

Prompts Versus Recasts in Dyadic Interaction

Roy Lyster

McGill University

Jesús Izquierdo

Universidad Juárez Autónoma de Tabasco

This study investigated the differential effects of prompts and recasts, in the context of dyadic interaction, on the acquisition of grammatical gender by adult second language learners of French. Participants were 25 undergraduate students enrolled in an intermediate-level French course at an English-speaking university. All students were exposed in class to a 3-hr form-focused instructional treatment distributed over 2 weeks and were then randomly placed in either the recast or prompt group. On two occasions outside of class, individual students participated in three different oral tasks during dyadic interaction with a native or near-native speaker of French who, following learner errors in grammatical gender, provided feedback in the form of either prompts or recasts. Pretests and immediate and delayed posttests included two oral production tasks and a computerized reaction-time binary-choice test. Results of repeated-measures ANOVA showed that both groups significantly improved accuracy and reaction-time scores over time, irrespective of feedback type. We conclude that learners receiving recasts benefited from the repeated exposure to positive exemplars as well as from opportunities

This study was funded by the Social Sciences and Humanities Research Council of Canada (410-2006-2212 and 410-2002-0988). Parts of this study were presented at the annual meeting of the American Association for Applied Linguistics in Washington DC on March 31, 2008, and at the annual meeting of the Canadian Association of Applied Linguistics in London, Ontario on May 31, 2005. We are grateful to the participating teacher and students, to José Correa for his assistance with the statistical analyses and to the following research assistants for contributions to various phases of this research: Susan Ballinger, Sophie Beaudoin, Andréanne Gagné, Carl Larouche, Elisa David, Andrea Sterzuk, Jean-Sébastien Vallée, and Yingli Yang. We gratefully acknowledge Iliana Panova, Leila Ranta, and four anonymous *Language Learning* reviewers for their helpful comments on earlier versions of this article.

Correspondence concerning this article should be addressed to Roy Lyster, Department of Integrated Studies in Education, McGill University, 3700 McTavish Street, Montreal, QC, Canada H3A 1Y2. Internet: roy.lyster@mcgill.ca

to infer negative evidence, whereas learners receiving prompts benefited from the repeated exposure to negative evidence as well as from opportunities to produce modified output.

Keywords recasts; prompts; corrective feedback; dyadic interaction; grammatical gender; form-focused instruction; focus-on-form; French as a second language

The relative effectiveness of different types of feedback continues to attract attention in the field of second language acquisition (SLA), equally so for researchers with theoretical interests in the cognitive processes triggered by feedback and for researchers and practitioners alike interested in how feedback can be effectively yet seamlessly integrated into classroom interaction in ways that drive second language (L2) development forward. A considerable amount of recent research, both in and out of classrooms, has concerned *recasts*: implicit reformulations of learners' nontarget utterances (for helpful reviews of this literature, see Ellis & Sheen, 2006; Nicholas, Lightbown, & Spada, 2001).

Knowledge about the characteristics and effectiveness of recasts continues to expand, and it is now widely accepted that their effectiveness is constrained by a wide range of variables. Specifically in classroom settings, the extent to which teachers' intentions and learners' perceptions overlap is known to affect recast effectiveness (Mackey et al., 2007), as is the interactional context in which recasts are provided. For example, Oliver and Mackey (2003) found in child English as a second language (ESL) classrooms that learners produced significantly more modified output following recasts in explicit language-focused exchanges than in exchanges that were content based, management related, or communicative in nature. In addition, recasts of phonological and lexical errors are more noticeable than recasts of errors in morphosyntax (Carpenter, Jeon, MacGregor, & Mackey, 2006; Han, 2008; Lyster, 2001; Mackey, Gass, & McDonough, 2000), as are recasts that are short and contain minimal changes (Egi, 2007; Philp, 2003; Sheen, 2006). Recasts benefit developmentally ready learners more than unready learners (Mackey & Philp, 1998) and learners with high accuracy scores in their use of the target forms more than learners achieving low accuracy scores (Ammar & Spada, 2006). Trofimovich, Ammar, and Gatbonton (2007) found not only that higher proficiency learners benefit more than lower proficiency learners from recasts but also that, independent of L2 proficiency, other learner characteristics associated with accurate L2 production after hearing a recast include phonological memory, attention control, and analytic ability.

Ellis and Sheen (2006) argued that recasts have received an inordinate amount of attention in SLA research, owing in large part to their ubiquity and consequent availability for study. They continued:

Despite the strong theoretical claims regarding the acquisitional potential of recasts, there is no clear evidence that recasts work better for acquisition than other aspects of interaction such as models, prompts, or explicit corrective strategies. Indeed, there is some evidence that the last two of these are more effective than recasts. (p. 597)

Similarly, Lyster (1998, 2002, 2007) argued that, especially in classroom settings that are more meaning-oriented than form-oriented, other types of feedback might serve more effectively than recasts as negative evidence (i.e., information about the incorrectness of an utterance; see Gass, 2003). A growing number of studies have recently investigated the effects of different types of feedback on L2 development in both classroom and laboratory settings but with different outcomes across these different settings.

Classroom Studies

Recasts are, by far, the most frequently used feedback across a spectrum of classroom settings, including elementary immersion classrooms (Lee, 2006; Lyster & Ranta, 1997; Mori, 2002), high school foreign language classrooms (Havranek, 2002; McNulty, 2007; Tsang, 2004), university-level foreign language classrooms (Doughty, 1994; Havranek, 2002; Roberts, 1995), Spanish literature courses (Zyzik & Polio, 2008), and adult ESL classrooms (Ellis, Basturkmen, & Loewen, 2001; Panova & Lyster, 2002). Classroom intervention studies have led the way in comparing recasts with other types of feedback and, to the best of our knowledge, there is as yet no published research showing that recasts are more effective than other types of feedback.

An alternative type of feedback that has been compared with recasts in classroom settings is referred to as *prompts* (Lyster, 2004, 2007; Lyster & Mori, 2006, 2008; Ranta & Lyster, 2007), because they provide signals that prompt learners to self-repair rather than providing them with a correct reformulation of their nontarget utterance, as do recasts. Prompts include clarification requests, repetition of learner error, metalinguistic clues, and elicitation moves (see Appendix A for examples). Several classroom studies have shown prompts to be more effective than recasts; for example, Havranek and Cesnik (2001) found repair following prompts to be the most effective feedback combination

in a range of English as a foreign language (EFL) classrooms. In a classroom study of adult ESL learners, Ellis, Loewen, and Erlam (2006) compared the effects of recasts versus prompts on students' use of the simple past tense in English. Operationalizing prompts as a repetition plus a metalinguistic clue (e.g., "you need the past tense"), they found significantly superior effects for prompts over recasts on delayed posttest measures. In a similar ESL context, Ellis (2007) compared the effects of recasts and prompts (again operationalized as a repetition plus a metalinguistic clue) on the acquisition of past tense *-ed* and comparative *-er* in English. He found that prompts were overall more effective than recasts but more so for the comparative than for past tense forms.

Higher accuracy rates for prompts over recasts were also found in two classroom studies conducted in elementary school settings. The first examined the differential effects of prompts and recasts in a form-focused intervention study targeting the acquisition of grammatical gender in French by fifth-grade immersion students (Lyster, 2004). In addition to implementing an instructional unit on grammatical gender, the three participating teachers each interacted with students in a specific way that permitted comparisons of three oral feedback options: prompts, recasts, and no feedback. The comparison group received no form-focused instruction or any preplanned feedback on grammatical gender. The analysis of eight proficiency measures (i.e., two oral tasks and two written tasks administered immediately following the instructional unit and then 2 months later) showed that the group receiving prompts distinguished itself by being the only group to significantly outperform the comparison group on all eight measures. The recast group significantly outperformed the comparison group on five of the eight measures, whereas the instruction-only group (receiving no feedback) significantly outperformed the comparison group on four of the eight measures, suggesting that recasts were more effective than no feedback, but only marginally so.

Ammar and Spada (2006) investigated the differential effects of prompts and recasts in form-focused instruction in three sixth-grade intensive ESL classrooms over a 4-week period. The form-focused intervention targeted third person possessive determiners in English (*his* and *her*), which are known to be difficult for francophone learners of English even after many years of ESL instruction (White, 1998). Students in all three classes received form-focused instruction, which included metalinguistic information and both controlled and communicative practice activities. During the practice activities, one class received feedback in the form of recasts, another received prompts, and the third received no feedback. Results of pretests, immediate posttests, and delayed posttests showed that all three groups benefited from the form-focused

instruction and that the two feedback groups benefited the most, outperforming the control group on both immediate and delayed oral posttests. The group receiving prompts significantly outperformed the recast group on written and oral posttests. Prompts were especially effective for learners who had pretest scores below 50%, whereas learners with pretest scores above 50% benefited similarly from both recasts and prompts.

Loewen and Nabei (2007) compared the effects on question formation in English of recasts and two specific types of prompts—one considered implicit (i.e., clarification requests: “Pardon?”) and the other considered more explicit (i.e., metalinguistic feedback: “Can you think about your question again?”). They found no significant difference across feedback types and suggested that one reason might be that the classroom treatment had lasted only 30 min. In their study with adult ESL learners, Loewen and Philp (2006) found that, on immediate posttests, prompts led to an accuracy rate of 75%, whereas recasts led to an accuracy rate of only 53%. Seemingly at odds with this finding, however, they concluded that the relatively low success rate for recasts “should be encouraging for classroom teachers,” because “a recast is time-saving, less threatening to student confidence, and less intrusive to the flow of interaction than, for example, elicitation of self repair” (p. 551). No supporting evidence was provided from their data to illustrate this claim, but the suggestion that recasts are “pedagogically expeditious” (Loewen & Philp, p. 551) arguably finds some support in common sense, especially if the effectiveness of recasts is not contingent on any follow-up moves on the part of students. Questionable, however, is the pedagogical soundness of teacher discourse that neither requires nor invites learner responses. Also open to discussion is whether focus-on-form techniques that simply speed up instruction should be considered more effective than those that actually increase target language accuracy.

Finally, Sheen’s (2007) classroom study is also noteworthy because it compared the effects of two distinct feedback types on ESL learners’ use of English articles. She compared recasts with metalinguistic corrections, which included provision of the correct form (unlike prompts, which, by definition, withhold correct forms), followed by metalinguistic explanation (e.g., “You should use the definite article ‘the’ because you’ve already mentioned ‘fox’”). The metalinguistic group significantly outperformed the recast and control group, whereas the recast group did not perform significantly better than the control group.

Han (2008) argued that for recasts to have any effect on morphosyntactic development in classroom settings, teachers would first need to have an unambiguous understanding of the meaning of the learner’s utterance and then, in those instances, provide recasts in an ongoing and systematic manner, focusing

on one grammatical morpheme. Such was likely the intent of the oft-cited study by Doughty and Varela (1998), which found “corrective recasting” more effective than no feedback. However, their study did not directly examine the effects of recasts, because the latter were used solely as secondary moves in the event that the primary move, which was a prompt that repeated verbatim the learner’s error (and thus not a recast), failed to elicit self-repair. The teacher consistently used repetition to draw attention to the error and then recast only when students made no attempt at repair. As Doughty and Varela observed, by the beginning of the second of three treatment sessions, “students were beginning to self-correct before the teacher had the opportunity to recast” (p. 135).

The findings of the intervention studies by Lyster (2004), Ammar and Spada (2006), Ellis (2007), and Ellis et al. (2006) indicate that classroom learners who are prompted to retrieve more targetlike forms are more likely to retrieve these forms during subsequent processing than learners provided with recasts of these forms. Important to stress is that these studies all included measures of target language development over time rather than considering immediate learner responses. Otherwise, given the ostensibly binary nature of target features such as gender attribution in French and possessive determiners in English, the rate of immediate repair could arguably have been affected by a seemingly simple on-the-spot computation (i.e., if *his* is wrong, then it must be *her*). Whether learners could be easily led in this way to repair their errors immediately following feedback, however, was not used as a measure of effectiveness in these intervention studies. Instead, whether learners could retrieve more targetlike forms at a later point in time was examined.

Laboratory Studies

For the most part, research demonstrating the effectiveness of recasts has been conducted in laboratory settings, where variables can more easily be controlled than in classroom settings, allowing for feedback to be delivered intensively in consistent ways on specific linguistic targets. Overall, these laboratory studies have shown positive effects for recasts on L2 development (e.g., Han, 2002; Ishida, 2004; Iwashita, 2001; Leeman, 2003; Long, Inagaki, & Ortega, 1998; Mackey & Philp, 1998; McDonough & Mackey, 2006).

Many laboratory studies demonstrating the effectiveness of recasts, however, lack rigorous comparisons with other clearly identified types of feedback. In Ishida’s (2004) study of recasts and their effects on the acquisition of aspectual forms in Japanese as a foreign language, there was no control group because the study used a time-series design to track the progress over time of the

four participants receiving recasts. In the laboratory studies by Han (2002) and McDonough and Mackey (2006), the recast groups were compared to control groups receiving no feedback. In Long et al.'s (1998) study, participants receiving recasts during experimentally controlled interaction were compared with participants receiving models, but models were provided to learners before they spoke and, thus, are not a comparable type of semantically contingent feedback. The studies by Mackey and Philp (1998) and Iwashita (2003) compared the progress made by participants engaged in interaction with recasts to progress made by participants engaged in interaction without recasts. Interaction without recasts was called "negotiated interaction" (Mackey & Philp) or "negotiation moves" (Iwashita), which contained moves such as confirmations and confirmation checks. Yet, Mackey and Philp drew on their own data to illustrate that, in some cases, recasts "are part of negotiation sequences and function as confirmation checks" (p. 342). Similarly, Loewen and Philp (2006) stated "Recasts provided in the context of conversation are often part of negotiation sequences and function as confirmation checks" (p. 540). Because recasts are forms that perform confirming functions, they cannot be categorically differentiated from confirmations and confirmation checks (see Braid, 2002; Egi, 2007; Ellis & Sheen, 2006; Lyster, 1998). Comparing recasts with negotiation that contains confirmations and confirmation checks continues to be problematic in research on interaction that aims to assess the effectiveness of recasts (Lyster, 2007).

Three laboratory studies that compared recasts to other clearly defined types of feedback are those by Carroll and Swain (1993), McDonough (2007), and Leeman (2003). Carroll and Swain's laboratory study compared the effects of explicit correction, recasts, two types of prompts, and no feedback on the acquisition of dative alternation in English. The prompts involved either "explicit utterance rejection" (telling participants they were wrong) or "implicit metalinguistic feedback" (asking participants if they were sure their response was correct whenever they made a mistake). All groups receiving feedback significantly outperformed the control group, whereas the explicit correction group outperformed the prompt groups on two measures and outperformed the recast and prompt groups on two other measures. The recast and prompt groups did not perform significantly differently from one another. McDonough investigated the effects of recasts and clarification requests relative to no feedback on the emergence of simple past activity verbs in English. She found no significant differences overall between recasts and clarification requests, although both feedback types were more effective than no feedback. A post hoc analysis of the data revealed that clarification requests, but not recasts, had a significant effect on the emergence of progressive activity verbs as well

as simple forms, leading to the tentative conclusion that “clarification requests may impact several forms across developmental stages simultaneously, whereas recasts may have a more concentrated impact on a single developmental feature” (McDonough, p. 337). Leeman also compared recasts with another type of feedback—a move referred to as negative evidence that contained a repetition of the error preceded by “but you said . . .,” but without opportunities for immediate repair in either condition. There were no statistically significant differences between the recast and negative evidence groups. However, the recast group, but not the negative evidence group, outperformed the control group but the negative evidence group did not, leading Leeman to conclude that the effectiveness of recasts derives more from positive evidence than from negative evidence.

Immediate Learner Repair

Classroom studies have revealed discrepant amounts of immediate repair following recasts according to instructional setting. Specifically, infrequent repair following recasts was observed in French immersion classrooms in Canada (Lyster & Ranta, 1997), English immersion classrooms in Korea (Lee, 2006), adult ESL classrooms in Canada (Panova & Lyster, 2002), EFL classrooms in Hong Kong secondary schools (Tsang, 2004), and Spanish foreign language secondary classrooms in the United States (McNulty, 2007). More frequent repair following recasts was observed in Japanese immersion classrooms in the United States (Mori, 2002), adult ESL classrooms in New Zealand (Ellis et al., 2001), and adult EFL conversation classes in Korea (Sheen, 2004).

As to whether repetitions of recasts are of any value, however, the verdict is still out. On the one hand, Mackey et al. (2000) found that when learners repeated a recast, they were more likely to have correctly perceived its corrective intention. In EFL classrooms, Havranek and Cesnik (2001) demonstrated through follow-up language tests that recasts eliciting immediate repetition by learners were more effective than recasts not eliciting immediate repetition. In ESL classrooms, Loewen (2005) found that learner repair in language-related episodes was significantly associated with subsequent accuracy scores on posttests. Others have suggested that, at the very least, a repetition of a recast might be “evidence that learners are noticing the feedback” (Lightbown, 2000, p. 447) or “facilitative of acquisition” (Ellis et al., 2001). In contrast, however, Loewen and Philp (2006) found in their classroom study with adult ESL learners that successful repair following recasts did not predict accuracy on posttest scores. Studies conducted in laboratory settings also showed that the potential

effectiveness of recasts may be unrelated to modified output following recasts (Leeman, 2003; Mackey & Philp, 1998; McDonough, 2007; McDonough & Mackey, 2006). It appears to be the case that some learners more than others have the requisite cognitive predisposition to compare their ill-formed utterances with recasts and to benefit without producing modified output. Such a predisposition may have been characteristic of the adult ESL learners enrolled in private language schools participating in the studies by Loewen and Philp and by Mackey and Philp and of the university-level EFL students participating in the studies by McDonough and by McDonough and Mackey. It could even be the case, however, that classroom learners who are relatively form-oriented do repeat recasts more than researchers are able to observe, by incorporating them as part of their "private speech" (i.e., "oral language addressed by the student to himself or herself"; Ohta, 2002, p. 52) as Ohta showed in her study of adult learners wearing lapel microphones in a form-oriented foreign language classroom.

Unlike recasts, the effectiveness of prompts has been associated with the modified output they elicit. For example, in the small-scale study by Nobuyoshi and Ellis (1993), the two learners who responded with self-completed repair following clarification requests improved more than the learner who did not modify his output following the feedback. In a more recent and larger scale study, McDonough (2005) found that modified output following clarification requests was a significant predictor of L2 development. Similarly, Loewen and Philp (2006) found that successful repair following prompts (but not following recasts) predicted accuracy on posttest scores. This would explain Leeman's (2003) finding that when no opportunities for immediate repair were provided, recasts proved more effective than repetition of error.

Theoretical Perspectives

Recasts have been hypothesized to create ideal opportunities for learners to notice the difference between their interlanguage forms and targetlike reformulations (e.g., Doughty, 2001; Long, 1996, 2007). For example, Long (1996) argued that conversational moves such as recasts benefit L2 development because they provide learners with a primary source of negative evidence. He argued that because recasts preserve the learners' intended meaning, they free up cognitive resources that would otherwise be used for semantic processing. Thus, with meaning held constant, recasts have the potential to enable learners to focus on form and to notice errors in their interlanguage production (see also Doughty). Others have argued, however, that this is the case only in

form-oriented classrooms, in which the emphasis on accuracy primes learners to notice the corrective function of recasts (Ellis & Sheen, 2006; Lyster, 2007; Lyster & Mori, 2006, 2008; Nicholas et al., 2001). Braidí (2002) and Leeman (2003) suggested that recasts serve as exemplars of positive evidence, which facilitates the encoding of new target representations. Similarly, Ellis and Sheen recently argued, "It is not possible to say with any certainty whether recasts constitute a source of negative evidence (as it is often assumed) or afford only positive evidence, as this will depend on the learner's orientation to the interaction" (p. 596). In addition, the ability of classroom learners to infer negative evidence from recasts by comparing them with their nontarget output depends on whether the discourse context in which the recasts are delivered enables learners to perceive them as disapproving (rather than approving) the use of nontarget forms. Acknowledging that recasts are well suited to communicative classroom discourse because they maintain the flow of communication and keep students' attention focused on meaning, Lyster argued that in many discourse contexts occurring during communicative and content-based instruction, prompts may be better suited than recasts to provide negative evidence precisely because they incite learners to switch their attention momentarily away from meaning toward form.

In addition to the negative evidence they provide, the effectiveness of prompts can be explained through skill acquisition theory, which describes L2 learning as a gradual change in knowledge from declarative to procedural mental representations (e.g., DeKeyser, 1998, 2001). The transformation of declarative knowledge into procedural knowledge involves a transition from controlled processing, which requires a great deal of attention and use of short-term memory, to automatic processing, which operates on automatized procedures stored in long-term memory (Shiffrin & Schneider, 1977). The transition from controlled to more automatic processing results from repeated practice in transforming declarative representations into production rules in contexts clearly linking form with meaning (DeKeyser, 1998, 2007). Designing practice activities that are both communicative in purpose and controlled in the sense of requiring the use of specific target forms, however, is challenging in any instructional context, and this is where prompts play a central role (Lyster, 2007). Given their aim to elicit modified output, prompts serve to scaffold opportunities for controlled practice in the context of communicative interaction. As with other types of practice, prompts aim to improve control over already-internalized forms by providing opportunities for "pushed" output, hypothesized by Swain (1985, 1988) to move interlanguage development forward, and by assisting learners in the transition from declarative to procedural knowledge (de Bot,

1996; DeKeyser, 1998; Ranta & Lyster, 2007). De Bot argued that L2 learners benefit more from being pushed to retrieve target language forms than from merely hearing the forms in the input, because retrieval and subsequent production stimulate the development of connections in memory. The results of research in experimental psychology on the “generation effect” also predict, for similar reasons, that prompts will be more effective than recasts. This line of experimental research has consistently found that learners remember information better when they take an active part in producing it, rather than having it provided by an external source (e.g., Clark, 1995; deWinstanley & Bjork, 2004).

Grammatical Gender

The linguistic feature that was targeted in the present study by form-focused instruction and interactional feedback was gender attribution in French, which is a well-known area of difficulty for anglophone learners of French (e.g., Tucker, Lambert, & Rigault, 1977). Grammatical gender markers are not salient in classroom discourse, in spite of their frequency, nor do they convey, in the case of inanimate nouns, any semantic distinctions (see Ayoun, 2007). Moreover, grammatical gender does not exist in English. Carroll (1989) proposed that native speakers of French acquire and process determiners and nouns as coindexed chunks, whereas anglophone learners of French acquire and process French determiners and nouns as they do in their first language (i.e., as distinct syntactic words and independent phonological units).

Influenced by the many French grammarians who claim grammatical gender is arbitrary and unsystematic in the case of inanimate nouns (e.g., Bosquart, 1998; Jacob & Laurin, 1994), teachers encourage students to learn gender attribution on an item-by-item basis, and they often do so through incidental reminders (Lyster, 2007). Yet there exists considerable evidence that gender attribution is largely rule-driven and based on word-internal structural properties. Tucker et al. (1977) found that grammatical gender entails a rule-governed subsystem, in which “distinctive characteristics of a noun’s ending and its grammatical gender are systematically related” (p. 64). Lyster (2006) corroborated this finding in a corpus analysis of nearly 10,000 nouns in *Le Robert Junior Illustré*. Operationalizing noun endings as orthographic representations of rhymes, which consist of either a vowel sound (i.e., a nucleus) in the case of vocalic endings or a vowel-plus-consonant blend (i.e., a nucleus and a coda) in the case of consonantal endings, Lyster classified noun endings as reliably masculine, reliably feminine, or ambiguous by considering as reliable predictors of

grammatical gender any noun ending that predicted the gender of at least 90% of all nouns in the corpus with that ending. Results revealed that 81% of all feminine nouns and 80% of all masculine nouns in the corpus proved to be rule governed, having endings whose orthographic representations systematically predict their gender.

At least two classroom studies have shown that drawing learners' attention to the relationship between noun endings and grammatical gender in this way leads to significant improvement in their ability to accurately assign grammatical gender. Harley's (1998) study showed that young second-grade learners made significant progress on all but one measure, which was composed of unfamiliar nouns. She concluded that the learners' progress may have involved exemplar-based learning rather than rule-based learning. Lyster (2004) found that the fifth-grade learners in his aforementioned classroom study made significant progress on all measures, including one with low-frequency unfamiliar nouns. He concluded that the learners developed abstract rule-based knowledge of grammatical gender, not merely exemplar-based knowledge.

With respect to gender attribution in French, it may be the case that improving L2 learners' accuracy through increased awareness of the predictive value of noun endings results more from rule-based learning than from exemplar-based binary choices. In French, unlike some other Romance languages, there are hundreds of noun endings that predict grammatical gender. Assigning grammatical gender accurately in French based on knowledge of noun endings does not simply involve a single binary choice but rather an iterative process of analysis and synthesis of interrelated structural elements: (a) nouns and their constituent endings, operationalized as orthographic representations of rhymes; (b) noun endings as predictors of grammatical gender; (c) nouns and their inherent gender attribution, manifest in the noun's constituent ending and also in a separate determiner (Carroll, 1989). The goal of the present study was to assess the effects of form-focused instruction and different types of feedback on L2 learners' ability to access these associative patterns and to perform the necessary computations online.

Research Question and Predictions

In classroom studies that have compared recasts and prompts, prompts have led overall to higher rates of accuracy on posttest measures. In contrast, laboratory studies have generally yielded positive results for recasts, but only when there has been either no control group (e.g., Ishida, 2004) or a control group receiving (a) no feedback (e.g., Han, 2002; McDonough & Mackey, 2006), (b)

models (e.g., Long et al., 1998), (c) feedback in the form of error repetition but with no opportunity for immediate repair (Leeman, 2003), or (d) a mixture of interactional moves referred to monolithically as “negotiation” or “negotiated interaction” (e.g., Iwashita, 2003; Mackey & Philp, 1998). In her comparison of recasts and clarification requests in a laboratory setting, McDonough (2007) found both feedback types to be equally effective. Similarly, Carroll and Swain (1993) found prompts and recasts to be equally effective but that participants receiving explicit correction performed even better. The present study adds to this body of research by comparing recasts with prompts provided in the context of dyadic interaction while controlling the output variable: Immediate learner repair was allowed in the prompt condition but not in the recast condition. The exclusion of opportunities for modified output in the recast condition and their inclusion in the prompt condition allowed us in our comparison of these feedback types to explore which combination is more effective: negative and positive evidence in the case of recasts, or, in the case of prompts, negative evidence without positive exemplars but with opportunities for self-repair.

We adopted an innovative design that allowed us to conduct the study both inside and outside the classroom. A form-focused instructional treatment was implemented in two classrooms by the same instructor so that all participating students received identical instruction on the target feature. The feedback treatments were then implemented outside the classroom in a laboratory setting in order to control for the provision of feedback more systematically and with more consistency than is possible in classroom settings in which individual learners cannot all receive equal amounts of feedback. In addition to regulating the provision of feedback, the laboratory setting allowed us to control (and measure) the amount of learner repair in the form of modified output following feedback. The inclusion of a classroom instructional treatment in both experimental conditions was crucial in the present study, given our interests in skill acquisition theory and the role of feedback as a pedagogical tool for automating the retrieval of already existing or emergent knowledge (Lyster, 2007; Ranta & Lyster, 2007). Our research question was formulated as follows:

In the context of dyadic interaction with a researcher, do adult L2 learners of French improve their accuracy in assigning grammatical gender more from prompts or from recasts?

Drawing on skill acquisition theory, the following prediction was made. 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. Recasts are

predicted to be less effective because they do not require a similarly deep level of processing entailing both retrieval and opportunities for production practice. Specifically with respect to accurate gender attribution during online production, the processing requires quick and coordinated access to previously disconnected elements, associating nouns with their constituent endings, noun endings with grammatical gender, and grammatical gender with nouns. One could thus predict, in accordance with the findings of Lyster's (2004) classroom study that also targeted grammatical gender in French, that learners with more opportunities to process these elements in a productive mode in response to prompts will benefit more than students provided with opportunities to process the same information in a receptive mode via recasting.

Method

Participants and Design

Twenty-five undergraduate university students, consisting of 21 females and 4 males, participated in this combined "classroom + laboratory" study. With the exception of one participant who was 49 years old, their ages ranged from 18 to 24, with a mean age of 21. Of the 25 participants, 21 were L1 speakers of English, 3 were bilingual (English/Armenian, English/Russian, English/Serbian), and 1 was a Japanese foreign exchange student. They were all undergraduate students in various faculties (i.e., Arts, Science, Management, and Religious Studies) at an English-speaking university in Quebec. They were enrolled in one of two sections of an elective intermediate-level French L2 course. According to the official course description, the course was "open to students in any degree program having an elementary knowledge of French" but was "not open to students from Quebec." Course content was described as "grammar review, comprehension, vocabulary development, selected readings, and group discussions." Both sections of this 13-week course (three 1-hr classes per week) were taught by the same instructor, who was a native speaker of French and had been teaching French L2 for 26 years at the university where the research took place.

The timeline of the instructional treatment and feedback treatments along with the testing sessions over the present study's 9-week period appears in Table 1. The study began just after the midway point of the winter semester and ended 9 weeks later; delayed posttests were administered 3 weeks after the course had ended. Because participation in this study required students to attend three testing and two treatment sessions outside of regular class time,

Table 1 Design and timeline

Week 1	Week 2	Week 3	Week 4	Week 5	Week 9
Pretest: <i>RT task A</i> <i>OI task A</i> <i>PD task A</i>	Classroom treatment: <i>La grande</i> <i>séduction</i>	Classroom treatment: <i>L'Odyssée . . .</i> Feedback treatment: <i>Riddles A</i> <i>OI task A</i> <i>PD task A</i>	Feedback treatment: <i>Riddles B</i> <i>OI task B</i> <i>PD task B</i>	Posttest 1: <i>RT task B</i> <i>OI task B</i> <i>PD task B</i>	Posttest 2: <i>RT task A</i> <i>OI task A</i> <i>PD task A</i> Exit questionnaire

Note. RT = binary-choice and reaction-time measures; OI = object-identification task; PD = picture-description task. A and B refer to the counterbalanced forms (the above schedule pertains to participants beginning with Form A and would be reversed for those beginning with Form B).

those who stayed involved in the study for its full 9 weeks were given a \$60 payment at the conclusion of the final testing session.

Classroom Instructional Treatment

The classroom treatment entailed a form-focused instructional unit created by the research team and implemented by the instructor in class for approximately 3 hr over a 2-week period. The unit included a 17-page student workbook designed to target a set of typically feminine endings (–*ie*, –*ine*, –*eine/–aine*, –*tion/–sion*, –*té*, –*se*, –*ence/–ance*, –*ette*, –*ure*, –*che*) and a set of typically masculine endings (–*in*, –*an/–ent*, –*age*, –*o/–ot*, –*eau*, –*eu*). The workbook was contextualized in a thematic unit about two Quebec films, *La grande séduction* (Pouliot, 2003) and *L'Odyssée d'Alice Tremblay* (Filiatrault, 2002). Synopses for each film that had appeared on the Internet for promotional purposes were used along with reviews that individuals had written and posted on the Internet. These reviews clearly indicated that one film was much more appreciated than the other, allowing the instructor to focus not only on grammatical gender but also on different genres and the fact that promotional blurbs are similar and always positive, whereas film reviews range from positive to negative, depending both on the film and the reviewer. These texts provided the context for a series of consciousness-raising exercises that students completed in pairs as a means of drawing their attention to noun endings that reliably predict grammatical gender. Table 2 displays the number of types and tokens of nouns with each target ending appearing in the materials (e.g., two occurrences of the word *sortie* count as one type but two tokens).

Table 2 Number of nouns with target endings occurring in the classroom instructional materials

	Types	Tokens
–ie	12	30
–ine, –eine/–aine	5	11
–tion/–sion/–gion	20	37
–té	21	37
–se	8	15
–ence/–ance	6	16
–ette	3	5
–ure	5	5
–che	5	8
–in	13	15
–an/–ant/–ent	23	13
–age	27	11
–o, –eau	12	25
–eu	14	5

In the first set of texts, which were about *La grande séduction*, the determiners and endings of target nouns were highlighted in bold to draw students' attention to the relationship between noun endings and gender attribution as they read and discussed the texts. An example follows:

Critique 3: Il y a **une** excellente performance de la part de Raymond Bouchard en particulier. C'est **une** superbe production et **un** excellent scénario. Quoi dire de plus? C'est **une** excellente comédie! 9/10

[Review 3: There is an excellent performance by Raymond Bouchard in particular. It's a superb production and an excellent screenplay. What more is there to say? It's an excellent comedy! 9/10]

Students were then given 15 target endings and asked to group together the nouns with those endings in the texts they had read and to identify whether the endings predicted masculine or feminine gender. In subsequent exercises, they were given three endings (–age, –eau, –tion) and asked to identify as quickly as possible 10 nouns with each ending; then they were asked to select any other two target endings and to identify 10 nouns with each ending and then to identify the gender of 60 nouns that had not appeared in any of the preceding texts. Exercises such as these were corrected as part of a whole-class activity.

In the next set of texts, which revolved around *L'Odyssée d'Alice Tremblay* and its less positive reviews, target nouns still appeared with their endings in bold but blank spaces preceded target nouns so that students had to add the correct determiners. An example follows:

Critique 14: Je m'attendais à beaucoup plus, j'ai été un peu déçu. _____
divertissement pour une petite soirée ennuyeuse, mais j'aurais préféré
_____ casino ou même _____ bingo! 6/10

[Critique 14: I was expecting a lot more, I was a bit disappointed. A distraction for a boring little evening, but I would have preferred the casino or even bingo! 6/10]

The exercises that followed these texts emphasized differences between feminine and masculine adjectives, drawing attention to the fact that many noun endings that can be distinguished as being either typically feminine or masculine are also found in adjectives.

Laboratory Feedback Treatment

The classroom instructional treatment was held constant for all participants insofar as they were all exposed to the same instructional materials taught by the same instructor. Different kinds of feedback treatments were then provided outside of the classroom in a laboratory setting in which individual students completed three different oral tasks, on two occasions, with one of three research assistants who were native or near-native speakers of French. The pretest scores from the computerized binary-choice test (see below) were immediately available, being computer generated and requiring no coding. They were therefore used to assign participants to either the recast group ($n = 14$) or the prompt group ($n = 11$), making sure that there were no significant differences between groups at the outset. The unequal distribution of participants in each group is the result of participant attrition that occurred prior to delayed posttesting, well after group assignment had taken place. The 25 participants each took part in two 30-min feedback sessions conducted 1 week apart, in Weeks 3 and 4, as outlined earlier, thus overlapping in Week 3 with the second (and final) week of classroom instruction. During the tasks, the researcher provided either prompts or recasts, depending on the feedback condition, following learner errors in grammatical gender.

Oral Tasks

There were two forms of each oral task (A and B) so that students alternated from Form A to Form B or from Form B to Form A throughout the study. Digital audio recordings were made of the interaction between the researcher and student participants so that all interactional data could be transcribed for verification and follow-up analyses.

Object-identification task. During the object-identification task, which was adapted from Harley (1998) and Lyster (2004), the researcher showed the student a set of two drawings of the same object. For example, as the researcher pointed to drawings of two planes (see Figure 1), she named these objects, *Ce sont deux rabots* “These are two planes”; she then covered one drawing, pointed to the other, and asked *Qu’est-ce que c’est?* “What is it?” as a means of creating an obligatory context for the singular gender-specific article in the right response, *C’est un rabot* “It’s a plane.” After doing an example with the student to ensure comprehension of the task, the researcher proceeded to present 38 such items.

Picture-description task. For the picture-description task, the researcher used a colorful 20 × 32-in. drawing of a street scene for Form A (see Appendix B), eliciting target words such as *pâtisserie, bicyclette, vélo, poubelle, échelle, cadeau, peinture, pharmacie, poussette*, and *magasin*, and a restaurant scene for Form B (see Appendix C), eliciting target nouns such as *assiette, farine, confiture, fraise, fromage, vin, fourchette, cuisine, baguette*, and *potage*. The student was asked to either tell a story about the scene or simply describe as much of the scene as possible. Each drawing included various vignettes

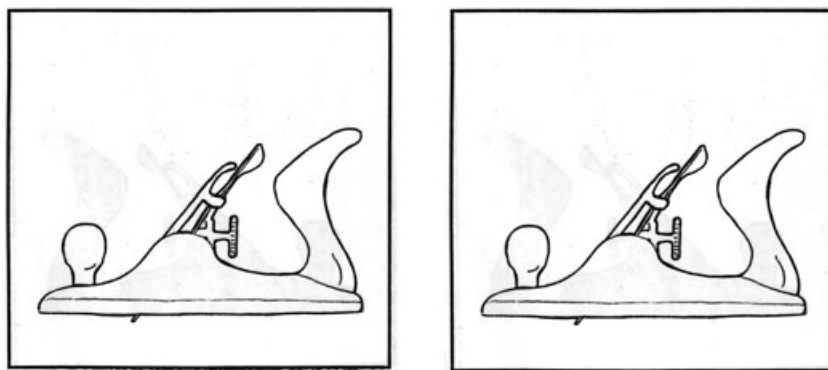


Figure 1 Sample item from object-identification task.

designed to pique the student's curiosity and to elicit a minimum of 25 target words.

Riddles. For the final interactional activity, 150 riddles were created—75 in Form A and 75 in Form B. For each feedback session, the researcher and the participant completed five sets of 15 riddles. For each set, the participant was given a card with 18 nouns void of any determiners, from which to select answers to the riddles read by the researcher and which needed to be preceded by the correct determiner (see the sample set in Appendix D). For example, the riddle, *Je sers à diviser le terrain de deux voisins* "I'm used to divide the land between two neighbors" elicited *une clôture* "a fence" as the correct response. The riddles were relatively simple in order to compensate for their high frequency, which was necessary to create multiple opportunities for learners to access nouns and determiners during online production and for the researcher to provide feedback.

Feedback Conditions

The two feedback types selected for comparison were recasts and prompts; the latter entailed clarification requests followed by a repetition of the error if necessary. Comparison of recasts with prompts operationalized in this way was thought to maintain consistency across feedback types with respect to implicitness, following Loewen and Nabei's (2007) classification of the six feedback types identified by Lyster and Ranta (1997) along a continuum of explicitness (see Figure 2). Recasts are situated at the implicit end of feedback types involving other-repair, whereas clarification requests and repetition of error are situated at the implicit end of feedback leading to self-repair. Because participants receiving prompts were provided feedback a second time if necessary, whereas participants receiving recasts received feedback only once, there was a slight imbalance across treatment groups, which was unavoidable in order

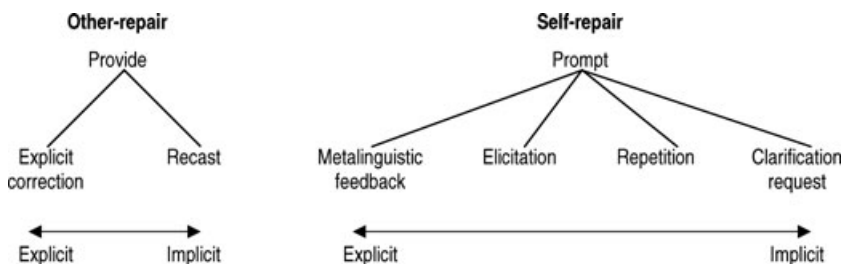


Figure 2 Options for corrective feedback (from Loewen & Nabei, 2007, p. 326).

to ensure a controlled comparison of feedback followed by self-repair with feedback followed by no repair. Our initial design even included recourse to a third and fourth prompt (elicitation and metalinguistic clue, respectively) but it was never necessary to do so and it was indeed rare to provide a second prompt (frequency of feedback in each group is reported in the Results section). Before the actual study, researchers participated in a 1-hr guided practice session with student volunteers, and they were given guidelines that provided scripted responses. The scripted responses illustrated below pertain to the exchanges involving riddles.

Recast condition. In the recast condition, researchers responded to errors in grammatical gender by using recasts that reformulated the participant's answer with the right gender marking. No metalinguistic information was provided, no intonational changes were made for emphasis, and no opportunity for the participant to repeat the recast was provided, as in the following example:

- Researcher: Je suis un vaisseau sanguin dans lequel circule le sang qui revient au cœur. [*I am a blood vessel through which blood circulates back to the heart.*]
Participant: Un veine. [*A vein-M.*]
Researcher: Une veine. Oui, on continue. [*A vein-F. Yes, let's continue.*]

Prompt condition. In the prompting condition, researchers were instructed to use prompts in a particular order, beginning with a clarification request and, if necessary, followed by repetition of the participant's error. Typically, only one prompt was used and, as in the following scripted example, this was a clarification request:

- Researcher: Je suis comme une ville mais plus petite et à la campagne. [*I am like a city but smaller and in the country.*]
Participant: Une village. [*A village-F.*]
Researcher: Pardon? [*Pardon me?*]
Participant: Un village. [*A village-M.*]
Researcher: Oui, on continue. [*Yes, let's continue.*]

If a second prompt was necessary, the researcher repeated the participant's (second) error with rising intonation, as in the following scripted example:

- Researcher: Je suis rempli d'encre et on se sert de moi pour écrire. [*I am full of ink and am used for writing.*]
Participant: Une stylo. [*A pen-F.*]

- Researcher: Pardon? [*Pardon me?*]
 Participant: Une stylo. [*A pen-F*]
 Researcher: Une stylo? [*A pen-F?*]
 Participant: Un stylo. [*A pen-M*.]
 Researcher: Oui, on continue. [*Yes, let's continue.*]

Both conditions. Scripted responses were created for both prompters and recasters to review beforehand to ensure that, in both conditions, researchers responded to participants in identical fashion in cases where (a) participants answered correctly with the right lexical item and the correct gender marking, (b) participants answered using the wrong lexical item, (c) participants did not use a determiner and (d) the researcher could not hear the gender marking. An example of each of these four cases follows.

(a) Participant uses article correctly:

- Researcher: J'ai un siège, un dossier et quatre pattes. Une personne peut s'asseoir sur moi. [*I have a seat, a back, and four legs. A person can sit on me.*]
 Participant: Une chaise. [*A chair.*]
 Researcher: C'est bien ça, on continue . . . [*That's good, let's continue . . .*]

(b) Participant chooses wrong lexical item:

- Researcher: Je suis un outil qui sert à percer des trous. [*I'm a tool used for making holes.*]
 Participant: Un marteau. [*A hammer.*]
 Researcher: Ce n'est pas le bon mot . . . [*That's not the right word . . .*]

(c) Participant does not use article:

- Researcher: Je suis ce qui est vrai. [*I am what is true.*]
 Participant: Vérité. [*Truth.*]
 Researcher: Il faut dire l'article . . . [*You have to say the article . . .*]

(d) Researcher cannot hear the article:

- Researcher: J'ai un oncle et une tante et ma sœur est leur nièce. Qui suis-je? [*I have an uncle and an aunt and my sister is their niece. Who am I?*]
 Participant: <?> neveu. [*<?> nephew.*]
 Researcher: Je n'ai pas entendu . . . [*I didn't hear . . .*]

Testing Instruments

The design of this experimental study entailed three testing sessions: pretests the week before the instructional treatment began, immediate posttests 4 weeks later (immediately following the final feedback session), and delayed posttests 3 weeks after the immediate posttests. Each testing session entailed two oral production measures and a computerized binary-choice test, the latter yielding both accuracy and reaction-time scores. Each test had two forms so that participants alternated forms from one session to the next.

Oral production measures. The oral production measures included both Form A and Form B of the object-identification task and picture-description task that were used during the treatment sessions to elicit target nouns and create opportunities to provide feedback. However, during the scheduled testing sessions, no feedback on accuracy was provided. Digital audio recordings were made of the oral interaction between the researcher and student participants during the testing sessions; they were transcribed and coded for accuracy by native or near-native speakers of French. The oral data were scored twice, in their entirety, to measure interrater reliability, which was 99% at each of the testing times in the case of the object-identification task and 98% in the case of the picture-description task.

Binary-choice and reaction-time measures. Using the C+ programming language, we designed a computerized binary-choice test that served as a measure of accuracy and included a reaction-time measure to assess fluency. Delivered through a Toshiba Protégé 3500 Tablet PC Laptop TOUCHSCREEN, this test was adapted from the pencil-and-paper version that was used in Lyster's (2004) classroom study with young learners, but without the drawings, which had served to contextualize discrete-point items and to reduce confusion for children unfamiliar with some of the lexical items. In the present study, the binary-choice test consisted of 80 lexical items that appeared on the computer screen one at a time (see Figure 3). The lexical items on Form A were randomized manually to determine their order of appearance; then both forms were made equivalent by using different lexical items but with the same distribution of target endings presented in the same order. For example, the first five items were *fourchette*, *différence*, *blouse*, *bougie*, and *train* on Form A and *trompette*, *naissance*, *berceuse*, *garantie*, and *bain* on Form B. Under each lexical item appeared the masculine and feminine forms of the indefinite article (*un*, *une*), each in a click-box so that participants could select one by simply touching the screen.

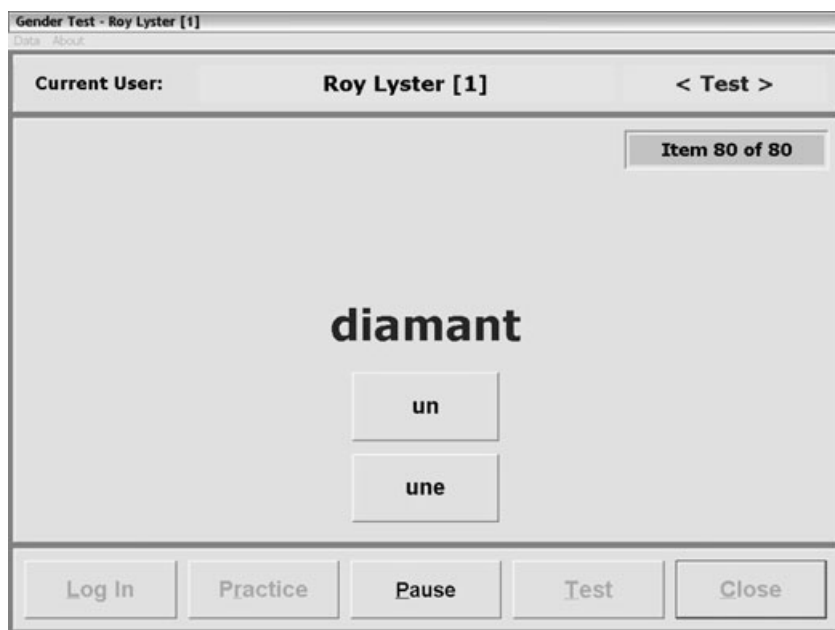


Figure 3 Computer screen sample.

Prior to the actual test, participants completed a short practice session, which they could complete as many times as needed, but none of the participants completed it more than twice. Participants were not told that they were being timed. During the actual test, a hidden timer counting in milliseconds ran from the moment the target item appeared on the screen until the participants made their gender choice by touching the screen. Then a new target item appeared and the timer restarted. For scoring purposes, the computer program generated a list with all the test items, participants' gender choice, and the amount of time that they had required to answer. Based on the number of correct answers, the program generated an accuracy score out of 80. Reaction-time scores were produced to convey the number of milliseconds that each participant required to provide all 80 answers. As is often the case in reaction-time measures, we also examined the amount of time learners spent generating only correct answers. However, these results were similar to those that obtained by adding up the time participants spent on generating all their answers; we therefore report the latter in the Results section.

Exit Questionnaire

At the end of the final testing session, participants were asked to complete a short six-item questionnaire that asked whether they had noticed receiving feedback and whether they had noticed receiving any specific types of feedback. They were also asked to comment on how easy or difficult the feedback was to notice, whether it was helpful, and whether they would have preferred not receiving any feedback. The questionnaire, which was not computerized, is reproduced in Appendix E.

Summary

A form-focused instructional unit was designed and implemented in two classrooms to draw the attention of L2 learners of French to target noun endings that predict grammatical gender. The classroom instruction was supplemented by a set of feedback treatments implemented in a laboratory setting in order to compare the effectiveness of prompts and recasts for improving learners' ability to accurately assign grammatical gender. The binary-choice and reaction-time measures were used solely for testing purposes, whereas the object-identification and picture-description tasks were employed without feedback for testing purposes but then again with feedback during the two feedback treatment sessions. The riddles were used only during the feedback treatments sessions, along with the object-identification and picture-description tasks in order to maximize opportunities for learners to access nouns and determiners repeatedly during online production and for the researcher to provide feedback. During online production, the participants in the present study were required to access rule-based knowledge of more than 20 high-frequency noun endings and the differential effects they have on the grammatical gender of hundreds of familiar and unfamiliar nouns.

Results

The audio recordings of the two feedback treatment sessions confirmed that the quantity of feedback was similar in both conditions, with 328 instances occurring in the prompt condition and 341 in the recast condition. As displayed in Table 3, participants in the prompt condition each received an average of 29.8 prompts ($SD = 12.4$) and participants in the recast condition each received an average of 24.4 recasts ($SD = 15.8$). The prompt group received slightly more feedback than participants in the recast group because, as mentioned earlier, a follow-up prompt in the form of a repetition was used in the event that the clarification request did not succeed in eliciting self-repair. As anticipated,

Table 3 Amount of feedback and immediate repair per group

Condition	Feedback			Repair		
	Total	Mean	<i>SD</i>	Total	Mean	<i>SD</i>
Prompts (<i>n</i> = 11)	328	29.8	12.4	323	29.4	12.5
Recasts (<i>n</i> = 14)	341	24.4	15.8	15	1.1	2.4

because participants were given the opportunity to modify their output following prompts but not recasts, those in the prompt group repaired 323 of their 328 ill-formed utterances (98%), whereas participants in the recast group modified only 15 of their 341 erroneous utterances (4%). That learner repair was not 100% in the prompt condition and not nil in the recast condition resulted from occasional but minimal use of the wrong feedback type in each condition.

Repeated-measures ANOVAs were used to assess differences in accuracy scores between the prompt and recast groups and any change over time (from pretest to immediate posttest to delayed posttest). Separate ANOVAs were used for the object-identification and picture-description tests, and the binary-choice test involved two separate ANOVAs: one for accuracy scores and one for reaction-time scores. The ANOVAs were conducted using the mixed-model procedure (PROC MIXED) of SAS, version 9.1 (SAS Institute, Inc., 2002–2003). The correlation of the repeated measures over time was explored by modeling three possible structures: compound symmetry (equal correlation), autoregressive (first-order correlation; i.e., the correlations decrease exponentially with the distance between the measurements), and unstructured correlation. The goodness-of-fit of the models with the different correlation structures was compared using two criteria: Akaike's Information Criterion and Schwartz's Bayesian Criterion (Littell, Milliken, Stroup, Wolfinger, & Schabenberge, 2006). In all ANOVAs, the unstructured correlation was chosen. Because the groups were unbalanced, a Kenward-Roger correction was applied to the degrees of freedom (Kowalchuk, Keselman, Algina, & Wolfinger, 2004). When a significant main or interaction effect was found, Bonferroni probability adjustments for multiple comparisons were used to identify significant mean differences. Effect sizes were calculated for all the mean contrasts within the significant main or interaction effect. Effect sizes were obtained by using Cohen's *d*, which was calculated as the difference between the two group means divided by the pooled within-group variance of the two groups. The data were checked to ensure they met the assumptions required for ANOVA. In the two cases in which the assumptions were not

met (picture-description scores and reaction-time scores), a natural logarithm transformation was performed on the scores to stabilize the variance.

At the time of pretesting *t*-Tests were conducted on the results obtained from Form A and Form B of each measure. No significant differences emerged between forms on any of the measures: object-identification test, $t = 1.04, p = .31$; picture-description test, $t = 1.21, p = .24$; computerized binary-choice test, $t = 0.58, p = .57$; reaction-time measures, $t = 0.64, p = .53$. This allowed for the conflation of both forms of each measure in the analyses. As reported below, no significant differences between the prompt and recast groups appeared on any pretest measures, indicating group equivalence at the onset of the study.

Object-Identification Test

The means and standard deviations resulting from the object-identification test for each group across three testing times appear in Table 4 and are displayed graphically in Figure 4. The ANOVA revealed a significant main effect for time, $F(2, 22) = 23.74, p < .001$. Mean contrasts revealed significant differences between pretest and immediate posttest ($p < .001, d = 1.32$) and between pretest and delayed posttest ($p < .001, d = 1.42$) but not between immediate and delayed posttests ($p = 1, d = 0.13$). There was neither a significant effect for feedback condition, $F(1, 23) = 0.12, p = .732$, nor a significant interaction between test time and feedback condition, $F(2, 22) = 0.48, p = .627$.

Picture-Description Test

For the picture-description test, Table 5 displays the means and standard deviations of the log-transformed values and Figure 5 displays graphically the log-transformed group means over time. The ANOVA revealed a significant main effect for time, $F(2, 22) = 21.68, p < .001$. Mean comparisons yielded significant differences between pretest and immediate posttest ($p < .001, d = 1.84$) and between pretest and delayed posttest ($p < .001, d = 1.27$) but not between immediate and delayed posttests ($p = .492, d = -0.33$). Again, no

Table 4 Object-identification test means over time by feedback condition (maximum score = 38)

		Pretest	Posttest 1	Posttest 2
Prompt group	Mean	30.36	35.27	36.09
	SD	6.02	2.53	2.21
Recast group	Mean	29.36	35.57	35.5
	SD	5.21	2.68	2.44

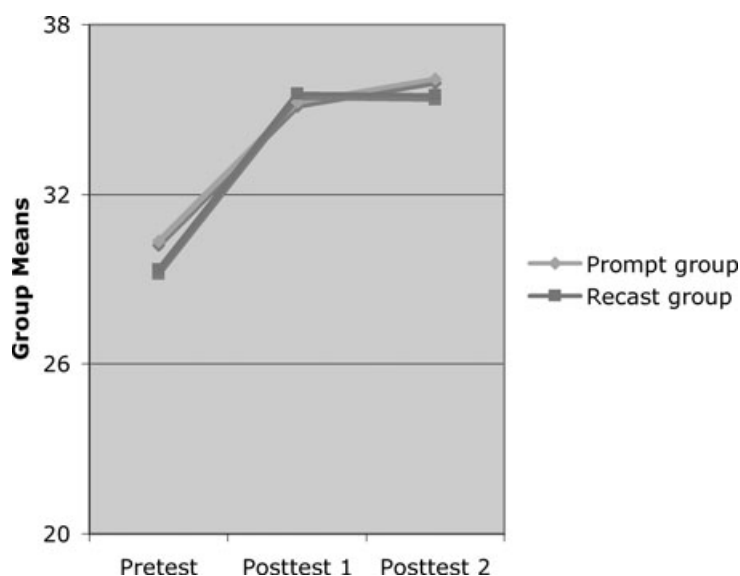


Figure 4 Object-identification test means.

significant differences were detected between groups, $F(1, 23) = 0.36$, $p = .552$, and no significant interaction effect was detected between test time and feedback condition, $F(2, 22) = 2.1$, $p = .146$.

Binary-Choice Test

The means and standard deviations for the accuracy scores obtained by both groups across time appear in Table 6; the means over time are displayed graphically in Figure 6. The ANOVA again revealed a significant main effect for time, $F(2, 22) = 64.89$, $p < .001$. The mean comparisons again revealed significant differences between pretest and immediate posttest ($p < .001$, $d = 2.90$) and between pretest and delayed posttest ($p < .001$, $d = 2.34$) but not between

Table 5 Log-transformed picture-description test means over time by feedback condition

		Pretest	Posttest 1	Posttest 2
Prompt group	Mean	2.98	3.18	3.11
	SD	0.13	0.05	0.15
Recast group	Mean	3.01	3.15	3.16
	SD	0.1	0.07	0.04

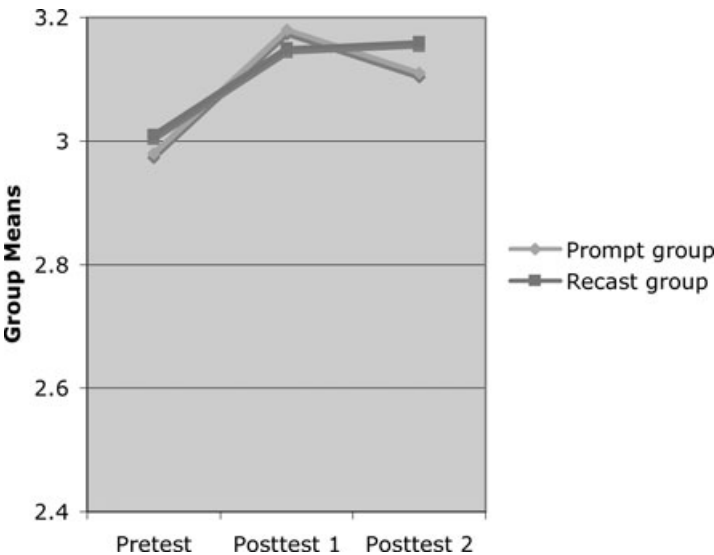


Figure 5 Log-transformed picture-description test means.

immediate and delayed posttests ($p = .258$, $d = -0.29$). No significant differences were revealed between feedback conditions, neither as a main effect, $F(1, 23) = 0.24$, $p = .626$, nor in their interaction with time, $F(2, 22) = 0.55$, $p = .584$.

Reaction-time scores on the binary-choice test were calculated by adding the reaction time in milliseconds for all 80 items. The means and standard deviations of the log-transformed values of reaction-time scores obtained by both groups across time are displayed in Table 7, and the log-transformed means over time are displayed graphically in Figure 7. Similar to all other measures, there was a significant effect for testing time, $F(2, 22) = 35.56$, $p < .001$, with significant mean contrasts between pretest and immediate posttest ($p = .005$, $d = 2.25$) and between pretest and delayed posttest ($p = .005$, $d = 2.85$) but

Table 6 Binary-choice test means over time by feedback condition (maximum score = 80)

		Pretest	Posttest 1	Posttest 2
Prompt group	Mean	56.27	74.36	71.45
	SD	7.54	4.23	7.67
Recast group	Mean	56.86	74.57	73.71
	SD	6.64	6.21	6.28

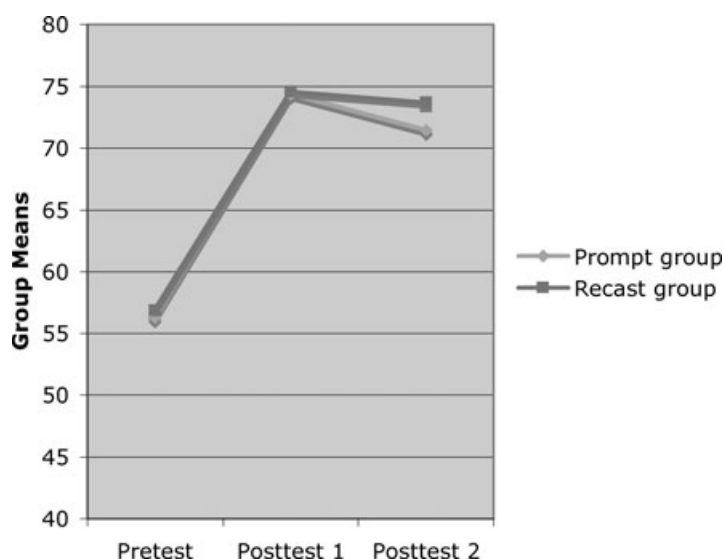


Figure 6 Binary-choice test means.

not between immediate and delayed posttests ($p = .060$, $d = -0.59$). Again similar to all other measures, the main effect for feedback condition was not significant, $F(1, 23) = 1.45$, $p = .241$, nor was the interaction effect between test time and feedback condition, $F(2, 22) = 0.63$, $p = .541$.

Questionnaire Results

The final questionnaire about feedback preferences was completed by 21 of the 25 participants. All but one of the respondents recalled receiving feedback; those who recalled receiving feedback all indicated that the feedback was easy to notice and helpful. Even though the questionnaire gave clear examples of a recast and a prompt, many thought they had received a mixture of both types,

Table 7 Log-transformed reaction-time means over time by feedback condition

		Pretest	Posttest 1	Posttest 2
Prompt group	Mean	5.41	5.11	4.92
	SD	0.27	0.16	0.2
Recast group	Mean	5.32	4.94	4.86
	SD	0.48	0.22	0.26

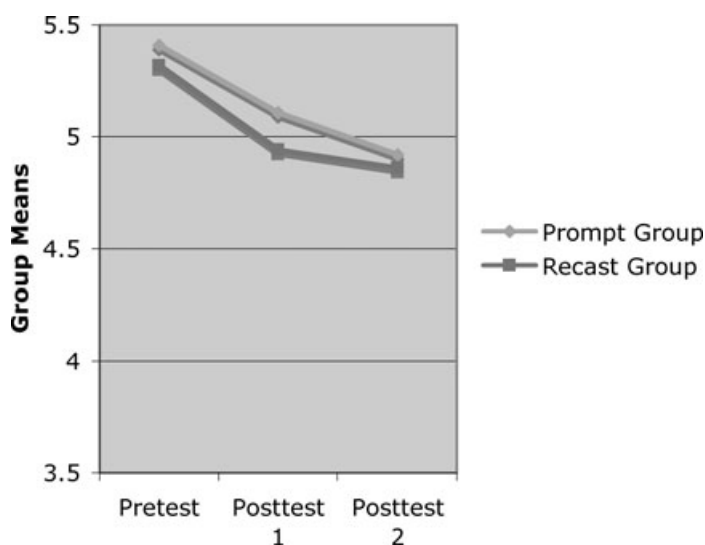


Figure 7 Log-transformed reaction-time means.

even though audio recordings of the feedback sessions confirmed that this was not the case. None indicated a preference for not receiving feedback. Most recalled sessions when they received no feedback (i.e., the testing sessions), and all but one indicated that they would have preferred to receive feedback during those sessions as well.

Discussion

The repeated-measures ANOVA yielded identical results on all four measures; that is, learners in both groups, regardless of feedback condition, made significant progress over time, improving from pretest to immediate posttest and then maintaining their improvement at the time of delayed posttesting. Their significant progress over time yielded large effect sizes, with d values ranging from 1.27 for the picture-description pretest-to-delayed-posttest contrast to 2.90 for the binary-choice pretest-to-immediate-posttest contrast.

Of interest, of course, is whether a larger number of participants would have had an impact not only on the main effect of testing time, as in the present study, but also on the interaction between feedback condition and testing time. Given the constraints we faced in accessing a larger number of participants, the sample in the present study was small. Moreover, the sample did not include a

comparison group that received instruction without any feedback treatment. It may have been the case, therefore, that the form-focused instruction provided over 3 hr during a 2-week period outweighed any potential impact that the two half-hour sessions of feedback might have had and was even sufficient on its own to cause the gains made by all participants. However, previous studies by Ammar and Spada (2006) and Lyster (2004) had already demonstrated that learners benefit from a combination of instruction and feedback more than from instruction only, with variable effects for learners receiving the same instruction but different types of feedback. In the following discussion, therefore, we assume the likely probability that the effects of the form-focused instruction provided in class were enhanced by the feedback sessions that took place out of class. However, further study is warranted that would compare the effects of instruction with and without feedback provided during dyadic interaction.

An unexpected finding in this study was that both groups made similar progress over time, regardless of the type of feedback they received. It had been predicted that students receiving prompts would outperform those receiving recasts, given the opportunities that prompts provide for a deeper level of processing that entails both retrieval and production mechanisms. We begin the discussion by asking why no significant differences emerged between feedback conditions in the present study, as had been the case in recent classroom studies by Ammar and Spada (2006), Ellis (2007), Ellis et al. (2006), and Lyster (2004).

A primary determinant in recast effectiveness has proven to be the extent to which a classroom's overall communicative orientation is more form-focused or more meaning-focused, with recasts being more easily noticed in form-focused classrooms than in meaning-focused classrooms (Ellis & Sheen, 2006; Lyster & Mori, 2006; Nicholas et al., 2001; Sheen, 2004). Not only does the overall communicative orientation of a classroom play a role in recast effectiveness, the orientation of individual learners does as well, as explained by Ellis and Sheen:

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. (pp. 596–597)

The intermediate learners of French in the present study can generally be considered as form-oriented, given their expressed interest in unraveling what appeared to them as the mysteries of grammatical gender, which required five separate visits to our research laboratory outside of regular class time.

In addition, because participants in the present study had just received form-focused instruction in class, they had been primed to notice the feedback and make choices between masculine and feminine gender attribution. As Nicholas et al. (2001) suggested, recasts may benefit language development when “the learner has already begun to use a particular linguistic feature and is in a position to choose between linguistic alternatives” (p. 752; see also Ishida, 2004).

Moreover, given the consistency with which the recasts were provided after errors in grammatical gender, they arguably fell short of being as implicit as we had initially intended. Based on their review of recast studies, Ellis and Sheen (2006, p. 585) concluded that recasts actually range along a continuum from implicit to explicit: “recasts cannot be viewed as a purely implicit form of negative feedback. In many cases, their illocutionary force as corrections is quite transparent and, therefore, they should be seen as a relatively explicit form of negative feedback.” The recasts used in the present study can be considered at the explicit end of the continuum, because they were always short, involving reformulation of no more than one noun phrase (see Egi, 2007; Loewen & Philp, 2006; Philp, 2003; Sheen, 2006). In addition, as Han (2002) concluded, recasts are effective in laboratory settings because they can be provided intensively and with consistency to learners receiving individualized attention (see also Lyster, 1998; Nicholas et al., 2001; Spada, 1997).

A laboratory setting was included in the present study in order to ensure intensive individualized feedback for all participants and to regulate their opportunities for modified output according to treatment condition. The impact of controlling variables in this way is that the laboratory component in the present study created conditions in which all feedback, whether prompts or recasts, proved equally effective. The effectiveness of recasts and prompts alike can be attributed to the form-focused orientation of the learners and to the laboratory setting itself, which afforded learners not only intensive feedback on one form but also individualized attention. In addition, prompts and recasts may have both proven effective as a result of the explicitness that superseded the implicitness that was initially intended for both feedback types. Not only did the recast condition entail short recasts with a single change; any ambiguity otherwise associated with either recasts or prompts was attenuated by the frequency of both feedback types and by the participants’ awareness of the linguistic targets—that is, the exit questionnaires revealed that students were well aware of the feedback they were given and its focus on grammatical gender. Most stated that they appreciated the feedback sessions and disliked the absence of feedback during the testing sessions.

Notwithstanding good reasons for both prompts and recasts to prove equally effective in a laboratory setting, this finding begs the question as to why participants with consistent opportunities for self-repair in the prompt group did not outperform the recast group, who had no such opportunities. Our results lend support to the observation that the effectiveness of recasts is not contingent upon learner repetition, whereas the effectiveness of prompts is associated with their propensity for eliciting modified output (Loewen & Philp, 2006). Loewen and Philp found that the effectiveness of recasts lies not in the opportunities they provide for learners to repair their nontarget output (see also Mackey & Philp, 1998; McDonough, 2007; McDonough & Mackey, 2006); it lies instead in specific input features related to length, number of changes, and intonation (see also Egi, 2007; Ellis & Sheen, 2006; Nicholas et al., 2001; Sheen, 2006).

In the present study, learners in the recast group benefited from repeated exposure to positive exemplars in the form of recasts that were made salient by the discourse context of interaction between a learner and a researcher (see Leeman, 2003). It is also likely the case that the salience of the recasts enabled learners to infer negative evidence by engaging in a cognitive comparison of the recast and their nontarget utterance. Learners in the prompt group had no exposure to positive exemplars of the linguistic targets and instead benefited from repeated exposure to negative evidence and opportunities to produce modified output. Therefore, although both types of feedback proved equally effective, the types of processing that led to their respective effectiveness are arguably very different, but the measurements used in the present study were not designed to capture those differences. Differences might have emerged if we had measured proficiency in terms of fluency, not only accuracy (Leila Ranta, personal communication, June 7, 2005); that is, fluency measures might reveal that prompts leading to self-repair are more effective than recasts without repair, because the benefits of prompts are more likely to appear in learners' control over gender attribution than in their levels of analysis (see Bialystok, 1994). We did measure reaction times, which decreased significantly over time for both groups, but we did so only for the computerized binary-choice test, not for oral production. Of further interest would be a study of even longer duration than ours to assess whether the deeper level of processing that we predicted to result from prompts leads to more robust change over a longer period of time than do recasts. If not, then the type of cognitive comparison initiated by focused recasts, as described by Loewen and Philp (2006; see also Doughty, 2001) and used in this study, must entail a similarly deep level of processing, at least within the context of dyadic interaction. A longer study, however, would make it more difficult to isolate the effects of feedback on development, because,

over time and outside of the experimental setting, learners would be exposed to opportunities for practice, which, arguably, would have a greater impact on development than feedback alone.

One final question left without a conclusive answer is the extent to which the learning that occurred in the present study consisted of rule-based or exemplar-based learning. This is because the testing instruments used in the present study included many familiar high-frequency nouns rather than only low-frequency items that were unknown to intermediate adult learners, and many of these familiar nouns occurred in the treatment sessions as well as the testing sessions. For example, about half of the items appearing on the computerized binary-choice reaction-time test were relatively familiar nouns such as *drapeau* “flag” that also occurred during the treatment tasks, whereas the other half were lower frequency nouns such as *narine* “nostril” that occurred only in the computerized test. With immediate posttest scores of both groups on the computerized test at a high average of 74 (out of 80), it may be the case that a combination of exemplar-based learning and rule-based learning was involved. As Skehan (1998) argued, rule-based and exemplar-based representational systems are not entirely separate and are both considered useful for language processing. He suggested that both systems are “in constant dialectic” (p. 92) in a way that enables learners to engage in complementary processes of analysis and synthesis.

Conclusion

Laboratory studies have generally yielded positive results for recasts, but not necessarily in comparison with other clearly defined types of feedback (e.g., Han, 2002; Ishida, 2004; Iwashita, 2003; Long et al., 1998; Mackey & Philp, 1998; McDonough & Mackey, 2006). Leeman’s (2003) study, however, which included a comparison of recasts and repetition of error without any opportunities for immediate repair, yielded superior results for recasts, whereas McDonough’s (2007) study, which compared recasts and clarification requests and included opportunities for immediate repair, showed both feedback types to be equally effective. The present study compared recasts and clarification requests (followed by a repetition if necessary) and included opportunities for immediate repair after the prompts but not after recasts, and it found both feedback types to be equally effective. Together, these results suggest that, in dyadic interaction with a researcher, learners are likely to benefit similarly from both feedback types—whether they be prompts or recasts—provided intensively and repeatedly on the same target feature, on the condition that learners be

given opportunities for immediate repair after prompts, but not necessarily after recasts. The finding that accuracy as well as reaction-time scores improved significantly over time in the present study, whether learners received recasts or prompts, leads to the conclusion that learners receiving recasts benefited from repeated exposure to positive exemplars as well as from opportunities to infer negative evidence, whereas learners receiving prompts benefited from repeated exposure to negative evidence as well as from opportunities to produce modified output.

In contrast to these laboratory studies, classroom studies conducted in a range of instructional settings have demonstrated that prompts lead to greater gains in accuracy than do recasts (Ammar & Spada, 2006; Ellis, 2007; Ellis et al., 2006; Havranek & Cesnik, 2001; Loewen & Philp, 2006; Lyster, 2004). These discrepant findings put into question the extent to which laboratory studies are apt to address pedagogically driven questions about the differential effectiveness of various feedback types provided in classrooms. As our study showed, there are many factors associated with intensive one-on-one interaction that contribute to the effectiveness of recasts; however, there are many reasons to not expect these contributing factors to carry over to the hurly-burly of communicative classrooms (e.g., Ellis & Sheen, 2006; Lyster, 1998; Spada, 2005). As Musumeci (1996) argued, "what learners will do in a small-group or one-to-one exchange with native speakers in the experimental setting may not generalize at all to the whole-class multiple-learners-one-teacher situation of the classroom" (p. 318; see also Foster, 1998; Lyster, 2002, 2007).

Ammar and Spada (2006) found that sixth-grade classroom learners with initially high accuracy scores in their use of the target forms benefited equally from recasts and prompts (much like the intermediate-level university students in the present study; see also Carroll & Swain, 1993; McDonough, 2007), whereas learners with lower accuracy scores benefited from prompts but not from recasts. Teachers, therefore, are not in a position to use only one type of feedback over another and instead need "to orchestrate, in accordance with their students' language abilities and content familiarity, a wide range of feedback types befitting of the instructional context" (Lyster, 2007, p. 124). The observation that some but not all students are able to benefit from recasts pushes us away from a simple binary comparison of prompts versus recasts and toward more fine-grained analyses of the discourse contexts, linguistic targets, and learner characteristics that are more amenable to one type of feedback than to another. It would be timely in future research to match the increasingly detailed information we now have about the characteristics and variable effectiveness of recasts with similar information about prompts. Of value, for example,

would be a research agenda designed to explore the long-term effectiveness of the various feedback moves comprising prompts relative to each other and to different types of recasts, comparing their differential effects on different linguistic features and on classroom learners distinguished by a range of individual differences. Many of the contributions in Mackey's (2007) recent edited volume, *Conversational Interaction and Second Language Acquisition*, have begun moving that agenda forward.

Revised version accepted 15 July 2008

References

- Ammar, A., & Spada, N. (2006). One size fits all? Recasts, prompts and L2 learning. *Studies in Second Language Acquisition*, 28, 543–574.
- Ayoun, D. (2007). The second language acquisition of grammatical gender and agreement. In D. Ayoun (Ed.), *French applied linguistics* (pp. 130–170). Amsterdam: Benjamins.
- Bialystok, E. (1994). Analysis and control in the development of second language proficiency. *Studies in Second Language Acquisition*, 16, 157–168.
- Bosquart, M. (1998). *Nouvelle grammaire*. Montreal: Guérin.
- Braidi, S. (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.
- Carroll, S. (1989). Second-language acquisition and the computational paradigm. *Language Learning*, 39, 535–594.
- Carroll, S., & Swain, M. (1993). Explicit and implicit negative feedback: An empirical study of the learning of linguistic generalizations. *Studies in Second Language Acquisition*, 15, 357–386.
- Clark, S. (1995). The generation effect and the modeling of associations in memory. *Memory & Cognition*, 23, 442–455.
- de Bot, K. (1996). The psycholinguistics of the output hypothesis. *Language Learning*, 46, 529–555.
- DeKeyser, R. (1998). Beyond focus on form: Cognitive perspectives on learning and practicing second language grammar. In C. Doughty & J. Williams (Eds.), *Focus on form in classroom second language acquisition* (pp. 42–63). Cambridge, UK: Cambridge University Press.
- DeKeyser, R. (2001). Automaticity and automatization. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 125–51). Cambridge, UK: Cambridge University Press.
- DeKeyser, R. (Ed.). (2007). *Practice in a second language: Perspectives from applied linguistics and cognitive psychology*. Cambridge, UK: Cambridge University Press.

- deWinstanley, P. A., & Bjork, E.L. (2004). Processing strategies and the generation effect: Implications for making a better reader. *Memory & Cognition*, 32, 945–955.
- Doughty, C. (1994). Fine-tuning of feedback by competent speakers to language learners. In J. Alatis (Ed.), *Georgetown University round table 1993: Strategic interaction and language acquisition* (pp. 96–108). Washington, DC: Georgetown University Press.
- Doughty, C. (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., & Varela, E. (1998). Communicative focus on form. In C. Doughty & J. Williams (Eds.), *Focus on form in classroom second language acquisition* (pp. 114–138). Cambridge, UK: Cambridge University Press.
- Egi, T. (2007). Interpreting recasts as linguistic evidence: The roles of linguistic target, length, and degree of change. *Studies in Second Language Acquisition*, 29, 511–538.
- Ellis, R. (2007). The differential effects of corrective feedback on two grammatical structures. In A. Mackey (Ed.), *Conversational interaction in second language acquisition: A series 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). Re-examining the role of recasts in L2 acquisition. *Studies in Second Language Acquisition*, 28, 575–600.
- Filiatrault, D. (Director). (2002). *L'Odysée d'Alice Tremblay* [Motion picture]. Canada: Cinéma Maginaire Inc.
- Foster, P. (1998). A classroom perspective on the negotiation of meaning. *Applied Linguistics*, 19, 1–23.
- Gass, S. (2003). Input and interaction. In C. Doughty & M. Long (Eds.), *Handbook of second language acquisition* (pp. 224–255). Oxford: Blackwell.
- Han, Z. (2002). A study of the impact of recasts on tense consistency in L2 output. *TESOL Quarterly*, 36, 542–572.
- Han, Z. (2008). On the role of meaning in focus on form. In Z. Han (Ed.), *Understanding second language process* (pp. 45–79). Clevedon, UK: Multilingual Matters.
- Harley, B. (1998) The role of form-focused tasks in promoting child L2 acquisition. In C. Doughty & J. Williams (Eds.), *Focus on form in classroom second language acquisition* (p. 156–174). Cambridge, UK: Cambridge University Press.
- Havranek, G. (2002). When is corrective feedback most likely to succeed? *International Journal of Educational Research*, 37, 255–270.

- Havranek, G., & Cesnik, H. (2001). Factors affecting the success of corrective feedback. *EUROSLA Yearbook*, 1, 99–122.
- 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. (2001). The effect of learner proficiency on interactional moves and modified output in nonnative-nonnative interaction in Japanese as a foreign language. *System*, 29, 267–287.
- Jacob, R., & Laurin, J. (1994). *Ma grammaire*. Montreal: Les Éditions françaises.
- Kowalchuk, R., Keselman, H., Algina, J., & Wolfinger, R. (2004). The analysis of repeated measurements with mixed model adjusted *F* tests. *Educational and Psychological Measurement*, 64, 224–242.
- Lee, J. (2006). *Corrective feedback and learner uptake in English immersion classrooms in Korea*. Unpublished master's thesis, International Graduate School of English, Seoul, Korea.
- Leeman, J. (2003). Recasts and second language development: Beyond negative evidence. *Studies in Second Language Acquisition*, 25, 37–63.
- Lightbown, P. M. (2000). Classroom SLA research and second language teaching. *Applied Linguistics*, 21, 431–462.
- Littell, R., Milliken, G., Stroup, W., Wolfinger, R., & Schabenberge, O. (2006). *SAS[®] for mixed models* (2nd ed.). Cary, NC: SAS Institute.
- Loewen, S. (2005). Incidental focus on form and second language learning. *Studies in Second Language Acquisition*, 27, 361–386.
- 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 series of empirical studies* (pp. 361–378). Oxford: Oxford University Press.
- Loewen, S., & Philp, J. (2006). Recasts in the adult English L2 classroom: Characteristics, explicitness, and effectiveness. *Modern Language Journal*, 90, 536–555.
- Long, M. (1996). The role of the linguistic environment in second language acquisition. In W. C. Ritchie & T. K. Bhatia (Eds.), *Handbook of second language acquisition* (pp. 413–468). San Diego, CA: Academic Press.
- Long, M. (2007). *Problems in SLA*. Mahwah, NJ: Lawrence Erlbaum.
- Long, M., Inagaki, S., & Ortega, L. (1998). The role of implicit negative evidence in SLA: Models and recasts in Japanese and Spanish. *Modern Language Journal*, 82, 357–371.
- Lyster, R. (1998). Recasts, repetition, and ambiguity in L2 classroom discourse. *Studies in Second Language Acquisition*, 20, 51–81.
- Lyster, R. (2001). Negotiation of form, recasts, and explicit correction in relation to error types and learner repair in immersion classrooms. *Language Learning*, 51(Suppl. 1), 265–301.

- Lyster, R. (2002). Negotiation in immersion teacher-student interaction. *International Journal of Educational Research*, 37, 237–253.
- Lyster, R. (2004). Differential effects of prompts and recasts in form-focused instruction. *Studies in Second Language Acquisition*, 26, 399–432.
- Lyster, R. (2006). Predictability in French gender attribution: A corpus analysis. *Journal of French Language Studies*, 16, 69–92.
- Lyster, R. (2007). *Learning and teaching languages through content: A counterbalanced approach*. Amsterdam: Benjamins.
- Lyster, R., & Mori, H. (2006). Interactional feedback and instructional counterbalance. *Studies in Second Language Acquisition*, 28, 269–300.
- Lyster, R., & Mori, H. (2008). Instructional counterbalance in immersion pedagogy. In T. Fortune & D. Tedick (Eds.), *Pathways to bilingualism and multilingualism: Evolving perspectives on immersion education* (pp. 131–151). Clevedon, UK: Multilingual Matters.
- Lyster, R., & Ranta, L. (1997). Corrective feedback and learner uptake: Negotiation of form in communicative classrooms. *Studies in Second Language Acquisition*, 19, 37–66.
- Mackey, A. (Ed.). (2007). *Conversational interaction in second language acquisition: A series of empirical studies*. Oxford: Oxford University Press.
- Mackey, A., Al-Khalil, M., Atanassova, G., Ham, M., Logan-Terry, A., & Nakatsukasa, K. (2007). Teachers' intentions and learners' perceptions about corrective feedback in the L2 classroom. *Innovation in Language Learning and Teaching*, 1, 129–152.
- Mackey A., Gass S., & McDonough, K. (2000). How do learners perceive interactional feedback? *Studies in Second Language Acquisition*, 22, 471–497.
- Mackey, A., & Philp, J. (1998). Conversational interaction and second language development: Recasts, responses, and red herrings? *Modern Language Journal*, 82, 338–356.
- 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 series 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.
- McNulty, A. (2007). *A study of corrective feedback and uptake patterns as observed in four foreign language classrooms at selected mid-western public schools*. Unpublished doctoral dissertation, University of Nebraska, Lincoln.

- Mori, H. (2002). *Error treatment sequences in Japanese immersion classroom interactions at different grade levels*. Unpublished doctoral dissertation, University of California, Los Angeles.
- Musumeci, D. (1996). Teacher-learner negotiation in content-based instruction: Communication at cross-purposes? *Applied Linguistics*, 17, 286–325.
- Nicholas, H., Lightbown, P., & 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. (2000). Rethinking recasts: A learner-centered examination of corrective feedback in the Japanese classroom. In J. K. Hall & L. Verplaeste (Eds.), *The construction of second and foreign language learning through classroom interaction* (pp. 47–71). Mahwah, NJ: Lawrence Erlbaum.
- Oliver, R., & Mackey, A. (2003). Interactional context and feedback in child ESL classrooms. *Modern Language Journal*, 87, 519–533.
- Panova, I., & Lyster, R. (2002). Patterns of feedback and uptake in an adult ESL classroom. *TESOL Quarterly*, 36, 573–595.
- 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.
- Pouliot, J.-F. (Director). (2003). *La grande séduction* [Motion picture]. Canada: Canadian Film.
- Ranta, L., & Lyster, R. (2007). A cognitive approach to improving immersion students’ oral language abilities: The Awareness-Practice-Feedback sequence. In R. DeKeyser (Ed.), *Practice in a second language: Perspectives from applied linguistics and cognitive psychology* (pp. 141–160). Cambridge, UK: Cambridge University Press.
- Roberts, M. (1995). Awareness and the efficacy of error correction. In R. Schmidt (Ed.), *Attention and awareness in foreign language learning* (Tech. Rep. No. 9, pp. 162–182). Honolulu: University of Hawai’i, Second Language Teaching and Curriculum Center.
- 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. (2007). The effects of corrective feedback, language aptitude, and learner attitudes on the acquisition of English articles. In A. Mackey (Ed.), *Conversational interaction in second language acquisition: A series of empirical studies* (pp. 301–322). Oxford: Oxford University Press.
- Shiffrin, R.M., & Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending, and a general theory. *Psychological Review*, 84, 127–190.

- Skehan, P. (1998). *A cognitive approach to language learning*. Oxford: Oxford University Press.
- Spada, N. (1997). Form-focussed instruction and second language acquisition: A review of classroom and laboratory research. *Language Teaching*, 29, 73–87.
- Spada, N. (2005). Conditions and challenges in developing school-based SLA research programs. *Modern Language Journal*, 89, 328–338.
- Swain, M. (1985). Communicative competence: Some roles of comprehensible input and comprehensible output in its development. In S. Gass & C. Madden (Eds.), *Input in second language acquisition* (pp. 235–253). Rowley, MA: Newbury House.
- Swain, M. (1988). Manipulating and complementing content teaching to maximize second language learning. *TESL Canada Journal*, 6, 68–83.
- 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 series of empirical studies* (pp. 171–196). Oxford: Oxford University Press.
- Tsang, W. (2004). Feedback and uptake in teacher-student interaction: An analysis of 18 English lessons in Hong Kong secondary classrooms. *Regional Language Centre Journal*, 35, 187–209.
- Tucker, R., Lambert, W. E., & Rigault, A. (1977). *The French speaker's skill with grammatical gender: An example of rule-governed behaviour*. Paris: Mouton.
- White, J. (1998). Getting the learners' attention: A typographical input enhancement study. In C. Doughty & J. Williams (Eds.), *Focus on form in classroom second language acquisition* (pp. 85–113). Cambridge, UK: Cambridge University Press.
- Zyzik, E., & Polio, C. (2008). Incidental focus on form in university Spanish literature courses. *Modern Language Journal*, 92, 53–70.

Appendix A

Examples of Prompts

Prompts are defined below, along with examples from Grade 4 French immersion classes, all of which address grammatical gender to better illustrate differences across feedback types while maintaining consistency in error type (M = masculine; F = feminine).

- Clarification request: The teacher indicates to the student, by using phrases such as “*Pardon me*” and “*I don’t understand*,” that the message has not been understood or that the utterance is ill-formed in some way, and that a repetition or a reformulation is required.

- S: *La marmotte, c'est pas celui en haut?*
 [The groundhog-F, isn't it the one-M at the top?]
 T: *Pardon?* [Excuse me?]
 S: *La marmotte, c'est pas celle en haut?*
 [The groundhog-F, isn't it the one-F at the top?]

- Repetition: The teacher repeats the student's erroneous utterance, adjusting the intonation to highlight the error.

- S: *Puis ma grand-mère a acheté du laine pour faire euh . . . tu sais . . .*
 [And my grandmother bought some wool-M to make um . . . you know . . .]
 T: *Du laine?* [Wool-M?]
 S: *De la laine.* [Wool-F.]

- Metalinguistic clues: The teacher provides comments, information, or questions related to the well-formedness of the student's utterance, without explicitly providing the correct form (e.g., "*Do we say 'goed' in English?*" "*We don't say that in French,*" "*No,*" "*Is it masculine?*").

- S: *Parce qu'elle cherche, euh, son, son carte.*
 [Because she's looking for, um, her, her card-M.]
 T: *Pas son carte.* [Not her card-M.]
 S: *Euh, sa carte?* [Um, her card-F.]

- Elicitation: The teacher directly elicits correct forms from students by asking questions such as "*How do we say that in French?*" or by pausing to allow students to complete the teacher's utterance (e.g., "*C'est un . . . ?*"), or by asking students to reformulate their utterance (e.g., "*Try again*").

- T: *Il vit où un animal domestique? Où est-ce que ça vit?*
 [Where does a pet live? Where does it live?]
 S: *Dans un maison.* [In a house-M.]
 T: *Dans . . . ? Attention.* [In . . . ? Careful.]
 S: *Dans une maison.* [In a house-F.]

Appendix B

Picture Description Task: Form A



Appendix C

Picture Description Task: Form B



Appendix D

Sample Set of Riddles And Answer Card

1^e JEU

1.	Je suis un synonyme du mot «auto».	voiture
2.	Je suis la qualité d'une personne qui reste calme, attend sans s'énervier et va jusqu'au bout de ce qu'elle a entrepris.	patience
3.	Je suis un aliment fabriqué avec du lait; je m'appelle souvent camembert ou brie.	fromage
4.	Je suis un concours de circonstances qui fait que deux événements se produisent en même temps.	coïncidence
5.	Je suis la partie vitrée d'un magasin où l'on expose les objets à vendre.	vitrine
6.	J'ai un oncle et une tante et ma sœur est leur nièce. Qui suis-je?	neveu
7.	Je suis le pays habité par le plus grand nombre de francophones au monde.	France
8.	Je sers à diviser le terrain de deux voisins.	clôture
9.	J'ai deux yeux, un nez, une bouche, deux joues et un menton.	visage
10.	Je suis un produit dangereux qui se trouve dans le tabac.	nicotine
11.	Je suis un petit fruit rouge qui pousse sur le fraisier.	fraise
12.	Je suis un objet creux en métal qui sert à sonner.	cloche
13.	Je suis comme une ville mais plus petite et à la campagne.	village
14.	Je suis l'événement qui se produit lorsque quelqu'un naît.	naissance
15.	Je suis le poil épais qui couvre certains mammifères et qui sert à fabriquer des vêtements très chauds.	fourrure

1^e JEU

<i>bouchon</i>	<i>France</i>	<i>patience</i>
<i>cloche</i>	<i>fromage</i>	<i>salopette</i>
<i>clôture</i>	<i>métier</i>	<i>village</i>
<i>coïncidence</i>	<i>naissance</i>	<i>visage</i>
<i>fourrure</i>	<i>neveu</i>	<i>vitrine</i>
<i>fraise</i>	<i>nicotine</i>	<i>voiture</i>

Appendix E

Exit Questionnaire

Thank you for participating in this study! We have a few more questions to ask you before your participation comes to a close. But first, read the following explanation of what ‘feedback’ is.

Feedback is a type of error correction that is used (usually by a teacher) after a language learner makes a mistake. The following examples are two different types of feedback:

TYPE 1		TYPE 2	
Tester:	<i>Je suis ce qui est vrai.</i>	Tester:	<i>Je suis ce qui est vrai.</i>
Student:	<i>Le vérité.</i>	Student:	<i>Le vérité.</i>
Tester:	<i>La vérité. ← FEEDBACK</i>	Tester:	<i>Pardon? ← FEEDBACK</i>
	<i>Oui, on continue . . .</i>	Student:	<i>La vérité.</i>
		Tester:	<i>Oui, on continue . . .</i>

- Do you recall receiving any feedback on your errors during some of your sessions?
 - ☐ *Yes*
 - ☐ *No*
- If yes, do you recall what type of feedback?
 - ☐ *Type 1*
 - ☐ *Type 2*
 - ☐ *Both types*
 - ☐ *Can't remember*
- If you remember receiving a certain type of feedback, check as many descriptors below as you like to describe how you felt about the feedback.
 - ☐ *It was hard to notice.*
 - ☐ *It was easy to notice.*
 - ☐ *It was helpful.*
 - ☐ *It wasn't very helpful.*
 - ☐ *I would have preferred receiving another kind of feedback.*
 - ☐ *I would have preferred receiving no feedback.*

4. Do you recall some sessions when you received no feedback after your errors?
☐ *Yes*
☐ *No*
5. If yes, comment on how you felt about receiving no feedback.
☐ *I liked receiving no feedback.*
☐ *I would have preferred to receive feedback.*
6. Feel free to add any other comments on the back of this sheet about your participation in this study.