they could use any feedback that they desired. The researchers then coded these videotaped interactions, and 9 examples of recasts, 9 repetitions, and 7 distracters were

chosen to use as the stimulus. The videos were then edited into two separate videos: one that showed the learners' utterances and the instructor's responses, and another that showed only the response of the instructor. A new group of 34 adult advanced level ESL students were selected to first carry out the same task that the first group of students had completed. Half watched only the utterance-response video, while the other half watched the response-only video. Both groups filled out a multiple choice answer sheet, marking what he or she thought the instructor was trying to do in each feedback episode (correct an error, repeat an utterance, give instructions, etc.). Seven learners from each group were also asked to think aloud as they filled in their answers.

Carpenter et al. (2006) found that learners correctly identified recasts more often in the utterance-response videos than in the response-only videos; however, both groups were equally likely to misidentify recasts as repetitions. As for the role of nonlinguistic cues, only one comment out of 252 total learner comments explicitly referred to the nonlinguistic cue of the teacher's facial expression, and in this case, the learner misidentified a recast as a repetition. The researchers found that learners generally focused on meaning, without evidence of focus on nonlinguistic cues. As a final *post hoc* analysis, the researchers coded the corrective feedback for the type of error: phonological, lexical, or morphosyntactic, to see whether there was a connection between the type of error and the learners' ability to identify it. The results of their coding showed that the teacher gave morphosyntactic corrections most frequently, but that learners were more accurate in recognizing corrective recasts that focused on lexical or phonological errors.

Nabei and Swain (2002) studied Shoko, a native Japanese speaker learning English in an EFL environment. The researchers wanted to find out how many opportunities Shoko had to hear

recasts from the instructor, and whether Shoko was aware of receiving the recasts. Furthermore, they wished to study what kind of connection existed between the recasts, Shoko's awareness of the recast, and her own L2 learning.

To carry out their study, Nabei and Swain (2002) observed and videotaped Shoko's 70-minute discussion-based class six times. Within a week of each observation, the researchers gave Shoko a short grammaticality judgment test, composed of 8 to 15 sentences based on recast episodes (REs) that had occurred in the previous class. At each of these testing sessions, Shoko was also asked to watch video from the previous classroom session and recall what she had been thinking at the moment a RE happened. The researchers also showed other types of feedback as distracters. Three weeks after the course was completed, the researchers gave Shoko a final delayed grammaticality judgment posttest consisting of items from the previous tests.

Nabei and Swain (2002) were surprised to find that the instructor provided feedback quite rarely; out of 420 minutes of observation, the instructor gave feedback only 25 times. Of these 25, recasts were used 23 times (92%), supporting previous research that recasting is the most common type of classroom feedback. As for Shoko's awareness of recasts, she tended not to listen to the other student's answers in class, although she was interested in the teacher's response to student answers. The researchers found that Shoko's attention was generally focused on meaning, and that she rarely noticed the teacher's corrections to form. The grammaticality judgment tests administered within a week after the class did not show significant improvement (Shoko correctly answered 56% correct answers overall). Nevertheless, the delayed posttest did show improvement (78% correct overall). The researchers did not discuss Shoko's delayed posttest improvement, except to point out that she scored higher in the feedback that was given during group work, as opposed to feedback given when the classroom was teacher-fronted.

Kim and Han (2007) carried out a study related to Nabei and Swain's (2002), in that it also investigated learner interpretation of recasts. However, Kim and Han looked more in depth at how much teacher intent and learner interpretation overlapped, as well as how accurately learners could judge the gap between their original utterance and the information in the recast they received. As a final component of their research question, the authors wanted to find out what elements of the interaction influenced learner recognition of gaps, such as the type of addressee, the type of linguistic target, or the form of the recast.

Kim and Han (2007) videotaped 200 minutes of four intermediate adult EFL classes in Korea. After each 50-minute recording session, the researchers held individual stimulated recall interviews with five students from that class, in which each was shown episodes in which a recast was given, and then asked to recall what he or she was thinking at that moment. The two teachers of the classes were also asked to provide stimulated recall after watching the recording of each recast they provided. After the interviews, the learners' and teachers' comments were transcribed, and the recast episodes were coded according to complexity, linguistic content, and form and meaning.

In their results, the Kim and Han (2007) found that teacher intent and learner interpretation overlapped significantly, with learners recognizing 69% of simple corrective recasts as corrective, and 58% of the complex corrective recasts. Similarly, learners recognized 76% of simple communicative recasts, and 80% of complex communicative recasts. In contrast to Nabei and Swain's (2002) findings on Shoko, Kim and Han found that direct and indirect addressees perceived corrective recasts equally well. Learners were able to recognize gaps more easily in corrective recasts than in communicative recasts, and more easily in simple recasts than complex ones. Finally, the data in this study showed that learner recognition of gaps was also

affected by the type of recast, with recognition of phonological correction noticed most frequently. In contrast, teachers were least likely to use recasts to correct phonological errors; the most common category of corrective recasts targeted morphological errors.

These results were similar to Sheen's (2006) descriptive study observing 12 hours of an adult intermediate ESL classroom and an adult high-intermediate EFL classroom, in which grammatical recasts were most frequently targeted by teachers (52% of all recasts were grammatical), but learners were least likely to notice grammatical recasts. Instead, learners were most likely to successfully repair pronunciation recasts (86.7% of all pronunciation uptake was successful). Whether a learner's repair may be interpreted as a sign of learning will be discussed in the following section.

UPTAKE

The role of learner uptake in recasts has inspired a lot of debate and therefore deserves to be mentioned here, as researchers continue to argue both for and against its importance in assessing the effectiveness of corrective feedback and, more specifically, recasts. Uptake is defined as the learner's immediate response to the teacher's corrective feedback. It may or may not be a repair of the original incorrect utterance. It may be important partly because it could be a sign of noticing, which is an integral step towards restructuring a learner's interlanguage. In a seminal study on corrective feedback and uptake, Lyster and Ranta (1997) recorded and transcribed six hours of French immersion classrooms in Quebec. They compared a large variety of corrective feedback forms, and assessed which forms were most common in the classroom, and which ones were followed most frequently by uptake. In their analysis, the researchers argued that recasts are less effective than other forms of feedback such as elicitations or

clarification requests, because recasts did not push learner output as much as the other forms. Because the teacher provided the corrected form for the student, there was not often a chance for the student to repeat the correction before the conversation had moved on. As there were no pre- or posttests administered in their study, the lack of uptake following recasts was the chief evidence Lyster and Ranta used to support their argument against the effectiveness of recasts.

However, some researchers have pointed out problems with Lyster and Ranta's (1997) analysis of the data. Doughty and Long (2003) discussed the problem of natural classroom discourse not allowing for uptake. Often it is impossible for a student to produce any kind of uptake after a recast because the teacher provides a recast with a continuation of the topic:

> S: There is some famous restaurant.
> T: Where are the famous restaurants? Are they near Kang-Nam?
> S: Well, I don't know … umm … I know one … near Yong-In.
>
> (Kim and Han, 2007, p. 281)

In other cases it would be inappropriate to make a comment, because the recast had been in the form of a *yes/no* question and the appropriate response would not be a repetition, but a *yes/no* response:

> S: It sounds fish to me.
> T: Fishy?
> S: Yes.
>
> (Kim and Han, 2007, p. 281)

Therefore, Lyster and Ranta (1997) may have been too quick to dismiss recasts as ineffective solely based on the lack of uptake directly following a recast. They did not distinguish between the finer details of classroom discourse that might have made it impossible for the student to produce any kind of uptake. As Sheen (2004) pointed out in her study examining Lyster and Ranta's data, 70% of recasts in their study were followed by topic continuation. Of the episodes

when learners were able to produce uptake, repairs accounted for 57% of all uptake, which is much higher than Lyster and Ranta's original figure of 18%.

Researchers such as Long (2007) and Sheen (2006) have also argued that the usefulness of uptake in determining learning is questionable. Long pointed out that no studies have proved that a learner's immediate repair signifies immediate corrective effects. He added that oral production is often one of the last indications of change in an underlying grammar rule. Sheen (2006) also stated that while successful uptake may be evidence of noticing, the opposite is not automatically true: lack of uptake does not mean that a learner has not noticed the form. In fact, In Loewen and Philp's (2006) study described above, the authors did not find a strong connection between successful uptake and higher posttest scores.

A recent study by Nassaji (2011) investigated the potential differences in uptake by testing whether learner self-repair in response to elicitations or teacher repair due to recasts led to greater success on posttests. The research questions asked whether there was a relationship between learner repair in response to interactional feedback and learning of targeted forms as measured in a posttest, and whether this relationship was affected by the type of repair (learner repair vs. teacher repair). Finally, the author wished to see whether the effects of repair in response to feedback were still visible in a delayed posttest.

To test these questions, the Nassaji (2011) studied 42 adult intermediate ESL learners. Since there was no focus on one type of feedback or one grammar form to be learned, it was impossible to prepare a pretest. Instead, learners were asked to look at a set of drawings and write down the events shown in the drawings. Then, each learner met with a teacher and orally described the same pictures, attempting to keep his oral descriptions as close as possible to the written one. The teacher provided feedback in the form of recasts or elicitations as necessary,

and immediately after the oral part was finished, the learner was told to look again at his written description and make any corrections to it based on the feedback just provided by the teacher. Each learner came back after two weeks and was given the original uncorrected written description, and asked to correct it again for errors.

Nassaji (2011) found 150 instances of errors receiving correction that had been produced in both the learners' written and oral descriptions, and these errors formed the data to be analyzed. Learners were able to successfully correct more than half of the errors they had repaired during interaction with the teacher, and they were more likely to correct an error if they had repaired it during interaction than if they had not. There was no significant difference in the immediate posttest between self-repair and teacher repair, although the delayed posttest showed that errors that had been self-repaired were more likely to be remembered than errors that had been corrected by the teacher.

Although studies such as Nassaji's (2011) provide valuable insight into the effects of different types of uptake, more research is needed to determine the extent to which uptake may be representative of noticing or learning.

DISCUSSION

Factors Influencing Efficacy

The data collected from the studies described above show varying amounts of success in using recasts as corrective feedback. Looking deeper than results, however, motivates an examination of the factors that may have influenced these results. Many of the studies brought up important aspects regarding the nature of the recasting environment, the subject's preparedness, the instructor's delivery of the recast, and other elements. In this section, I will address six of the

most prominent factors influencing the efficacy of recasting, according to the studies that used tests to assess student improvement.

Many studies found that the *length of the recast* affected learners' success on tests. The instructor in Lyster and Izquierdo's (2009) study always used short recasts that involved no more than one noun phrase; this may have made the error more salient to the learner, and may have ultimately made the implicit feedback more explicit. Kim and Han (2007) found that learners were able to recognize gaps between their utterance when the recast was simple rather than complex. Loewen and Philp (2006) also found that learners were more likely to receive higher test scores if they were presented with recasts consisting of five morphemes or fewer, and no more than one change. An example of this type of recast appears below:

> S: somebody steal my paper (.) stolen
> T: someone stole your paper?
> (Loewen & Philp, p. 550):

In this exchange, the instructor used a short response that corrected only one aspect of the learner's utterance.

Learners were more likely to notice their errors and correct them if they received *individualized attention*. As Han (2002) pointed out, this type of environment is most common in a laboratory setting, in which there are generally fewer students than in a classroom. She described the setting as "akin to a clinic to which students went for symptomatic treatment" (p. 568). Han (2002) and Lyster and Izquierdo's (2009) studies, which both showed recasts to be successful, studied learner groups composed of 4 learners and 12 learners, respectively. These were small class sizes compared to the average classroom, and the individualized attention may have improved the learners' success rate. Additionally, Nabei and Swain (2002) found that Shoko performed better on grammaticality judgment questions that were taken from recast

episodes in a group setting, rather than a teacher-fronted setting. They found that "recasts provided in group interaction [rather than teacher-fronted interaction] were more likely perceived 'accurately' as correction" (p. 58). Therefore, it appears Shoko responded more to recasts that were directed at her; even if she was part of a group, this was more individualized attention than the entire class.

Han (2002) argued that systematic recasts *focusing only on one form*, such as tense, make a correction more salient to the learner. In her study, Han found that recasts focused almost completely on tense appeared to make the learner more aware of the pedagogical focus of the instruction. The exit questionnaires in Lyster and Izquierdo's (2009) study also revealed that the subjects in their study were aware of receiving feedback only on grammatical gender. In this case, the implicit recasting had effectively become explicit. In contrast, Loewen and Philp (2006) did not limit the focus of the recasts in their study; learners received and were tested on a variety of morphosyntactic and phonetic recasts. Perhaps due to this lack of consistency, the learners in this study did not show a significant testing difference between those who received recasts and those who received other forms of corrective feedback.

Learners appeared to be more successful when recasts were provided in an *intensive environment* in which they could practice the forms regularly and receive a high amount of feedback. For example, in Han's (2002) study, instruction took place in eight classes over a period of four weeks, and the subjects received recasts as their only form of instruction. Lyster and Izquierdo's (2009) study was similarly intensive in that students attended additional classes outside of their regular French class, which focused on grammatical gender through exclusively recasts or using prompts. A lack of intensity may have also influenced the results of Nabei and Swain's (2002) study on Shoko. In this classroom, the teacher did not use a lot of corrective

feedback in general. When recasts were used, there was often no chance for students to process the feedback because the teacher would continue the topic without pausing for the student to notice any kind of correction. The fact that the instructor's recasts were infrequent and not carried out in a form-focused way may explain why Shoko seemed rarely aware of form-based corrections.

Another key factor that appears to influence the learner's successful test scores is whether the learner is *developmentally ready* to acquire the form being recast. Loewen and Philp (2006) found in their study that successful uptake of a recast did not result in higher test scores. The researchers stated that although learners repeated the recast, the form may have represented new knowledge to the listener, which he or she may not have been able to use in production. Han (2002) also emphasized that her subjects were chosen for the study based largely upon the fact that they were upper-intermediate level English learners. Therefore, they had generally acquired the knowledge of when to use present and past tense forms, but they lacked control of these forms. Lyster and Izquierdo (2009), too, selected subjects who were at the intermediate level of French, and who therefore were familiar with French grammatical gender forms before they participated in the study.

A final relevant factor that may influence efficacy of recasts relates to *the learner's orientation*. Lyster and Izquierdo (2009) cited the following explanation by Ellis and Sheen (2006):

> If learners treat language as an object to be studied, then they may detect the corrective force of recasts and thus derive negative evidence from them. But if they act as language users and treat language as a tool, then they are less likely to see recasts as corrective (cited in Lyster & Izquierdo, 2009, p. 483).

The researchers found the subjects in their study to be extremely focused on acquiring form, as these were learners who had volunteered to attend five extra laboratory sessions that were advertised as being specifically designed to help them learn French grammatical gender. Thus, the researchers posited that their subjects were indeed treating language as "an object to be studied." Mackey et al. (2006) cited a study by Sheen (2004), in which Sheen found that the number of both recasts and repairs was much higher in EFL and ESL classroom contexts than in immersion contexts. The researcher hypothesized that this difference may have been due to the students in the ESL and EFL classes being more oriented to attend to linguistic form rather than the students in the immersion classes, who were probably more focused on meaning. Unfortunately, it is difficult to objectively assess the learners' orientation to the material in most of the studies described in this paper, because more detailed information on the subjects' backgrounds is not provided.

Potential Problems and Limitations

Many problems with the studies on recasts apply not only to this type of corrective feedback, but also to studies on all types of corrective feedback. This section will discuss some of the more prominent issues that have come up in multiple studies on corrective feedback.

One of the most noticeable problems with studies on recasts to date is the lack of a unified definition of recasts. As Long (2007) pointed out, binary oppositional terms such as *corrective/noncorrective*, *simple/complex*, *complete/partial*, etc., should be clear and consistent across studies. Sheen (2006) also addressed this problem, pointing out that recasts in ESL/EFL contexts are often much more explicit than recasts in immersion classrooms, which tend to be

more like confirmation checks and are often followed by topic continuation. Despite the different discourse functions of these recasts, they are generally not viewed separately in the literature.

Another potential problem relating to assessing learners' acquisition of the language is the role of delayed posttests. As described in the studies of Nassaji (2009), and Nabei and Swain (2002), the immediate posttests from these studies did not show noticeable differences between learners' acquisition when comparing the recast group with other feedback groups. It was only on the delayed posttests that differences began to appear among the different groups. However, Long (2007) cautioned against putting too much stock in delayed posttest results. He argued that while it is necessary to study the durability and stability of change achieved due to recasts, the "importance of demonstrating lasting effects for recasts or any other (intentional or unintentional) instructional event may be somewhat overstated in the SLA literature" (p. 112). For example, higher scores on delayed posttests may be in part a result of continued exposure, such as a learner's continued study of the language, rather than solely the result of feedback that the learner has been processing from an interaction that took place weeks earlier. In short, there is a need for a more reliable way to test long-term effects of recasts.

Another significant problem of the studies testing learner improvement in response to corrective feedback is that data for these studies has generally been gathered from laboratory environments, not from a typical classroom. The issue with this method of gathering data is that the intensive, consistent, and individualized characteristics of recasts that these lab results have shown to be most effective are not the classroom norm. Nassaji's (2011) study approached a more naturalistic corrective feedback setting in that the instructor was allowed to provide feedback on any type of error. However, the instructor was only allowed to use two types of feedback (elicitation or recast), and the setting was still a dyadic laboratory environment. Typical

language classroom settings are more like those described in Loewen and Philp's (2006) and Nabei and Swain's (2002) observational research, in which the instructors provided recasts in a variety of language forms; not all of the recasts were five morphemes or fewer, or focusing on one grammatical form. One of the larger hindrances of using more natural lesson environments is the impossibility of predicting which forms the learners will receive feedback on, which in turn makes it impossible to pretest learners on those forms. Thus, the pretest-posttest-delayed posttest has not yet been successfully applied to a truly "organic" classroom setting.

A final problem with the studies comparing implicit feedback of recasts with more explicit feedback is the fact that most studies showing explicit feedback to be the more successful of the two take place over a relatively short time, usually with laboratory sessions lasting only two or three hours. Additionally, these studies usually focus on a simple grammatical target, such as articles or formation of regular past tense. Long (2007) argued that the design of these studies inherently favors more explicit learning, so the results of these studies should therefore be treated with caution.

CONCLUSION

The studies described above provide strong evidence that recasts are effective in general and do improve student accuracy on tests, although recasts may not necessarily be more effective than more explicit forms of corrective feedback such as metalinguistic feedback or elicitations. Studies have also shown that learners were generally able to successfully identify recasts, although they seem to be less successful at distinguishing corrective recasts from repetitions. Factors such as length of recast, individualized attention, consistency, intensity, developmental readiness, and learner orientation appeared to influence learners' later success on written and oral

tests. It is worth noting that many of the factors that influence learner success also increase the recast's saliency, which brings attention to the fact that recasts vary considerably in how implicit or explicit they are, and that perhaps the more explicit the recast is, the more effective it ultimately is.

More research into the effectiveness of recasts in comparison to other forms of feedback is still needed. While studies so far have provided valuable insight into the factors potentially influencing the efficacy of recasting, a unified definition of recasts, a better system for testing long-term effects of recasts, a way to test practical application of the factors described above in a typical classroom, and a greater number of longitudinal studies investigating acquisition of more complicated grammatical forms are some of the ways that future research can shed more light on this topic.

REFERENCES

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

Doughty, C. J., & Long, M. H. (2003). The scope of inquiry and goals of SLA. In C. J. Doughty & M. H. Long (Eds.), *Handbook of second language acquisition* (pp. 3-15). Oxford, England: Blackwell.

Ellis, R. (2005). Measuring implicit and explicit knowledge of a second language: A psychometric study. *Studies in Second Language Acquisition, 27*, 141-172.

Ellis, R., & Sheen, Y. (2006). Re-examining the role of recasts in L2 acquisition. *Studies in Second Language Acquisition, 28*, 575–600.

Gass, S., & Selinker, L. (2008). *Second Language Acquisition: An Introductory Course* (3rd ed.). New York, NY: Routledge.

Han, Z. (2002). A Study of the Impact of Recasts on Tense Consistency in L2 Output. *Tesol Quarterly, 36*, 543-572.

Kim, J., & Han, Z. (2007). Recasts in communicative EFL classes: do teacher intent and learner

interpretation overlap? In A. Mackey (Ed.), *Conversational Interaction in Second Language Acquisition* (pp. 269-297). Oxford, England: Oxford University Press.

Loewen, S., & Philp, J. (2006). Recasts in the Adult English L2 Classroom: Characteristics, Explicitness, and Effectiveness [Electronic version]. *The Modern Language Journal, 90*, 536-556.

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

Lyster, R., & Izquierdo, J. (2009). Prompts Versus Recasts in Dyadic Interaction. *Language Learning, 52*, 453-498.

Lyster, R., & Ranta, L. (1997). Corrective Feedback and Learner Uptake: Negotiation of Form in Communicative Classrooms. *Studies in Second Language Acquisition, 19* (1), 37-66.

Nabei, T., & Swain, M. (2002). Learner Awareness of Recasts in Classroom Interaction: A Case Study of an Adult EFL Student's Second Language Learning. *Language Awareness, 11*, 43-63.

Nassaji, H. (2009). Effects of Recasts and Elicitations in Dyadic Interaction and the Role of Feedback Explicitness. *Language Learning, 59*, 411-452.

Nassaji, H. (2011). Immediate learner repair and its relationship with learning targeted forms in dyadic interaction. *System, 39*, 17-29.

Panova, I., & Lyster, R. (2002). Patterns of Corrective Feedback and Uptake in an Adult ESL Classroom. *TESOL Quarterly, 36*, 573-595.

Rezaei, S., & Derakhshan, A. (2011). Investigating Recast and Metalinguistic Feedback in Task-based Grammar Instruction. *Journal of Language and Teaching Research, 2*, 655-663.

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.