SSLA, **26**, 399–432. Printed in the United States of America.
DOI: 10.1017/S0272263104043025

DIFFERENTIAL EFFECTS OF PROMPTS AND RECASTS IN FORM-FOCUSED INSTRUCTION

Roy Lyster
McGill University

Four teachers and their eight classes of 179 fifth-grade (10–11-year-old) students participated in this quasi-experimental classroom study, which investigated the effects of form-focused instruction (FFI) and corrective feedback on immersion students' ability to accurately assign grammatical gender in French. The FFI treatment, designed to draw attention to selected noun endings that reliably predict grammatical gender and to provide opportunities for practice in associating these endings with gender attribution, was implemented in the context of regular subject-matter instruction by three of the four teachers, each with two classes, for approximately 9 hours during a 5-week period, while the fourth teacher taught the same subject matter without FFI to two comparison classes. Additionally, each of the three FFI teachers implemented a different feedback treatment: recasts, prompts, or no feedback. Analyses of pretest, immediate-posttest, and delayed-posttest results showed a significant increase in the ability of students exposed to FFI to correctly assign grammatical gender. Results of the written tasks in particular, and to a lesser degree the oral tasks, revealed that FFI is more effective when combined with prompts than with recasts or no feedback, as a means of enabling L2 learners to acquire rule-based representations of grammatical gender and to proceduralize their knowledge of these emerging forms.

This study was funded by the Social Sciences and Humanities Research Council of Canada (nos. 410-98-0175 and 410-2002-0988). Parts of this study were presented at the annual meeting of the American Association for Applied Linguistics in Salt Lake City on April 7, 2002; the Congress of the Social Sciences and Humanities in Toronto on May 26, 2002; and at the Congress of the International Association for Applied Linguistics in Singapore on December 12, 2002. I am grateful to the participating teachers and their students, to Lucy Fazio for her role as research associate in the data collection, 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, Kristina Eisenhower, Andréanne Gagné, Sophie Beaudoin, Laura-Annie Bouffard, France Bourassa, Sophie Bourgeois, Elisa David, Mélanie Mathieu, Sophie Prince, Andrea Sterzuk, and David Syncox. I gratefully acknowledge Leila Ranta, Iliana Panova, and four anonymous *SSLA* reviewers for their helpful comments on an earlier version of this paper.

Address correspondence to: Roy Lyster, Department of Integrated Studies in Education, McGill University, 3700 McTavish Street, Montreal, QC, Canada H3A 1Y2; e-mail: roy.lyster@mcgill.ca.

© 2004 Cambridge University Press 0272-2631/04 $12.00

https://doi.org/10.1017/S0272263104263021 Published online by Cambridge University Press

Previous studies of content-based second language (L2) programs have shown that students learning their L2 through immersion—where they learn much of their subject matter through the L2—develop high levels of comprehension skills as well as considerable fluency and confidence in L2 production (for recent overviews, see Lyster, 1999; Rebuffot, 1993; Swain, 1997). These studies have also pointed to certain shortcomings in grammatical accuracy that persist in immersion students' interlanguage development even after years of immersion education (e.g., Harley, Cummins, Swain, & Allen, 1990). This leveling-off effect has been explained by the fact that many grammatical features, such as certain verb tenses, occur only minimally in classroom discourse, whereas other features, such as grammatical gender, occur with considerable frequency yet lack saliency in the discourse (e.g., Allen, Swain, Harley, & Cummins, 1990; Swain, 1988). In both cases, researchers have underscored the importance of integrating form-focused instruction (FFI) into regular subject-matter instruction to allow students to notice, in the context of meaningful interaction, these otherwise infrequent or nonsalient features. FFI, according to Ellis (2001, pp. 1–2), refers to "any planned or incidental instructional activity that is intended to induce language learners to pay attention to linguistic form" (see also Spada, 1997; regarding "focus on form," see Doughty & Williams, 1998; Long, 1996). Previous quasi-experimental classroom studies have shown positive effects, in varying degrees, of FFI on immersion students' acquisition of aspectual and modal distinctions (Day & Shapson, 1991; Harley, 1989), sociolinguistic variation (Lyster, 1994), and grammatical gender (Harley, 1998).

Grammatical gender is a language feature that immersion students fail to master, even after many years of classroom exposure to French, and despite the fact that it occurs with considerable frequency in the input. Carroll (1989), for example, noted:

> English-speaking children in immersion programs have problems producing gender markers not only in spontaneous production but also in controlled experimental situations. They do not appear to have anything resembling native competence. Their acquisition of French gender appears to be as problematic as is that of English-speaking adults. (p. 575)

In light of these findings, Harley (1998) conducted a study in Grade 2 immersion classrooms, using form-focused activities designed to draw their attention to primarily phonological clues in noun endings that reliably predict grammatical gender. She found that these noticing activities helped these young children to learn the grammatical gender of high-frequency words they already knew but that they were unable to apply the rules to less familiar or unknown words.

Conducted at a higher grade level with 10–11-year-old students in Grade 5, the present study emphasizes indexes of grammatical gender that are not only phonological but also orthographic in nature. The FFI included noticing activities based on typographically enhanced texts, adapted from the students'

https://doi.org/10.1017/S0272263104263021 Published online by Cambridge University Press

subject-matter curriculum and inductive tasks designed to lead students to notice a set of orthographic and phonological patterns that determines grammatical gender in French. Additionally, the present study has as one of its primary objectives to examine the effects not only of FFI but also, more specifically, the effects of instruction that incorporates different oral feedback options—namely, recasts, prompts, or no feedback. The effects of these instructional variables are investigated from an information-processing perspective, drawing on Skehan's (1998) framework for understanding the relationship between cognitive processing and representational systems.

REPRESENTATION AND PROCESSING

Various researchers have invoked information-processing models to describe L2 learning as the acquisition of complex cognitive skills (e.g., DeKeyser, 1998, 2001; Hulstijn, 1990; Johnson, 1996; McLaughlin, 1987, 1990; O'Malley & Chamot, 1990; Towell & Hawkins, 1994). The notion of information processing was developed in cognitive psychology to describe skill acquisition as a gradual change in knowledge from declarative to procedural mental representations (Anderson, 1983, 1985) and in executive control from controlled to automatic processing of those mental representations (Shiffrin & Schneider, 1977). Regarding language, declarative knowledge refers to knowledge of the language system, such as word definitions and rule-based representations, whereas procedural knowledge refers to knowledge about how to perform cognitive activities, including language comprehension and production. Skill development depends on transforming declarative representations, through practice, into production rules that represent procedural knowledge. This development is described as a transition from controlled processing, which requires a great deal of attention and use of short-term memory, to automatic processing, which operates on routinized procedures available in long-term memory. The proceduralization of rule-based representations occurs through practice and feedback (Anderson, Corbett, Koedinger, & Pelletier, 1995), which are considered crucial elements in information-processing models of L2 development, because they engage learners in processes of restructuring interlanguage representations and, thus, have particular relevance in immersion contexts (Lyster, 1990; Ranta & Lyster, 2003).

An alternative view of skill acquisition is Logan's (1988) instance theory, whereby automatization involves, rather than proceduralization of rule-based representations with increasingly less attention, a transition from rule- to memory-based performance (see DeKeyser, 2001; Robinson & Ha, 1993; Schmidt, 1992, 2001). In this view, procedures initially deriving from rule-based representations become available as memory-based chunks, which then operate autonomously. With minimal computational demands, retrieval from the memory-based system involves more efficient processing, enables fluent performance, and is thus considered synonymous with automaticity. However, to

https://doi.org/10.1017/S0272263104263021 Published online by Cambridge University Press

account more specifically for L2 learning and variability in L2 performance, still from an information-processing perspective, Skehan (1998) described a dual-coding system, which instead reconciles both rule- and memory-based systems as equally important representational systems for L2 learners.

According to Skehan (1998), the dual-mode system comprises two interrelated representational systems: an analytic rule-based system and a memory-driven exemplar-based system. The rule-based system is generative and parsimoniously organized, with the advantage of its propensity for creative and precise computations but the disadvantage of a heavy processing burden during ongoing language use. In contrast, the exemplar-based system is composed of lexicalized chunks, stored in a redundant unstructured memory system, which have the advantage of being retrieved quickly during ongoing communication because they do not require excessive internal computation but the disadvantage of having only limited potential for expressing new and precise meanings. Access to the dual-mode system thus affects fluency, as learners engage differentially in retrieval, leading to either computed rule-based performance or memory-driven exemplar-based performance. Skehan argued that during online communication, "where communicative pressure and accessibility are paramount," the exemplar-based system will be the system of choice, thus "reducing the likelihood that the more open and generative rule-based system will be used" (p. 62). These two representational systems, however, are not entirely separate; instead, they are "in constant dialect," enabling L2 learners to engage in complementary processes of analysis and synthesis: "The analysis is necessary to enable the learner to gain generativity and flexibility, and the synthesis is necessary to enable fluency and control to be achieved" (Skehan, p. 92).

With respect to L2 development, Skehan (1998) argued that interlanguage changes are more effectively activated through the rule-based system than through the exemplar-based system. Moreover, he claimed that the rule-based system "is also more likely to be more sensitive to feedback since the precision and system which accounts for rule-organization will make the feedback more informative" (p. 88), whereas feedback may not be so effective in effecting change in the exemplar-based system, because exemplars are not part of a structured system with connections to rules that can produce general change. In further consideration of the potential effectiveness of feedback, or what he called "feedback appreciation," Skehan suggested that conscious awareness of feedback ("consciousness enhanced processing") predisposes the learner toward a rule-based perspective, which in turn is more likely to lead to longer-term change.

RECASTS AND PROMPTS

This section introduces a distinction between two very different types of feedback, which have been observed in immersion classrooms (Lyster & Ranta,

https://doi.org/10.1017/S0272263104263021 Published online by Cambridge University Press

1997) and presented as complementary teacher moves with different purposes for different learners in different contexts (Lyster, 2002a, 2002b). Notwithstanding the pedagogical justification for teachers to use both these moves "in accordance with their students' language abilities and content knowledge . . . without abandoning one at the expense of the other" (Lyster, 2002b, p. 251), this section presents a theoretical justification, by drawing on the information-processing framework described in the previous section, for comparing recasts with prompts in terms of their relative effectiveness in a classroom experiment. In the next section, the rationale is developed further by drawing on processing and representational difficulties specific to grammatical gender attribution.

Long (1996) and others (e.g., Gass, 1997; Pica, 1994) have proposed that the negotiation of meaning, defined as various input modifications and interactional moves including semantically contingent feedback such as repetition and reformulation, provides learners with implicit negative evidence and thus serves to benefit L2 development. One type of implicit reformulation move that has received increasing attention in both L1 and L2 contexts is the recast—a well-formed reformulation of a learner's nontarget utterance with the original meaning intact, an example of which follows (this and subsequent examples of feedback types were observed in either the present study or Lyster & Ranta's [1997] study and were selected because they target grammatical gender; M and F in parentheses indicate masculine and feminine determiners, respectively):

(1) Student: *Le guimauve*. "(M) Marshmallow."
Teacher: *La guimauve. Oui*. "(F) Marshmallow."

There is nothing disconfirming in the teacher's recast (*La guimauve*) that actually incites the learner to notice that *le guimauve* is ungrammatical; the learner could even infer from this and other classroom input that *le* and *la* are interchangeable variants. Some researchers, however, have argued that the juxtaposition of the learner's ungrammatical utterance with the teacher's reformulation provides the learner with an ideal opportunity to make a cognitive comparison and to notice the gap between the targetlike and nontargetlike forms, especially because meaning is held constant and so the learner's processing resources are freed up to focus on form (e.g., Doughty, 2001; Long; Long & Robinson, 1998).

A sampling of studies investigating the use of recasts demonstrates their effectiveness in some contexts, but not in comparison with other clearly identified feedback. For example, during experimentally controlled interaction with adults, Long, Inagaki, and Ortega (1998) showed that recasts can be more effective for some target features than *models*, which are provided to learners before they speak and, thus, are not a comparable type of semantically contingent feedback. Mackey and Philp (1998) showed that experimentally controlled interaction during which a researcher provides L2 learners with intensive recasts,

https://doi.org/10.1017/S0272263104263021 Published online by Cambridge University Press

which focus consistently on one type of structure, is more effective, in the case of adults developmentally ready to learn the target form, than interaction without intensive recasts (see also Han, 2002). With young learners in content-based classrooms, Doughty and Varela (1998) showed that *corrective recasting*, which includes a recast preceded by a repetition of the learner's error and intonational stress added to both repetition and recast to emphasize the incorrect and correct forms, is also more effective than no feedback. Ohta (2000) showed that four adult learners who produced private speech with moderate to high frequency while wearing lapel microphones in a form-oriented foreign language classroom tended to repeat recasts, regardless of whether they were the intended recipients of the recasts (cf. Nabei & Swain, 2002). For a comprehensive review of recast studies, see Nicholas, Lightbown, and Spada (2001).

Based on my own observation studies of immersion classrooms (e.g., Lyster, 1998b, 2002b), I have argued that recasting—as defined in the L1 literature (e.g., Marcus, 1993; Saxton, 1997) and as observed in immersion classrooms (i.e., an implicit targetlike reformulation of a learner's utterance)—is not the most effective way of providing young L2 learners with negative evidence in classrooms where the primary focus is on subject matter, especially in comparison with other feedback options. My analyses of immersion classroom discourse suggested that recasts are, for the most part, embedded in meaning-focused negotiation and are thus ideal for facilitating the delivery of complex subject matter, because they provide supportive, scaffolded help, which serves to move the lesson ahead when the target forms in question are beyond the students' current production abilities. Recasts used in this way compete with many other demands on attention during content-based instruction and appear to be ambiguous, because they share discourse functions with a similar proportion of teacher repetitions of well-formed utterances (Lyster, 1998b); that is, recasts of ill-formed utterances and repetitions of well-formed utterances together appear to confirm or disconfirm the meaning of a learner's message, not its form. Moreover, recasts often co-occur with signs of approval (i.e., positive feedback moves including affirmations such as *Oui*, *C'est ça*, and *OK* and praise markers such as *Très bien*, *Bravo*, and *Excellent*; see Lyster, 1998b), as teachers respond to the veracity of learners' messages. Similarly, Musumeci (1996) and Pica (2002) documented the difficulty of effectively drawing attention to form specifically during content-based instruction, even in the case of adult L2 learners. Indeed, researchers discussing a wide range of classroom contexts have noted the ambiguity of implicit feedback (e.g., Allwright, 1975; Chaudron, 1977; Fanselow, 1977; Long, 1977; Netten, 1991).

In addition to recasts, I observed that teachers have at their disposal a set of interactional moves that provide less ambiguous and more cognitively engaging feedback, referred to interchangeably as the "negotiation of form" (Lyster, 1998a, 2002a; Lyster & Ranta, 1997), "form-focused negotiation" (Lyster, 2002b), and "prompts" (Lyster, 2002b; Ranta & Lyster, 2003). Prompts provide a solution to Swain's (1985) call for immersion teachers to "push" their students to

https://doi.org/10.1017/S0272263104263021 Published online by Cambridge University Press

be more accurate in their output. They include the following four types of teacher response:

1. *Clarification requests* are phrases such as "Pardon me" and "I don't understand" used to indicate that the student's message has either been misunderstood or ill formed. For example:
   Student: *Et le coccinelle . . .* "And the (M) ladybug."
   Teacher: *Pardon?* "Sorry?"
   Student: *La coccinelle . . .* "The (F) ladybug."
2. *Repetitions* replicate the student's error verbatim, usually with rising intonation and stress to highlight the error. For example:
   Student: *La chocolat . . .* "(F) Chocolate."
   Teacher: *La chocolat?* "(F) Chocolate?"
   Student: *Le chocolat.* "(M) Chocolate."
3. *Metalinguistic clues* provide comments, information, or questions related to the well-formedness of the student's utterance. For example:
   Student: *Parce qu'elle cherche, euh, son, son carte.*
   "Because she's looking for, um, her, her (M) card."
   Teacher: *Pas son carte.* "Not her (M) card."
   Student: *Euh, sa carte?* "Um, her (F) card?"
4. *Elicitation* entails direct questions such as "How do we say that in French?" or pauses that allow students to complete the teacher's utterance. For example:
   Teacher: *Il vit où un animal domestique? Où est-ce que ça vit?*
   "Where does a pet live? Where does it live?"
   Student: *Dans un maison.* "In a (M) house."
   Teacher: *Dans . . . ? Attention.* "In . . . ? Careful."
   Student: *Dans une maison.* "In a (F) house."

Although these four prompting moves, used separately or in combination, represent a wide range of feedback types, they have one crucial feature in common: They withhold correct forms (and other signs of approval) and instead offer learners an opportunity to self-repair by generating their own modified response. This approach resembles the "clueing" procedure or "withholding phenomenon" identified by McHoul (1990) in his study of feedback in subject-matter classrooms. In contrast, recasts provide learners with correct target forms, which frequently co-occur with signs of approval.

In their study of cognitive processes generated by output, Swain and Lapkin (1995) posited that feedback enables learners to notice problems in their output and pushes them to conduct an analysis leading to modified output. Swain and Lapkin contended that what occurs between the first and second outputs is part of the process of L2 learning. I would add that the extent to which restructuring processes are activated between the learner's first and second outputs depends on the type of feedback. That is, not all types of feedback lead to similar degrees of pushing. Recasts, for example, tend not to push learners to modify their nontarget output, at least not immediately following feedback (e.g., Chaudron, 1977; Lyster & Ranta, 1997; Panova & Lyster, 2002;

https://doi.org/10.1017/S0272263104263021 Published online by Cambridge University Press

Pica, 1988, 2002). From a discourse perspective, a student's repetition of a teacher's recast can be seen as a redundant move in an exchange in which the teacher, by virtue of recasting, both initiates and completes the repair within a single move. On the small number of occasions when learners do modify their ill-formed utterances following recasts, the modification may simply be a repetition of the alternative form, involving retrieval from short-term memory rather than from long-term memory.

Ellis (1997) and others (e.g., Bialystok & Sharwood Smith, 1985; de Bot, 1996) have distinguished between two types of acquisition: acquisition as the internalization of new forms, and acquisition as an increase in control over forms that have already been internalized. In this view, as exemplars of positive evidence (see Braidi, 2002; Leeman, 2003), recasts occurring in appropriate discourse contexts can facilitate the encoding of new declarative knowledge. Prompts, on the other hand, given their aim to elicit modified output, can enhance control over already-internalized forms—that is, prompts serve to assist learners in the transition of declarative to procedural knowledge (regarding the psycholinguistic effect of output, see de Bot, p. 549).

In the immersion context, because learners have had years of exposure to L2 input, including the target forms that they consistently have problems acquiring, they need to be pushed, when their focus is on academic content, to use target forms that are in competition with highly accessible interlanguage forms (Ranta & Lyster, 2003; Swain, 1985). Prompts, therefore, may be particularly beneficial in immersion classrooms and other meaning-focused instructional contexts where continued recasting of what students already know may prove to be less effective for promoting the restructuring of interlanguage representations and the proceduralization of competing targetlike representations. This was the case in Ammar's (2003) recent classroom study, which revealed superior effects for prompts over recasts in the acquisition of possessive determiners by Grade 6 Francophone learners of English as a second language (ESL). She also found that prompts were particularly effective for lower proficiency learners, whereas higher proficiency learners appeared to benefit similarly from both recasts and prompts. Other studies as well have suggested that low-proficiency learners might be at a disadvantage in their limited ability to notice recasts as corrective feedback (Lin & Hedgcock, 1996; Mackey & Philp, 1998; Netten, 1991).

Whereas Doughty (2001) argued that recasting is an ideal type of feedback, because L2 learners are able to store the target reformulation in working memory and to make a direct comparison between input and output, de Bot (2000) suggested that "there is never a direct comparison between input and output because the input information is immediately processed and not stored in memory in that form" (p. 228). Conversely, in response to prompts, learners must attend to the retrieval from long-term memory of previously encoded representations, retrieving either an alternate exemplar or a rule for computing a more targetlike form. De Bot (1996, 2000) argued that this increased level of activation enhances the likelihood of the retrieved item being selected again,

https://doi.org/10.1017/S0272263104263021 Published online by Cambridge University Press

because the attention required for retrieval from long-term memory and subsequent production stimulates the development of strong connections in memory. By retrieving target forms stored in long-term memory, therefore, L2 learners are more likely to restructure existing interlanguage representations than by merely hearing the forms in the input (de Bot, 1996). The restructuring of knowledge representations and concomitant processing constraints specific to grammatical gender are addressed in the next section.

GRAMMATICAL GENDER

Grammatical gender, which is best defined in contrast to biological (or natural) gender, encompasses the gender of generic nouns—that is, inanimate nouns (e.g., *un chapeau* "a hat," *une casquette* "a cap") as well as a class of animate nouns, that refer generically, without identifying sex, to many living entities such as insects (e.g., *une mouche*, "a fly"), marine life (e.g., *une crevette* "a shrimp," *une truite* "a trout"), birds (e.g., *un moineau* "a sparrow," *une alouette* "a lark"), and various other animals (e.g., *une gazelle* "a gazelle," *un serpent* "a snake"). Biological gender, on the other hand, encompasses gender determined by sex, in reference to humans (e.g., *une fille* vs. *un garçon* "a girl" vs. "a boy"; *une actrice* vs. *un acteur* "an actress" vs. "an actor") and many animals (e.g., *un chat* vs. *une chatte* "a cat"; *un loup* vs. *une louve* "a wolf").

Carroll (1989) described a useful distinction between gender agreement and gender attribution. On the one hand, gender agreement, a variable characteristic of modifiers, is systematically derivative and rule-based and depends on the context of occurrence of modifiers with nouns. Gender attribution, on the other hand, is an inherent feature of nouns that remains invariable and independent of the context of occurrence. Through exposure to language arts materials often designed for and by native speakers (NSs) of French (Lyster, 1990), immersion students develop some knowledge of gender agreement, as they typically engage in grammatical exercises requiring the transformation of adjectival forms from masculine to feminine (*nouveau* → *nouvelle* "new"). Similarly, with respect to gender attribution, immersion students become familiar with morphological rules governing animate nouns whose inherent gender is determined by sex (e.g., *un chanteur* → *une chanteuse* "a singer"; *un chat* → *une chatte* "a cat"), again because these types of transformational exercises appear in their language arts materials. However, their materials do not focus on the inherent gender of generic nouns, because this information is acquired early by NSs and does not involve rules of grammatical spelling, which permeate grammar-based materials designed for NSs of French and their language arts classes (e.g., Fazio & Lyster, 1998). Consequently, the resulting gap in immersion students' knowledge of gender attribution seriously constrains their ability to compute appropriate matches between determiners and generic nouns.

Carroll (1989) shed further light on the processing difficulties experienced by immersion students by arguing that Anglophone learners of French, unlike

https://doi.org/10.1017/S0272263104263021 Published online by Cambridge University Press

NSs of French, do not learn determiners as parts of nouns but learn them instead as distinct syntactic words and independent phonological units. The morphosyntactic representations of nouns are thus different for Francophone and Anglophone children, and so it follows that the processing of these different representations entails different types of computation. In the case of learning French L1, a phonological representation initially co-occurs with a nominal constituent to form an unanalyzed chunk composed of determiner + noun (e.g., laporte [lapɔʀt] = the door). "When the learner figures out that the determiners are distinct lexical items, the phonological representations will be reduced and the morphosyntactic representation will be augmented to include [gender specification]" (Carroll, p. 578). Thus, NSs of French process determiners and nouns, at least initially, as coindexed chunks; when these chunks are later analyzed as separate constituents, the noun still retains its inherent gender specification. Anglophone learners of French L2 appear instead to acquire nouns and their determiners as separate entities, applying, on the one hand, their knowledge of definiteness, indefiniteness, and possession to acquire determiners such as *le/la*, *un/une*, and *mon/ma*, and, on the other, their conceptual and morphosyntactic knowledge of nouns in English to acquire corresponding lexical entries in French. Thus, Carroll suggests, in the case of Anglophone learners of French L2, the functions deriving underlying phonological representations do not chunk determiners with nouns, nor do the functions deriving morphosyntactic representations of nouns include gender specification. In her view, this explains the difficulty that Anglophone learners, both children and adults, have in retrieving the gender specification of nouns during online production, "although they might be able to resort to some heuristic, a testable prediction" (Carroll, p. 578).

The problem with resorting to some heuristic or rule-driven mnemonic device is that many French grammarians (e.g., Bosquart, 1998; Grevisse & Goosse, 1995) and, thus, L2 teachers and their instructional materials claim that gender attribution is arbitrary and unsystematic in the case of generic nouns (i.e., to be learned on an item-by-item basis). Yet, considerable counterevidence to the claim that gender attribution is not rule governed suggests instead that gender attribution, even in the case of generic nouns, is largely rule driven, based on word-internal structural properties. For example, based on lexicographical analyses of nouns appearing in *Le Petit Larousse*, Tucker (1967, p. 13) concluded that "gender does co-occur, in a systematic and predictable manner, with noun ending" (see also Batchelor & Offord, 1993; Huot, 1991; Surridge, 1986, 1995; Tucker, Rigault, Lambert, & Segalowitz, 1968). Some predictors are primarily orthographic, showing up in the spelling of suffixes. For example, the suffix *-elle* is reliably feminine, whereas the homophonous ending *-el* reliably predicts masculine gender. Other predictors are phonological, such as the reliably masculine ending /o/ and the reliably feminine ending /ɛt/, each with variable orthography: /o/ = *-eau*, *-ot*, *-op*; /ɛt/ = *-aite*, *-ète*, *-ette*.

Research undertaken in preparation for the present study also confirmed an unexpectedly high number of orthographic and phonological patterns at

https://doi.org/10.1017/S0272263104263021 Published online by Cambridge University Press

play in predicting grammatical gender (Lyster, 2004). Using a CD-ROM version of *Le Robert Junior Illustré*, I compiled a corpus of more than 10,000 common nouns in French organized according to some 100 endings that are either reliably feminine, reliably masculine, or ambiguous. Close to 80% of the nouns in this corpus were found to have endings that reliably predict their grammatical gender. This finding, which is at odds with traditional grammars, provided the impetus for the present study to investigate the extent to which FFI and corrective feedback effect change in immersion students' rule-based representations of grammatical gender.

RESEARCH QUESTIONS AND PREDICTIONS

The purpose of the present study is to test the prediction that immersion students will improve in their use of grammatical gender, if provided with helpful opportunities to notice the co-occurrence of particular noun endings with appropriate gender attribution. The study aims to effect change in learners' rule-based representational system by providing learners with opportunities for consciousness-enhanced processing via FFI, including opportunities for feedback appreciation induced to varying degrees through the provision of different types of feedback (prompts and recasts), which are hypothesized to trigger different degrees of cognitive processing and awareness. The study's research questions are thus formulated as follows:

1. Will FFI improve French immersion students' ability to accurately assign grammatical gender?
2. Is FFI more effective with feedback than without feedback?
3. Which type of feedback is more effective in FFI—recasts or prompts?

With respect to question 1, it is predicted that immersion students participating in form-focused activities will outperform students whose attention is not drawn to grammatical gender through such activities. This is because, without FFI, immersion students are left to their own devices to process grammatical-gender patterns, which may occur frequently in classroom input but without salience (in the absence of FFI) and are thus not easily noticeable—especially not in a sufficiently robust way that might make the rules accessible in subsequent production.

With respect to question 2, it is predicted that immersion students receiving feedback in addition to FFI will outperform immersion students receiving FFI without feedback. This prediction is based on the premise that feedback can be effective at drawing learners' attention to mismatches (Ellis, 1994; Schmidt & Frota, 1986) and on Skehan's (1998) claim that feedback appreciation effectively contributes to changes in the learner's rule-based representations.

https://doi.org/10.1017/S0272263104263021 Published online by Cambridge University Press

It remains difficult to predict outcomes pertaining to question 3 because, if the FFI is effective, then it seems reasonable to expect that students will be sufficiently primed to notice and to process either type of feedback. However, a slight advantage for learners receiving prompts is predicted, for three reasons. First, prompts will provide more opportunity for output practice and thus for proceduralization of newly acquired rule-based representations. Second, prompts are more likely than recasts to draw learners' attention to feedback, thus creating conditions for conscious awareness, which adds efficiency to working memory operations (Schmidt, 1990) and predisposes the learner toward a rule-based perspective that is more likely to effect change in long-term memory (Skehan, 1998). Third, according to Carroll (1989), the processing demands necessitated by accurate gender attribution during a learner's online production involve the coordinated retrieval of previously disconnected elements: a noun plus its gender attribution, manifest in a separately stored determiner but also apparent in the noun's internal structural properties (as highlighted by the FFI). One could thus expect learners with more opportunities to process these elements in a productive mode in response to prompts to benefit more than students provided with opportunities to process (or not) the same information in a receptive mode via recasting. This prediction is in keeping with Skehan's argument that interlanguage change is unlikely to arise in a receptive mode and then later become available to production (see also Swain, 1985).

Given the binary nature of grammatical gender in French, it could be argued that the learner's retrieval of more targetlike representations of grammatical gender does not require much conscious awareness, because it involves a simple on-the-spot computation (i.e., if *un* is wrong, then it must be *une*).[1] However, building on and moving beyond previous descriptive studies of feedback, the present study is designed to investigate not immediate learner responses (i.e., uptake) but rather the effects on L2 learning, in the short- and long-term, of prompts and recasts integrated into FFI. Whether learners can be led easily to repair their gender errors immediately following feedback is not used as a measure of effectiveness in the present study. Instead, whether learners can retrieve more targetlike forms at a later point in time is examined. Overall, it is expected that comparisons of feedback types that activate different cognitive processes will have much to contribute to theoretical discussions of oral feedback and its role in L2 development.

METHOD

Participants

The participants in this study were four Francophone teachers (three females and one male) and their eight classes of 10–11-year-old students in an early French immersion program. The students were all in Grade 5 at three differ-

https://doi.org/10.1017/S0272263104263021 Published online by Cambridge University Press

ent schools in the same school board on the island of Montreal. Their school day in Grade 5 was conducted 60% in French, and had been 60% French in Grade 4, 85% in Grades 1–3, and 100% in Kindergarten.

Each of the four teachers had two classes, and these two classes together formed one group or condition. Three groups (the treatment groups) received FFI on grammatical gender for approximately 5 weeks while the fourth group (the comparison group) continued with its regular curriculum, without receiving any special FFI. Of the three treatment groups receiving FFI, two groups each received a particular type of feedback. One group, designated as the *FFI-recast group*, received, in addition to FFI, recasts following their errors in grammatical gender, whereas the other group, designated as the *FFI-prompt group*, received prompts following errors in gender, in addition to FFI. The third treatment group, designated as the *FFI-only group*, received FFI but no particular type of feedback.

At a workshop held for participating teachers prior to the study, teachers were given the opportunity to select the feedback conditions that best suited their own teaching style. To underscore the present study's ecological validity, it is important to report that the three teachers had little difficulty in selecting their feedback condition, as they concurred that their respective selection reflected their preferred interactional style. Thus, the teacher of the FFI-recast group chose the recast condition, considering recasts sufficiently unobtrusive and in keeping with her instructional style; similarly, the teacher of the FFI-prompt group claimed that prompting best reflected her overall pedagogical orientation; and the teacher of the FFI-only group was enthusiastic about providing no feedback, reporting that he generally did not provide oral feedback because his focus as an immersion teacher, he said, was on meaning, not form.

Class sizes ranged from 22 to 28, with an average of 25 students per class. Specifically, the two classes in the FFI-recast group each had 22 and 23 students; each class in the FFI-prompt group had 28 students; the two classes in the FFI-only group consisted of 23 and 24 students; and the control group consisted of one class of 27 students and another with 28. There were thus 203 students participating in the treatment activities. However, 12 students did not return permission forms to participate in the testing, and another 12 identified French as a language spoken at home and obtained perfect or near-perfect scores on the pretests. These 24 students, evenly distributed across the four groups, were not used in the statistical analyses of progress over time, which resulted in a total of 179 students in the sample.

Procedures

The FFI was implemented in the six experimental classrooms for approximately 8–10 hours over 5 weeks in February and early March. During this time, members of the research team took turns periodically observing the imple-

https://doi.org/10.1017/S0272263104263021 Published online by Cambridge University Press

mentation of the treatment activities in all six classrooms and recorded field notes. Pretests were administered in January, immediate posttests in March (posttest 1), and delayed posttests in May (posttest 2) to students in all eight classrooms.

The comparison group received neither FFI nor any special feedback treatment, continuing with its regular program of study, which included the same commercially produced curriculum materials used in the treatment classrooms (minus the instructional unit designed for this study). The comparison group thus received the same subject-matter instruction in French as the treatment groups. The importance of including a group not exposed to the FFI is to ascertain whether mere exposure over a 4-month period to extensive input alone, via subject-matter materials replete with target features, is sufficient to effect any change in these students' ability to correctly assign grammatical gender.

The study was initially designed so that the teachers of the FFI-recast and FFI-prompt groups would continue to provide their respective type of feedback on gender errors after the instructional period, during the 8-week period separating posttest 1 and posttest 2, whereas the teacher of the FFI-only group was expected to continue withholding feedback on gender errors during this time. However, classroom observations confirmed that it was too difficult for teachers of the FFI-recast and FFI-prompt groups to continue providing feedback consistently and specifically on gender errors during this time, given that they had moved on to new curriculum material and were focusing on new subject matter and different language activities. As a result, the classroom observations that were made during the period from March to May confirmed that continued reinforcement of grammatical gender assignment via feedback in the FFI-recast and FFI-prompt groups occurred occasionally but inconsistently during the 8-week period between posttest 1 and posttest 2. Therefore, for the purposes of this study, the FFI and feedback treatments refer only to the 5-week period between the pretest administered in January and posttest 1 administered in March.

Workshop for Teachers

A daylong workshop was held for the participating teachers in December, approximately 6 weeks prior to the onset of the instructional treatment. At this time, they were provided with a 16-page teachers' guide summarizing the gender patterns and the feedback types. The first half of the morning session was devoted to the patterns of grammatical gender on which teachers would focus during the treatment. The second half of the morning session was devoted to the treatment materials, which were presented activity by activity. The afternoon session was devoted entirely to different types of corrective feedback, using examples to illustrate each type from previous classroom observation studies.

https://doi.org/10.1017/S0272263104263021 Published online by Cambridge University Press

Target Features

To select the 15 or so target endings used in this study, the research team analyzed the students' regular curriculum materials (see next section) spanning a 3-month period and selected the most frequent endings appearing in these materials. Of these, only endings that are reliable predictors of gender over 90% of the time were selected for use in this study; they appear in Table 1 along with typical examples from the students' materials. The preponderance of feminine endings reflects the finding that immersion students and other learners of French L2 tend to overassign the masculine unmarked form (Harley, 1998; Lyster, 2004) and thus need their attention drawn to the marked feminine form, as a means of inciting them to notice grammatical gender distinctions.

Treatment Materials

The form-focused instructional unit was implemented in the context of subject-matter instruction within an iterative process comprising three interrelated pedagogical components: (a) noticing activities employing typographically enhanced texts; (b) awareness activities employing inductive rule-discovery tasks and metalinguistic explanation; and (c) practice activities employing both

Table 1. Target endings

Endings	Examples	
Feminine targets		
-ie	*une poulie*	"a pulley"
-sion, -tion	*une mission, la population*	"a mission," "the population"
-té	*la vérité*	"the truth"
-ance, -ence	*la distance, la différence*	"the distance," "the difference"
-et(t)e, -aite	*une casquette, la traite*	"a cap," "the trade"
-otte	*une botte*	"a boot"
-aine, -eine	*une semaine, une veine*	"a week," "a vein"
-ine	*une machine*	"a machine"
-elle	*une rondelle*	"a puck"
-ure	*la nourriture*	"the food"
-se	*une valise*	"a suitcase"
-che	*une branche*	"a branch"
Masculine targets		
/ɑ̃/	*un continent, un gant*	"a continent," "a glove"
/ɛ̃/	*un dessin, un train*	"a drawing," "a train"
/o/	*un mot, un bureau*	"a word," "an office"
-age	*un voyage*	"a trip"

https://doi.org/10.1017/S0272263104263021 Published online by Cambridge University Press

analysis- and fluency-based tasks. The treatment unit was designed around the children's regular curriculum—a commercially produced set of materials, called *Mémo mag* (Larose, Le Petitcorps, Jutras, & Bissonnette, 1994), which, although designed for NSs of French, is used in many French immersion classroom in Quebec. These materials, which integrate language arts, history, and science into monthly dossiers, provided the communicative context for the present study. Because the endings of generic nouns and the grammatical gender they convey carry no meaning or communicative value on their own, it was necessary to adopt the children's regular subject-matter curriculum as the meaningful context for the form-focused intervention.

For the purposes of this study, the dossier for the month of February was selected, and the research team created a 33-page student workbook, titled *Mon cahier de découvertes: le masculin et le féminin* "My discovery journal: Masculine and feminine," a copy of which was provided to each student and teacher in the treatment groups. Teachers also received a binder, which included a 27-page set of guidelines specifying the objectives, procedures, and answers for each activity. Additionally, a compact disk, various flash cards, word lists, and transparencies required for the completion of some activities, were included in the binder.

The treatment materials served to supplement the students' regular curriculum in two ways. First, the workbook contained simplified and shorter versions of texts found in *Mémo mag* and thus served to facilitate the students' comprehension of some otherwise difficult texts intended for NSs of French. Second, these simplified texts provided the context for drawing students' attention to noun endings as predictors of grammatical gender. Typographical enhancement was used to highlight, in bold, the endings of target nouns embedded in these texts. As students read each text, they were asked to fill in the missing definite or indefinite article before each noun by checking its gender in the original text in *Mémo mag*. Once the cloze text was completed, students were asked to organize target nouns according to their gender and to induce the morphophonological rules governing the patterns they observed. Some of these tasks revolved around the students' history program (e.g., the founding of Quebec City and Montreal), whereas others pertained to their science program (e.g., simple machines).[2]

Teachers were also provided with an additional set of activities, which focused on the same endings but were not directly related to the regular curriculum, to use at any time throughout the study. These included various crossword puzzles and word searches related to themes such as sports, transportation, clothing, and Valentine's Day, as well as activities highlighting patterns of grammatical gender in songs and rhyming verses (*comptines*), after which students could create their own rhyming verses. Also included for teachers to use at any time throughout the treatment as a whole-class or small-group activity was a set of 152 laminated flash cards with riddles eliciting target words, either from the students' curriculum or other high-frequency lexical items. Finally, a set of 15 laminated posters, one for each noun ending and

https://doi.org/10.1017/S0272263104263021 Published online by Cambridge University Press

each listing many high-frequency nouns with that particular ending, were provided to teachers to place on the classroom walls to serve as a quick reference for students throughout the instructional unit.

MEASURES

There were four tests used to assess the students' ability to assign grammatical gender: two written tasks and two oral tasks.

Written Tasks

The written tasks involved two different tests: a binary-choice test and a text-completion test. The tests were administered to intact classes by a research assistant, and students were given sufficient time, which never exceeded more than 45 minutes, to complete both. There were two forms of each written test, A and B, which allowed for a counterbalanced test administration designed to reduce test-retest effects. Whereas lexical items varied across the two forms, target endings were held constant: for example, the words *château* "castle" and *piscine* "swimming pool" appearing on form A were replaced by *bureau* "desk" and *machine* "machine" on form B. Half the students, randomly selected, completed the written tasks following ABA test order across the three testing sessions, while the other half completed the tasks following BAB test order. Analysis of variance conducted on the pretest scores of each test showed that there were no differences between forms A and B, neither on the binary-choice test, $F(1, 178) = 0.11$, $p = .74$, nor on the text-completion test, $F(1, 178) = 0.36$, $p = .55$.

Binary-Choice Test. On each form of the binary-choice test, there were 48 items, each of which included a noun (e.g., *fraise* "strawberry," *fromage* "cheese," *peinture* "paint") preceded by a masculine and a feminine article, along with a drawing to serve as the noun's referent. Students were asked to circle the right article. The visual support was provided to contextualize to some extent these discrete-point items and to reduce any confusion or anxiety in the case of students unfamiliar with some of the lexical items. The Kuder-Richardson formula was used to measure test reliability based on pretest scores; a high coefficient of reliability of .86 was obtained.

Text-Completion Test. The text-completion test consisted of two different tasks. The first task was a cloze exercise designed as a reading activity, requiring students to assign grammatical gender to target nouns embedded in a coherent text. Students were asked on form A to read a text about different ways of spending a vacation; on form B, they read two short texts: one about how the heart works and the other about violence on television. The texts on

https://doi.org/10.1017/S0272263104263021 Published online by Cambridge University Press

each form contained 30 target nouns preceded by the masculine and feminine forms of the definite or indefinite article. Students were again asked to circle the appropriate article.

The second task on the text-completion test was a production task that required students to write a recipe in the form of a short text. Students were given a list of 23 food items (e.g., *1 piment* "pepper," *1 tomate* "tomato," *1 aubergine* "eggplant") and were asked to choose 10 ingredients that they would use to make either a pizza (on form A) or a magic potion (on form B). They were required to write their recipe using singular definite articles and were restricted from using the number "1" that had preceded each food item on the list of available ingredients.

Again using the Kuder-Richardson formula to measure test reliability on the basis of pretest scores, the 30-item cloze exercise on the text-completion test yielded a high reliability coefficient of .90. The 10-item production task did not lend itself to this statistical procedure, given the slightly more creative nature of the task and the fact that the lexical choices students could make were less constrained.

Oral Tasks

The two oral production tasks, both adapted from Harley (1998), included an object-identification test and a picture-description test. Each test was respectively administered to a randomly selected subsample of 60 students (15 per group) on three occasions. Unlike the written tasks, the oral measures were not counterbalanced, having only one form each, and so participants in this subsample completed the same oral tasks on three occasions. Each participant was withdrawn from class and tested individually by a researcher who conducted the test as an interview in a small room while she audio-recorded their interaction.

In an earlier pilot study using both oral tests, the interviewer had intervened as little as possible, but, in the end, we were left with a considerable amount of untranscribable data. As documented by Harley (1998), immersion students have developed a strategy of using hybrid forms that sound like a combination of both *un* and *une* or *le* and *la*. This strategy eases the L2 learner's burden of having to accurately mark grammatical gender so frequently. Data collected during the piloting caused hours of discussion among members of the research team as we debated whether students said *un* or *une* or a hybrid form. Therefore, for the purposes of the present study, the interviewer asked students to clarify their responses when she was uncertain whether they used masculine or feminine forms.

A team of three NSs and one near-native speaker of French transcribed all the interaction that occurred between the student and the interviewer. Then, this same team coded the oral data for accuracy by scoring learners' first use of a target item (or immediate self-initiated self-repair), to capture as much as

https://doi.org/10.1017/S0272263104263021 Published online by Cambridge University Press

possible their spontaneous production and to minimize the effects of the interviewer's role in the interaction. Interrater reliability was high for both oral production tasks: 98% agreement for the object-identification test and 92% for the picture-description test.

Object-Identification Test. During the object-identification test, the interviewer showed the student a set of two drawings of the same object. For example, as the interviewer pointed to drawings of two rafts, she named these objects, *Ce sont deux radeaux* "These are two rafts," 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 radeau* "It's a raft." After doing an example with the student to ensure comprehension of the task, the interviewer proceeded to present 20 such items.

The majority of items used on this test were low-frequency noncognate items, presumed to be relatively unfamiliar to these Grade 5 students: for example, *lotte* "monkfish," *tondeuse* "lawnmower," *perceuse* "electric drill," *volant* "steering wheel," and *barbotte* "catfish." These items were selected as a way of testing whether or not students could correctly assign grammatical gender to relatively unfamiliar words. For this reason, the interviewer always provided the lexical item to the student, but not its grammatical gender, by using the gender-neutral plural determiner.

Picture-Description Test. For the picture-description test, the interviewer used a colorful 20 × 32 inch drawing of an imaginary scene on a farm, which depicted a cow jumping over the moon, a turtle resting on a raft, a rabbit eating a giant carrot in a vegetable garden, a snake playing a trumpet, and a variety of other unusual images that were designed to pique the student's curiosity and elicit a minimum of 25 target words. The student was asked to either tell a story about the scene or simply describe as much of the scene as possible. This test was thus more open-ended than the object-identification test, in that it was more interactive and allowed for more extensive speech. On the one hand, this test was more demanding in that students were generally not provided with the lexical items and instead were expected to retrieve the vocabulary as well as the correct grammatical information regarding gender. On the other hand, to accommodate these demands, the lexical items that the students were expected to generate were, for the most part, high-frequency nouns such as *gateau* "cake," *cadeau* "gift," *nuage* "cloud," and *voiture* "car" and included familiar cognates such as *carotte* "carrot," *bicyclette* "bicycle," *jardin* "garden," and *chaise* "chair." Nonetheless, in the case of students who were unable to generate certain lexical items, the interviewer was allowed to provide the item, without identifying its gender, by saying something like: *On dit* "*poubelle*." "We say 'trash can.'"

https://doi.org/10.1017/S0272263104263021 Published online by Cambridge University Press

RESULTS

During the 5-week form-focused treatment period, each of the three treatment teachers was observed by one of two researchers for an average of about 10 hours. During the 8-week interval between posttests 1 and 2, the FFI-recast and FFI-prompt groups were observed, on average, for an additional 17.5 hours each. Classroom observations confirmed that all three teachers devoted a similar amount of time to the instructional unit, ranging from about 8 hours for the FFI-prompt group, to 8.5 hours for the FFI-only group, and 10 hours for the FFI-recast group.

Binary-Choice Test Results

The descriptive statistics for the binary-choice test, including group means and standard deviations for each group over time, appear in Table 2, and the group means are plotted on the graph in Figure 1. These results show that the FFI-prompt group outperformed the other three groups on posttests 1 and 2.

The analysis of variance confirmed that there were significant effects for group, $F(3, 175) = 15.88$, $p < .01$, and for time, $F(2, 174) = 52.24$, $p < .01$, as well as a significant Time × Group interaction, $F(6, 348) = 9.99$, $p < .01$. Tukey's post hoc pairwise comparisons were used to detect the source of the significance (an alpha level of .05 was set for these and all subsequent post hoc analyses), revealing, first, that there were no differences among the four groups at the time of pretesting. Second, the FFI-prompt group performed significantly better than the other three groups on posttest 1, at which time the FFI-recast group also significantly outperformed the comparison group. No significant differences were detected on posttest 1 between the FFI-recast and FFI-only groups, nor between the FFI-only and comparison groups.

At the time of posttest 2, the FFI-prompt group continued to perform significantly better than the other three groups. Additionally, the FFI-only group made some gains at posttest 2, which resulted in a clear treatment effect for the FFI, as all three treatment groups significantly outperformed the compar-

Table 2. Group means and standard deviations for binary-choice test

	Pretest		Posttest 1		Posttest 2	
Groups	*M*	*SD*	*M*	*SD*	*M*	*SD*
FFI-prompt ($n = 49$)	33.3	5.8	43.0	6.3	42.5	5.6
FFI-recast ($n = 38$)	33.1	6.7	38.6	5.9	38.6	7.2
FFI-only ($n = 41$)	34.4	7.1	35.8	7.6	36.3	7.2
Comparison ($n = 51$)	31.1	6.6	32.2	6.6	32.1	6.9

https://doi.org/10.1017/S0272263104263021 Published online by Cambridge University Press

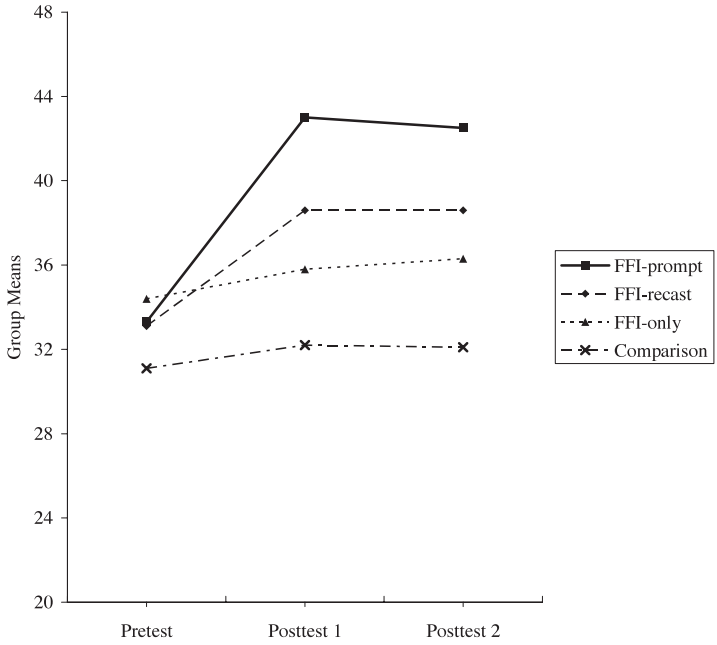


Figure 1. Group means on binary-choice test over time.

ison group. However, there were no statistically significant differences between the FFI-recast and FFI-only groups at the time of posttest 2.

Text-Completion Test Results

Group means and standard deviations for the text-completion test appear in Table 3, and the means are plotted on the graph in Figure 2. These results again show that the FFI-prompt group outperformed the other three groups on posttests 1 and 2.

Results of the analysis of variance again confirmed significant effects for group, $F(3, 175) = 11.41$, $p < .01$ and for time, $F(2, 174) = 83.70$, $p < .01$, as well as a significant Time × Group interaction, $F(6, 348) = 15.10$, $p < .01$. The post hoc comparisons again revealed that there were no differences among the four groups at the time of pretesting and confirmed the following significant differences at the time of posttest 1. First, the superior performance of the FFI-prompt group again proved to be significantly different from all other groups. Second, treatment effects for the FFI, regardless of feedback type, again proved to be significant, as all three treatment groups outperformed the comparison group at a level of significance. However, no differences were detected between the FFI-recast and FFI-only groups.

https://doi.org/10.1017/S0272263104263021 Published online by Cambridge University Press

Table 3. Group means and standard deviations for text-completion test

	Pretest		Posttest 1		Posttest 2	
Groups	*M*	*SD*	*M*	*SD*	*M*	*SD*
FFI-prompt (n = 49)	25.0	4.7	35.3	5.0	33.8	5.8
FFI-recast (n = 38)	24.6	6.8	29.7	6.4	28.8	6.0
FFI-only (n = 41)	26.6	5.1	29.6	6.1	28.8	6.1
Comparison (n = 51)	25.4	4.9	26.3	4.3	25.9	5.1

At the time of posttest 2, the significant effect for FFI was not consistently maintained across treatment groups, as the post hoc comparisons detected no differences among the FFI-recast, FFI-only, and comparison groups. However, the FFI-prompt group maintained its superior gains, significantly outperforming all three groups.

Object-Identification Test Results

The object-identification test was scored out of a possible 20 points. Table 4 presents the group means and standard deviations for each condition over

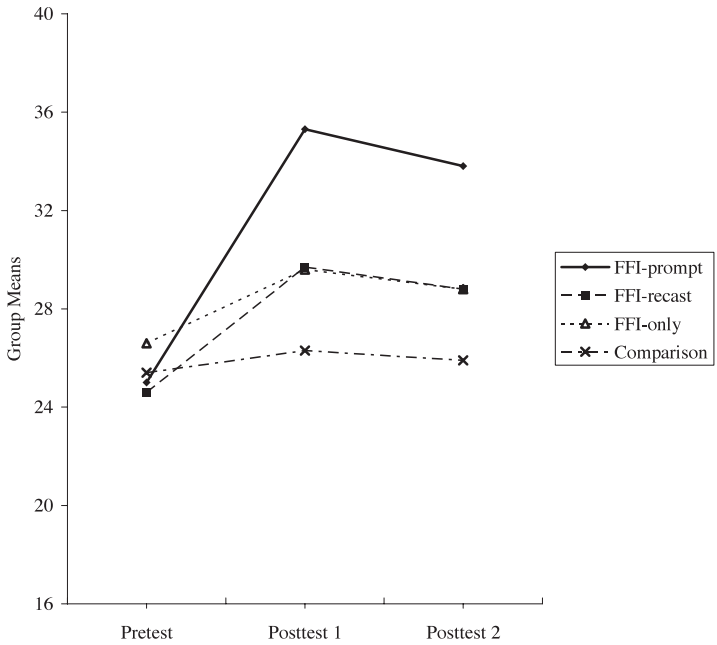


Figure 2. Group means on text-completion test over time.

https://doi.org/10.1017/S0272263104263021 Published online by Cambridge University Press

Table 4. Group means and standard deviations for object-identification test

	Pretest		Posttest 1		Posttest 2	
Groups	*M*	*SD*	*M*	*SD*	*M*	*SD*
FFI-prompt (n = 15)	11.5	3.0	15.8	4.3	16.6	3.3
FFI-recast (n = 15)	12.5	3.8	15.9	3.9	16.3	3.9
FFI-only (n = 15)	14.1	3.3	15.6	3.9	16.4	3.5
Comparison (n = 15)	10.1	2.1	11.8	3.0	10.9	3.0

time, and Figure 3 displays the group means graphically. Unlike the results of the written tasks, these descriptive statistics show that, although the three treatment groups clearly outperform the comparison group, their performance appears to converge over time in a similar pattern.

The analysis of variance revealed that there was a significant effect for group, $F(3, 56) = 7.73$, $p < .01$, and for time, $F(2, 55) = 23.57$, $p < .01$, and a significant Time × Group interaction, $F(6, 110) = 2.28$, $p < .05$. However, Tukey's post hoc pairwise comparisons detected significant differences between the FFI-only group and the comparison group at the time of pretesting, thus neces-

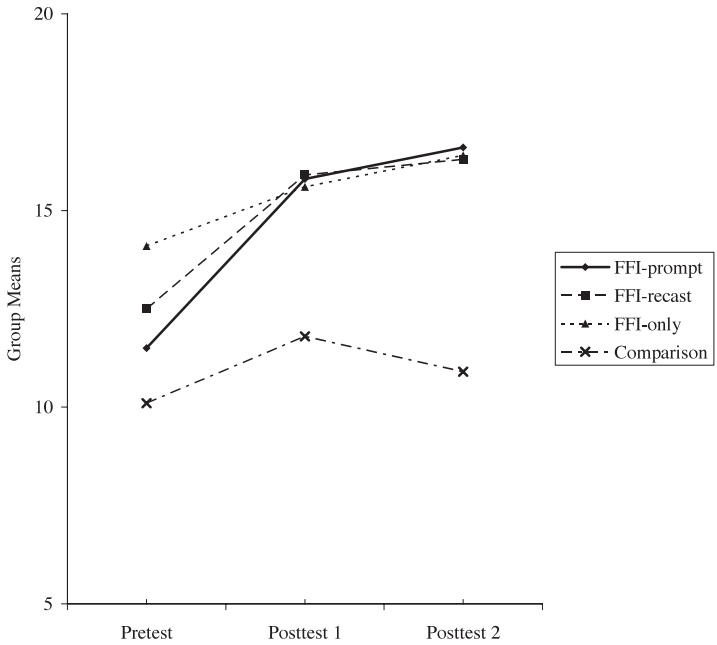


Figure 3. Group means on object-identification test over time.

https://doi.org/10.1017/S0272263104263021 Published online by Cambridge University Press

sitating an ANCOVA to adjust for initial discrepancies. Following Bonate's (2000) recommendation and Conover and Iman's (1982) finding that rank transformation procedures increase statistical power and robustness in analyses of covariance, raw test scores from the object-identification test were converted to rank scores. Table 5 displays the mean ranks for posttests 1 and 2, which have been adjusted by using mean pretest ranks as covariates.

The analysis of covariance, performed on the adjusted posttest rank scores, revealed a significant group effect at the time of posttest 1, $F(3, 55) = 3.09$, $p < .05$. The post hoc comparisons, using the Tukey-Kramer adjustment for multiple comparisons, revealed a significant difference between the FFI-prompt group and the comparison group. The significant group effect was maintained and strengthened at the time of posttest 2, $F(3, 55) = 5.83$, $p < .01$, at which time all three treatment groups, according to the post hoc comparisons, proved to be significantly different from the comparison group.

Picture-Description Test Results

The picture-description test was scored out of 25 points. The group means and standard deviations for each condition over time appear in Table 6, and the means are graphed in Figure 4. The means reveal gains made by all treatment groups in comparison to the lack of gains made by the comparison group and show a pattern wherein the FFI-prompt group outperforms the others. ANOVA was again used to confirm the statistical significance of these apparent differences.

The ANOVA results revealed significant effects for group, $F(3, 56) = 3.37$, $p < .05$, and for time, $F(2, 55) = 22.81$, $p < .01$, as well as a significant Time × Group interaction, $F(6, 110) = 3.35$, $p < .01$. Tukey's post hoc pairwise comparisons revealed no differences at the time of pretesting and, at the time of posttest 1, detected a significant difference only between the FFI-prompt group and the comparison group. This significant difference was maintained at the time of posttest 2, at which point the FFI-recast and FFI-only groups made further progress so that all three treatment groups significantly outperformed the comparison group.

Table 5. Adjusted mean rank scores used in ANCOVA of object-identification posttests 1 and 2

Groups	Posttest 1	Posttest 2
FFI-prompt ($n = 15$)	37.6	37.7
FFI-recast ($n = 15$)	35.4	34.8
FFI-only ($n = 15$)	30.1	32.4
Comparison ($n = 15$)	21.4	18.2

https://doi.org/10.1017/S0272263104263021 Published online by Cambridge University Press

Table 6. Group means and standard deviations for picture-description test

	Pretest		Posttest 1		Posttest 2	
Groups	*M*	*SD*	*M*	*SD*	*M*	*SD*
FFI-prompt (n = 15)	14.3	3.9	19.9	3.9	19.9	3.5
FFI-recast (n = 15)	14.1	4.9	18.2	5.2	18.9	5.3
FFI-only (n = 15)	16.2	4.1	17.7	5.1	18.6	5.3
Comparison (n = 15)	13.9	3.1	14.2	3.0	14.1	3.3

Summary of Results

Question 1 asked: Will FFI improve French immersion students' ability to accurately assign grammatical gender? The analysis of variance conducted on each of the four measures consistently revealed significant effects for time and for group, as well as significant Time × Group interactions. The post hoc comparisons revealed that the three treatment groups outperformed the comparison group on posttest 1 of the text-completion test and on posttest 2 of the binary-

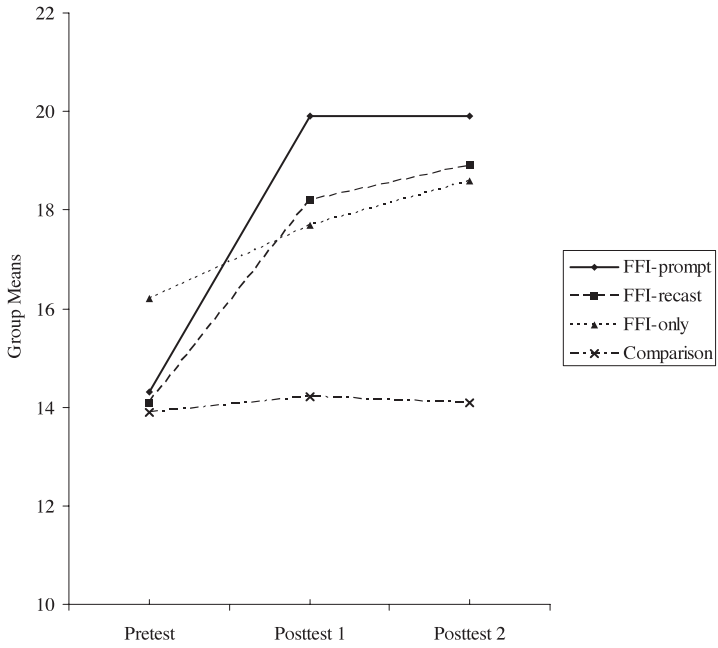


Figure 4. Group means on picture-description test over time.

https://doi.org/10.1017/S0272263104263021 Published online by Cambridge University Press

choice test, the object-identification test, and the picture-description test. Although the effects for instruction were significant for all three treatment groups at the time of posttest 1 on only one measure, all three treatment groups significantly outperformed the comparison group on the other three measures at the time of posttest 2, which occurred 2 months after the instruction. These results allow an unequivocally affirmative response to question 1, confirming the prediction that FFI would improve students' ability to accurately assign grammatical gender. The comparison group made no significant progress on any measures over time, which thus confirms that, without instruction designed to draw attention to co-occurrences of gender attribution with particular noun endings, students in the comparison group were unable to simply infer the patterns from the input and thereby improve their ability to accurately assign grammatical gender.

Question 2 asked: Is FFI more effective with feedback than without feedback? Test results do not permit an unequivocally positive affirmation of the effectiveness of feedback in general terms because the answer to this question varies according to type of feedback and task modality. In written tasks, students receiving prompts significantly outperformed students receiving recasts, who, contrary to initial predictions, performed similarly to students receiving no feedback. In oral tasks, also contrary to initial predictions, all three treatment groups performed similarly, regardless of feedback condition.

Finally, question 3 asked: Which type of feedback is more effective in FFI—recasts or prompts? The post hoc comparisons summarized in Table 7 reveal that the FFI-prompt group distinguished itself by being the only group to significantly outperform the comparison group on all eight measures.

The FFI-recast group significantly outperformed the comparison group on five of the eight measures, whereas the FFI-only group significantly outperformed the comparison group on four of the eight measures. FFI with recasts thus proved to be only marginally more effective than FFI without feedback, with no statistically significant differences appearing between the FFI-recast and FFI-only groups on any measures.

Table 7. Significant contrasts detected by post hoc comparisons of posttest measures

	Written tasks		Oral tasks	
Test	Binary choice	Text completion	Object identification	Picture description
Posttest 1	P > R, I, C R > C	P > R, I, C R, I > C	P > C	P > C
Posttest 2	P > R, I, C R, I > C	P > R, I, C	P, R, I > C	P, R, I > C

Note. P = FFI-prompt; R = FFI-recast; I = FFI-only; C = Comparison.

https://doi.org/10.1017/S0272263104263021 Published online by Cambridge University Press

The superior performance of the FFI-prompt group was especially evident in the written tasks, as it significantly outperformed all other groups on both posttests. Its performance in oral production was strong but did not set it apart from the other treatment groups to the same extent that it did on the written tasks. The FFI-prompt group was alone in significantly outperforming the comparison group on both oral tasks at posttest 1 but then performed similarly to the FFI-recast and FFI-only groups at posttest 2, when all three groups significantly outperformed the comparison group.

DISCUSSION

The results of the written and oral production tasks together indicate that FFI is clearly effective at improving students' ability to correctly assign grammatical gender. The results of the written tasks in particular and, to a much smaller degree, the results of the oral production tasks suggest that FFI is even more effective when combined with feedback in the form of prompts. The oral production results, while showing a clear effect for FFI, did not distinguish among feedback conditions to the same extent as the written tasks, although the FFI-prompt group displayed a significant advantage at the time of immediate posttesting on the oral tasks. It might be argued, therefore, that prompting affects online oral production skills only minimally, serving instead to increase students' metalinguistic awareness and their ability to draw on declarative rule-based representations on tasks where they have sufficient time to monitor their performance.

However, it may be argued instead that the overall similarity of the three treatment groups in oral performance, in contrast to the clear differences that emerged on the written tasks, was the result of a large task effect that led the subsample of 60 students to perform as a single group throughout this study. By virtue of being randomly selected as subsample participants, withdrawn individually from classes on three occasions to complete the oral tasks with the interviewer, this subsample of 60 students received more individualized attention than other participants in this study. Each session with the interviewer lasted approximately 30 minutes and was purposely administered with rigor and in identical fashion across all conditions. The oral tasks, unlike the written tasks, were not counterbalanced, so the students were exposed to the same routine on three occasions. Moreover, as previously mentioned, to counter the students' well-attested strategy of using a hybrid article and to thereby ensure the collection of transcribable data, the interviewer consistently prompted students to articulate gender-specific determiners as clearly as possible. It seems to be the case that these participants, regardless of their feedback condition, benefited greatly from the three intensive testing sessions, because the opportunities to interact one-on-one with a near-native speaker of French provided them with valuable oral practice to an extent that was impossible to match in class. Although participants from the comparison

https://doi.org/10.1017/S0272263104263021 Published online by Cambridge University Press

group were also required to interact with the interviewer during oral testing sessions, they had not been exposed to the FFI and, thus, had not been given the same opportunities to develop any rule-based representations of grammatical gender on which to draw during the oral tasks. That they were clearly outperformed by students exposed to the FFI underscores the effectiveness of providing L2 learners with opportunities to acquire declarative rule-based representations of grammatical gender through noticing and practice activities in the context of their subject-matter instruction.

It is noteworthy to mention here the effect of intervening variables in other quasi-experimental classroom studies. Harley (1989), Day and Shapson (1991), and Spada and Lightbown (1993) all reported that teachers in their comparison groups provided instruction that resembled the experimental treatment, the effects of which neutralized potentially significant differences between treatment and comparison groups. In the present study, it was not the comparison teacher who employed aspects of the form-focused treatment but rather the interviewer herself who, by necessity, employed prompting moves in her interaction with students during the oral production tasks. Similarly, Ellis (1984) noted that learners in his study also appeared to benefit from the interaction in which they engaged during audio-recorded test sessions administered before and after the classroom treatment.

It is important to reiterate that the interviewer refrained from providing students with the correct form and, instead, consistently provided prompts for clarification. Participants who had the benefit of FFI appear to have benefited significantly from this type of interaction, whose effectiveness may well have neutralized the effects of specific feedback types that were more distinguishable on the written tasks completed by all students. This interpretation presents a strong argument in favor of providing learners with rich opportunities for language practice and particularly for intensive practice involving interaction that aims to elicit target forms that are the current focus of instruction in class. For similar arguments, see DeKeyser's (1998) skill acquisition perspective on task sequencing and Ranta and Lyster's (2003) proposal for a pedagogical sequence of awareness, practice, and feedback as a means of improving immersion students' oral production abilities.

CONCLUSION

The present study revealed significant improvement in written and oral tasks for all groups receiving FFI. Their overall success on all measures, including those with high-frequency lexical items as well as one with low-frequency unfamiliar lexical items, suggests that the learning resulting from the FFI enabled learners to develop abstract rule-based knowledge of grammatical gender and not merely exemplar-based knowledge. This differs from Harley's (1998) finding that FFI presented to Grade 2 students was "more successful in inducing 'item learning' than 'system learning'" (p. 168). It is likely the case that the

https://doi.org/10.1017/S0272263104263021 Published online by Cambridge University Press

older learners in the present study were more predisposed, by virtue of their greater cognitive maturity, to induce rules and that, concomitantly, the form-focused intervention was presented with more intensity and repetitiveness, owing in part to the older learners' access to a much larger set of stored vocabulary items.

As for differences among feedback conditions, results of the written tasks indicated that (a) prompts were more effective than recasts and more effective than no feedback, and (b) students receiving recasts or no feedback performed similarly. These findings support the claim that learners who are prompted to retrieve more targetlike forms are more likely to retrieve these forms during subsequent processing than learners merely hearing recasts of these forms (de Bot, 1996, 2000; Lyster, 2002b). That learners benefited less from recasting may also be due to the ambiguity created specifically by recasts of gender errors or more generally to the difficulty learners have in noticing recasts of morphosyntactic errors (Mackey, Gass, & McDonough, 2000). It seems reasonable to conclude as well that students receiving prompts developed more "feedback appreciation" as a result of increased opportunities for conscious awareness of their teacher's feedback and were thus more predisposed toward a rule-based perspective, which in turn led to robust changes in their rule-based representations of grammatical gender, as predicted by Skehan's (1998) proposal for consciousness-enhanced processing. Finally, providing learners with opportunities to engage with feedback in a productive mode via prompting proved to be more effective in improving their ability to assign grammatical gender in written tasks than engaging students more receptively via recasting.

In oral production, differences among the three treatment groups were marginal. The lack of more significant differentiation on oral tasks across treatment groups is arguably due to a large task effect, which makes it difficult, on the one hand, to draw clear conclusions about the effects of different types of feedback on students' oral performance but, on the other, provides support for the effectiveness of prompting. That is, the subsample of students receiving FFI who participated in the oral production tasks performed more or less similarly as a single group throughout this study, because they invariably benefited from the one-on-one oral interaction with the interviewer and from the prompts she provided. Notwithstanding the subsample's overall similarity in oral performance, the FFI-prompt group distinguished itself from the other treatment groups on both oral tasks by significantly outperforming the comparison group at the time of immediate posttesting. The finding that students receiving prompts proved more likely than other students, at least in the short term, to manage the retrieval of grammatical information during online production is underscored here as a valuable topic for further research.

The instructional choices made by teachers throughout this study, beginning with their choice of feedback, appropriately evoke a bigger picture where choice of feedback and other pedagogical preferences reveal different language teaching approaches. Comparisons of these approaches in the present

https://doi.org/10.1017/S0272263104263021 Published online by Cambridge University Press

study point to the effectiveness of FFI overall, particularly when implemented in combination with prompts, as a means of enabling L2 learners, in the context of subject-matter instruction, to acquire rule-based representations of grammatical gender and to proceduralize their knowledge of these emerging forms. The present study thus contributes to theoretical arguments underpinning FFI by demonstrating its effectiveness when implemented in the context of subject-matter instruction within an iterative process comprising three interrelated pedagogical components:

1. Learners are led to notice frequent co-occurrences of appropriate gender attribution with selected noun endings, contrived to appear salient by means of typographical enhancement.
2. Learners' metalinguistic awareness of orthographic and phonological rules governing gender attribution is activated through inductive rule-discovery tasks and metalinguistic explanation.
3. Learners engage in complementary processes of analysis and synthesis (Klein, 1986; Skehan, 1998) through opportunities for practice in associating gender attribution with noun endings.

With respect to corrective feedback, the study contributes significantly to debates with both theoretical and practical relevance by confirming that recasts, when compared to other feedback options, are not necessarily the most effective type of feedback in communicatively oriented classrooms. In terms of pedagogical implications, one might not expect these findings to generalize to other instructional contexts, yet Ammar (2003) drew similar conclusions in the context of intensive ESL instruction, which entails communicative language teaching but not content-based instruction. In the latter type of instructional setting, immersion teachers can expect to continue their balancing act, varying their use of prompts and recasts according to context (Lyster, 2002a, 2002b) but need to do so in more systematically planned ways, taking into account the cognitive dimensions of processing and representation associated with specific target features in L2 learners' developing interlanguage system.

(*Received 12 December 2003*)

NOTES

1. Although binary at the word level, the choice actually requires multiple computations at the sentence level (and is likely to have further repercussions at the discourse level). Consider the following two sentences (in which meaning is held constant: "His new white bicycle is beautiful"):

Son nouveau vélo blanc est beau.
Sa nouvelle bicyclette blanche est belle.

Although the choice of gender is binary (*vélo* is masculine, not feminine; *bicyclette* is feminine, not masculine), the effects on morphosyntax are multiple. The focus of the present study, however, is restricted to gender attribution (at the word level) and not on noun-adjective agreement.

https://doi.org/10.1017/S0272263104263021 Published online by Cambridge University Press

2. For example, *Mémo mag*'s version of the founding of Quebec City ("*La fondation de Québec*") was presented in the students' workbook as a three-paragraph synopsis of this text that contains blanks preceding target nouns with highlighted endings, in bold, as in the following extract:

> *Québec ressemblait de plus en plus à ___ vrai vill**age**, doté notamment de ___ deuxième habita**tion** de Champlain, d' ___ chap**elle**, d' ___ magas**in** et d'autres bâtiments. Sur ___ plat**eau** au-dessus du cap Diamant, il y avait un fort, quelques maisons et ___ petite égl**ise** avec son presbytère.*
>
> "Quebec was looking more and more like a real village, featuring in particular Champlain's second residence, a chapel, a store, and other buildings. On the plateau above Cape Diamond, there was a fort, some houses, and a small church with a presbytery-house."

REFERENCES

Allen, P., Swain, M., Harley B., & Cummins, J. (1990). Aspects of classroom treatment: Toward a more comprehensive view of second language education. In B. Harley, P. Allen, J. Cummins, & M. Swain (Eds.), *The development of second language proficiency* (pp. 57–81). New York: Cambridge University Press.

Allwright, D. (1975). Problems in the study of the language teacher's treatment of learner error. In M. Burt & H. Dulay (Eds.), *New directions in second language learning, teaching, and bilingual education: On TESOL '75* (pp. 96–109). Washington, DC: TESOL.

Ammar, A. (2003). *Corrective feedback and L2 learning: Elicitation and recasts*. Unpublished doctoral dissertation, McGill University, Montreal.

Anderson, J. (1983). *The architecture of cognition*. Cambridge, MA: Harvard University Press.

Anderson, J. (1985). *Cognitive psychology and its implications*. New York: Freeman.

Anderson, J., Corbett, A., Koedinger, K., & Pelletier, R. (1995). Cognitive tutors: Lessons learned. *The Journal of Learning Sciences*, *4*, 167–207.

Batchelor, R., & Offord, M. (1993). *Using French: A guide to contemporary usage*. New York: Cambridge University Press.

Bialystok, E., & Sharwood Smith, M. (1985). Interlanguage is not a state of mind. *Applied Linguistics*, *6*, 101–117.

Bonate, P. (2000). *Analysis of pretest-posttest designs*. Boca Raton, FL: Chapman & Hall/CRC.

Bosquart, M. (1998). *Nouvelle grammaire* [New grammar]. Montreal: Guérin.

Braidi, S. (2002). Reexamining the role of recasts in native-speaker/nonnative-speaker interactions. *Language Learning*, *52*, 1–42.

Carroll, S. (1989). Second-language acquisition and the computational paradigm. *Language Learning*, *39*, 535–594.

Chaudron, C. (1977). A descriptive model of discourse in the corrective treatment of learners' errors. *Language Learning*, *27*, 29–46.

Conover, W., & Iman, R. (1982). Analysis of covariance using the rank transformation. *Biometrics*, *38*, 715–724.

Day, E., & Shapson, S. (1991). Integrating formal and functional approaches to language teaching in French immersion: An experimental study. *Language Learning*, *41*, 25–58.

de Bot, K. (1996). The psycholinguistics of the output hypothesis. *Language Learning*, *46*, 529–555.

de Bot, K. (2000). Psycholinguistics in applied linguistics: Trends and perspectives. *Annual Review of Applied Linguistics*, *20*, 224–237.

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). New York: Cambridge University Press.

DeKeyser, R. (2001). Automaticity and automatization. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 125–151). New York: Cambridge 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). New York: Cambridge University Press.

https://doi.org/10.1017/S0272263104263021 Published online by Cambridge University Press

Doughty, C., & Williams, J. (1998). Pedagogical choices in focus on form. In C. Doughty & J. Williams (Eds.), *Focus on form in classroom second language acquisition* (pp. 197–261). New York: Cambridge University Press.
Ellis, R. (1984). Can syntax be taught? A study of the effects of formal instruction on the acquisition of WH questions by children. *Applied Linguistics*, *5*, 138–155.
Ellis, R. (1994). A theory of instructed second language acquisition. In N. C. Ellis (Ed.), *Implicit and explicit language learning* (pp. 79–114). San Diego, CA: Academic Press.
Ellis, R. (1997). *SLA research and language teaching*. Oxford: Oxford University Press.
Ellis, R. (2001). Investigating form-focused instruction. *Language Learning*, *51*(Suppl. 1), 1–46.
Fanselow, J. (1977). The treatment of error in oral work. *Foreign Language Annals*, *10*, 583–593.
Fazio, L., & Lyster, R. (1998). Immersion and submersion classrooms: A comparison of instructional practices in language arts. *Journal of Multilingual and Multicultural Development*, *19*, 303–317.
Gass, S. (1997). *Input, interaction, and the second language learner*. Mahwah, NJ: Erlbaum.
Grevisse, M., & Goosse, A. (1995). *Nouvelle grammaire française* [New French grammar]. Louvain-La-Neuve, Belgium: Duculot.
Han, Z. (2002). A study of the impact of recasts on tense consistency in L2 output. *TESOL Quarterly, 36*, 542–572.
Harley, B. (1989). Functional grammar in French immersion: A classroom experiment. *Applied Linguistics*, *10*, 331–359.
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* (pp. 156–174). New York: Cambridge University Press.
Harley, B., Cummins, J., Swain, M., & Allen, P. (1990). The nature of language proficiency. In B. Harley, P. Allen, J. Cummins, & M. Swain (Eds.), *The development of second language proficiency* (pp. 7–25). New York: Cambridge University Press.
Hulstijn, J. (1990). A comparison between the information-processing and the analysis/control approaches to language learning. *Applied Linguistics*, *11*, 30–45.
Huot, C. (1991). *Dictionnaire des genres* [Dictionary of genders]. Montreal: Éditions du Méridien.
Johnson, K. (1996). *Language teaching and skill learning*. Oxford: Blackwell.
Klein, W. (1986). *Second language acquisition*. New York: Cambridge University Press.
Larose, C., Le Petitcorps, F., Jutras, L., & Bissonnette, L. (1994). *Mémo mag 5* [Memo mag 5]. Montreal: Graficor.
Leeman, J. (2003). Recasts and second language development: Beyond negative evidence. *Studies in Second Language Acquisition*, *25*, 37–63.
Lin, Y.-H., & Hedgcock, J. (1996). Negative feedback incorporation among high-proficiency and low-proficiency Chinese-speaking learners of Spanish. *Language Learning*, *46*, 567–611.
Logan, G. (1988). Toward an instance theory of automatization. *Psychological Review*, *95*, 492–527.
Long, M. (1977). Teacher feedback on learner error: Mapping cognitions. In H. D. Brown, C. A. Yorio, & R. H. Crymes (Eds.), *On TESOL '77* (pp. 278–293). Washington, DC: TESOL.
Long, M. (1996). The role of the linguistic environment in second language acquisition. In W. C. Ritchie & T. K. Bhatia (Eds.), *Handbook of language acquisition: Vol. 2. Second language acquisition* (pp. 413–468). San Diego, CA: Academic Press.
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.
Long, M., & Robinson, P. (1998). Focus on form: Theory, research, and practice. In C. Doughty & J. Williams (Eds.), *Focus on form in the classroom second language acquisition* (pp. 15–41). New York: Cambridge University Press.
Lyster, R. (1990). The role of analytic language teaching in French immersion programs. *Canadian Modern Language Review*, *47*, 159–176.
Lyster, R. (1994). The effect of functional-analytic teaching on aspects of French immersion students' sociolinguistic competence. *Applied Linguistics*, *15*, 263–287.
Lyster, R. (1998a). Negotiation of form, recasts, and explicit correction in relation to error types and learner repair in immersion classrooms. *Language Learning*, *48*, 183–218.
Lyster, R. (1998b). Recasts, repetition, and ambiguity in L2 classroom discourse. *Studies in Second Language Acquisition*, *20*, 55–85.
Lyster, R. (1999). Immersion. In B. Spolsky (Ed.), *Concise encyclopedia of educational linguistics* (pp. 626–632). New York: Pergamon.
Lyster, R. (2002a). The importance of differentiating negotiation of form and meaning in classroom interaction. In P. Burmeister, T. Piske, & A. Rohde (Eds.), *An integrated view of language devel-*

https://doi.org/10.1017/S0272263104263021 Published online by Cambridge University Press

opment: Papers in honor of Henning Wode (pp. 381–397). Trier, Germany: Wissenschaftlicher Verlag Trier.
Lyster, R. (2002b). Negotiation in immersion teacher-student interaction. *International Journal of Educational Research*, *37*, 237–253.
Lyster, R. (2004, May). *How to predict the grammatical gender of 8,000 French nouns*. Paper presented at the meeting of the American Association for Applied Linguistics, Portland, OR.
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., Gass, S., & McDonough, K. (2000). How do learners perceive implicit negative 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.
Marcus, G. (1993). Negative evidence in language acquisition. *Cognition*, *46*, 53–85.
McHoul, A. (1990). The organization of repair in classroom talk. *Language in Society*, *19*, 349–377.
McLaughlin, B. (1987). *Theories of second-language learning*. London: Arnold.
McLaughlin, B. (1990). Restructuring. *Applied Linguistics*, *11*, 113–128.
Musumeci, D. (1996). Teacher-learner negotiation in content-based instruction: Communication at cross-purposes? *Applied Linguistics*, *17*, 286–325.
Nabei, T., & Swain, M. (2002). Learner awareness of recasts in classroom interaction: A case study of an adult EFL student's second language learning. *Language Awareness*, *11*, 43–63.
Netten, J. (1991). Towards a more language oriented second language classroom. In L. Malavé & G. Duquette (Eds.), *Language, culture, and cognition* (pp. 284–304). Clevedon, UK: Multilingual Matters.
Nicholas, H., Lightbown, P., & Spada, N. (2001). Recasts as feedback to language learners. *Language Learning*, *51*, 719–758.
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: Erlbaum.
O'Malley, M., & Chamot, A. (1990). *Learning strategies in second language acquisition*. New York: Cambridge University Press.
Panova, I., & Lyster, R. (2002). Patterns of feedback and uptake in an adult ESL classroom. *TESOL Quarterly*, *36*, 573–595.
Pica, T. (1988). Interlanguage adjustments as an outcome of NS-NNS negotiated interaction. *Language Learning*, *38*, 45–73.
Pica, T. (1994). Research on negotiation: What does it reveal about second-language learning conditions, processes, and outcomes? *Language Learning*, *44*, 493–527.
Pica, T. (2002). Subject-matter content: How does it assist the interactional needs of classroom language learners? *Modern Language Journal*, *86*, 1–19.
Ranta, L., & Lyster, R. (2003). *A cognitive approach to improving immersion students' oral language abilities: The Awareness-Practice-Feedback sequence*. Unpublished manuscript submitted for publication.
Rebuffot, J. (1993). *Le point sur l'immersion au Canada* [Update on immersion in Canada]. Montreal: Éditions CEC.
Robinson, P., & Ha, M. (1993). Instance theory and second language rule learning under explicit conditions. *Studies in Second Language Acquisition*, *13*, 413–438.
Saxton, M. (1997). The Contrast Theory of negative input. *Journal of Child Language*, *24*, 139–161.
Schmidt, R. (1990). The role of consciousness in second language learning. *Applied Linguistics*, *11*, 129–158.
Schmidt, R. (1992). Psychological mechanisms underlying second language fluency. *Studies in Second Language Acquisition*, *14*, 357–385.
Schmidt, R. (2001). Attention. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 3–32). New York: Cambridge University Press.
Schmidt, R., & Frota, S. (1986). Developing basic conversational ability in a second language: A case study of an adult learner of Portuguese. In R. Day (Ed.), *Talking to learn* (pp. 237–326). Rowley, MA: Newbury House.
Shiffrin, R., & Schneider, W. (1977). Controlled and automatic human information processing: Part 2. 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.

https://doi.org/10.1017/S0272263104263021 Published online by Cambridge University Press

Spada, N. (1997). Form-focused instruction and second language acquisition: A review of classroom and laboratory research. *Language Teaching*, *30*, 73–87.
Spada, N., & Lightbown, P. (1993). Instruction and the development of questions in L2 classrooms. *Studies in Second Language Acquisition*, *15*, 205–224.
Surridge, M. (1986). Genre grammatical et dérivation lexicale en français [Grammatical gender and lexical derivation in French]. *Canadian Journal of Linguistics/Revue canadienne de linguistique, 31*, 267–283.
Surridge, M. (1995). *Le ou la? The gender of French nouns*. Clevedon, UK: Multilingual Matters.
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.
Swain, M. (1997). French immersion programs in Canada. In J. Cummins & D. Corson (Eds.), *Encyclopedia of language and education: Vol. 5. Bilingual education* (pp. 26–269). Dordrecht: Kluwer.
Swain, M. & Lapkin, S. (1995). Problems in output and the cognitive processes they generate: A step towards second language learning. *Applied Linguistics*, *16*, 370–391.
Towell, R., & Hawkins, R. (1994). *Approaches to second language acquisition*. Clevedon, UK: Multilingual Matters.
Tucker, R. (1967). *The French speaker's skill with grammatical gender: An example of rule-governed behaviour*. Unpublished doctoral dissertation, McGill University, Montreal.
Tucker, R., Rigault, A., Lambert, W., & Segalowitz, N. (1968). A psychological investigation of French speakers' skill with grammatical gender. *Journal of Verbal Learning and Verbal Behavior*, *7*, 312–316.

https://doi.org/10.1017/S0272263104263021 Published online by Cambridge University Press