

Focus on Form Through Interaction Enhancement: Integrating Formal Instruction Into a Communicative Task in EFL Classrooms

Hitoshi Muranoi
Tohoku Gakuin University

This study examines the impact of interaction enhancement (IE) on the learning of English articles. IE is a treatment that guides learners to focus on form by providing interactional modifications and leads learners to produce modified output within a problem-solving task (strategic interaction). Two different IE treatments were employed: IE plus formal debriefing (IEF), and IE plus meaning-focused debriefing (IEM). Outcomes of these treatments were compared with the effects of non-enhanced interaction in a quasi-experimental study involving 91 Japanese EFL learners. Progress was measured with a pretest and two post-tests, yielding these major findings: (1) IE had positive effects on the learning of English articles;

Hitoshi Muranoi, Department of English.

The findings reported here are based on my doctoral dissertation for the PhD in applied linguistics, submitted to Georgetown University in December 1996. I am very grateful to Cathy Doughty for guiding me in this research. I also thank Andrea Tyler, Ronald Leow, and Ken Schmidt for their insightful suggestions and helpful advice, and the editor and anonymous reviewers of *Language Learning* for their valuable comments. All errors are, of course, my own. I am also deeply grateful to the students who participated in this study.

Parts of the preliminary analyses of the data were reported at the Seventh International University of Japan Conference on Second Language Research in Japan (Tokyo, January 20, 1996; see Fujimura, Kato, Ahmed, & Leong, 1996, pp. 42–58), and the American Association for Applied Linguistics 1996 Annual Conference (Chicago, March 26, 1996).

Correspondence concerning this article may be sent to Hitoshi Muranoi, Department of English, Tohoku Gakuin University, 1-3-1 Tsuchitai, Aoba, Sendai 980-8511, Japan. Internet: muranoih@tssc.tohoku-gakuin.ac.jp

(2) the IEF treatment had a greater impact than the IEM treatment.

Careful examination of the effectiveness of purely meaning-focused communicative language teaching has led a number of second language (L2) researchers to claim that communicative instruction should involve systematic treatments to draw L2 learners' attention to linguistic forms to develop well-balanced communicative competence (Doughty & Williams, 1998; Lightbown & Spada, 1990; Long & Robinson, 1998; Loschky & Bley-Vroman, 1993; Skehan, 1996; Swain, 1985).¹ Though a number of empirical studies have been conducted to identify the conditions in which formal instruction facilitates L2 acquisition, there remain a vast number of issues to be investigated. Most importantly, we need to know more about precisely how and why formal instruction incorporated into communicative language teaching promotes interlanguage development. To obtain significant data for this issue, this study examined how a communicative instructional technique in which interaction was enhanced by means of teacher feedback affected L2 learning.

Long (1991) conceptualized the need to incorporate form-focused instruction into meaning-oriented communicative language teaching with the term "focus on form." Focus on form, as Long defined it, is a type of instruction in which the primary focus is on meaning and communication, with the learner's attention being drawn to linguistic elements only as they arise incidentally in lessons. This is in sharp contrast with traditional grammar instruction, or "focus-on-forms" instruction, which places a focus on forms themselves in isolation (Long, 1991, pp. 45–46). The significance of focus-on-form instruction has been recognized by L2 teachers and researchers, and a number of empirical studies aimed at determining the effect of focus on form have been conducted. Doughty (1991), for example, examined the effect of computer-based instruction on the learning of relativization by university ESL students, comparing three treatments: meaning-

oriented treatment, rule-oriented treatment, and no treatment. She reported that both types of instructional treatment had positive effects and that meaning-oriented instruction, in which saliency and redundancy of the target form were enhanced, was not detrimental to the formal learning of relativization. This has significant pedagogical implications in that it suggests that teachers can direct learners' attention to language forms effectively within meaning-oriented instruction if saliency and redundancy of forms are controlled sufficiently.

Lightbown and Spada (1990), observing communicative ESL courses in Quebec, also reported positive effects of focus on form. They found that a class in which form-focused instruction was provided within a communicative language teaching framework contributed to high levels of linguistic knowledge and improved command of progressive-*ing* and adjective-noun order in noun phrases. Spada and Lightbown (1993) then conducted a quasi-experimental study on the effects of form-focused instruction and corrective feedback on the development of interrogative constructions in the oral performance of ESL learners in Quebec. They provided form-focused activities and exercises (e.g., unscrambling tasks, guessing games, preference tasks) and corrective feedback over a two-week period (approximately nine hours) within the context of an intensive ESL program. Results support their earlier conclusion (Lightbown & Spada, 1990) that form-focused instruction within communicative language teaching is beneficial in L2 acquisition by ESL learners.

A number of researchers have attempted to implement form-focused instruction in communicative language teaching by providing interactional modifications such as recasts and requests for repetition in L2 classrooms. Based on positive reports on the impact of recasts on first and second language development (e.g., Doughty, 1994; Farrar, 1992), Doughty and Varela (1998) conducted an empirical study examining the effects of recasting on L2 learning in the context of a content-based ESL science class in the United States. In their study the instructor provided learners with corrective recasts whenever past or conditional errors

occurred in speaking and writing. Results indicate that learners who were given recasts showed greater improvements in accuracy and a higher total number of attempts at past-time reference than the control group. Long, Inagaki, and Ortega (1998) reported the results of two experiments conducted to assess the relative utility of models and recasts in L2 Japanese and Spanish. Each study provided some evidence of the ability of adult L2 learners to learn from implicit negative feedback, and, in their Spanish L2 study, they found evidence that reactive implicit negative feedback (recasts) were more effective than preemptive positive input (models) in achieving at least short-term improvements on a previously unknown L2 structure. Mackey and Philp (1998) also investigated the effect of recasts on ESL learners' interlanguage development of question forms by comparing groups of ESL learners who received interactionally modified input with learners who received the same input containing recasts. Their results suggest that for more advanced learners, interaction with recasts may be more beneficial than interaction alone in facilitating an increase in production of targeted morphosyntactic forms. Mackey and Philp also reported that these positive developmental effects were found for recasts even though recasts were usually not repeated and rarely elicited modification by the learners. That is, their results suggest that recasts may be beneficial even when they are not incorporated in learners' immediate responses. It should be noted that in these studies, which reported positive effects of recasts, the instructional treatments were targeted on specific grammatical items such as past-tense forms and question formation. However, in studies that examined the effects of recasts without any particular target, weaker impact has been reported. For instance, observing L2 learners of French in communicative classrooms, Lyster and Ranta (1997) found that recasts, the type of feedback used most frequently by teachers, were less effective in eliciting student-generated repair than other types of feedback: clarification requests, repetition, metalinguistic feedback, and elicitation. They reported that only 18% of teacher recasts were immediately followed by student repair. Lyster and Ranta's results, along with

the results of other recast studies focusing on particular forms (e.g., Doughty & Varela, 1998; Mackey & Philp, 1998), suggest that recasts might have weaker effects on L2 learning when they do not focus on specific forms. Lyster (1998) further examined the role of recasts in French immersion classrooms at the primary level. He obtained data suggesting that teachers did not consistently use recasts for corrective purposes; that is, they sometimes used recasts in reaction to students' accurate use of the L2 as well as for corrective purposes. He claimed, therefore, that "it remains unlikely . . . that students perceive any corrective purpose in such recasts" (1998, p. 75). Lyster (1998) argued that learners were pushed to modify their output, not by recasts, but by signals of noncomprehension, provided in forms such as clarification requests.

Applying Vygotskian sociocultural theory to L2 research, Aljaafreh and Lantolf (1994) conducted a study that investigated the effect of negative feedback on L2 learning by ESL adult learners. They reported that negative feedback provided during dialogic negotiation by a tutor who constantly tried to discover the learner's zone of proximal development (the distance between a learner's actual developmental level and the level of potential development) was vital for L2 learning. They claimed that it was important for a tutor to continuously assess the learner's needs and provide appropriate help. Aljaafreh and Lantolf made the strong claim that "this process can be accomplished only through the collaborative interaction of the expert and the novice" (1994, p. 468).

In an empirical study of question formation in ESL, Mackey (1999) found that interactional modification (negotiation with a recast, or negotiation without a recast) led to L2 development and that more active involvement in negotiated interaction led to greater development. Mackey also reported that learners who interacted with a native speaker during goal-based tasks increased in both their production of question forms at higher levels and their developmental stages and that watching interaction or taking part in interaction without negotiation had some limited

effects. Mackey claimed that the results of her study supported Long's (1996) updated version of the interaction hypothesis that implicit negative feedback, which can be obtained through negotiated interaction, facilitates L2 development.

The impact of interactional modifications on L2 acquisition has also been investigated in terms of learner output. Following Swain's (1985, 1995) output hypothesis, which claims that pushing learners to produce output is essential for L2 acquisition, Pica, Holliday, Lewis, and Morgenthaler (1989) examined the effects of output production. Analyzing interactions between adult ESL learners and a native speaker of English, they found that learners' production of comprehensible output was influenced by the linguistic demands placed on them by the native speakers' signals of comprehension difficulty (e.g., a request for clarification or a model to repeat or acknowledge) and by the nature of the communication tasks. Positive effects of a treatment pushing learners to produce output within a communicative task were reported by Takashima (1995). He provided Japanese learners of English as a foreign language (EFL) with form-focused feedback aimed at leading the learners to modify their output. Takashima claimed that the results of his experiment supported the facilitative effects of his output-oriented treatment on the learning of the target grammar (i.e., past-tense forms). On the basis of think-aloud data taken from immersion students engaging in a writing task, Swain and Lapkin (1995) argued that in L2 production, learners noticed a linguistic problem through either internal or external feedback, and that noticing triggered mental processes (e.g., the generation and assessment of alternatives) that led to modified output. Swain and Lapkin claimed that what went on mentally between the original output and its reprocessed form was part of the process of L2 learning. Kowal and Swain (1994) and Swain (1998) reported data suggesting that immersion students processed L2 syntactically in dictogloss tasks in which they worked in pairs or small groups to reconstruct a text read aloud by a teacher. Kowal and Swain (1994) concluded that collaborative language production tasks promoted L2 learning by (1) making learners aware of

gaps in their existing knowledge, which they would subsequently seek to fill; (2) raising their awareness of the links among form, function, and meaning; and (3) providing them with opportunities to obtain feedback.

A positive role of output was also reported in Ellis and He's (1999) study comparing the effect of a negotiated output treatment, a premodified input treatment, and an interactionally modified input treatment on ESL university students' acquisition of new L2 words. Results revealed that the negotiated output treatment, which provided the learners with opportunities to produce and modify their output, was more effective than the other two treatments.

These studies provide strong support for the assumption that a timely combination of formal instruction and communication-oriented instruction is highly beneficial to L2 learners. What is now urgently needed in focus-on-form research is the identification of more diverse ways of making this timely combination possible. L2 researchers particularly need to examine the differential effects of the various types of focus-on-form instruction available to L2 teachers and the effects of combinations of different treatments. Most importantly, the choice of options in focus-on-form instruction must be made based on psycholinguistically motivated predictions, because previous effect-of-instruction studies have suggested that instruction taking psycholinguistic and cognitive factors into account has beneficial effects on L2 acquisition (DeKeyser, 1998; Doughty & Williams, 1998; Ellis, 1997; Long, 1991; Long & Robinson, 1998; Skehan, 1998; Sorace, 1985).

Another move that has guided the further integration of form-focused instruction into communication-oriented instruction comes out of recent critical evaluation of communicative tasks. Though communicative tasks (e.g., information-gap tasks and problem-solving tasks) have been used widely in L2 classroom settings because they are useful in creating opportunities to use the target language, some pitfalls have been recently pointed out. For example, Loschky and Bley-Vroman (1993) argued that "in most common information gap tasks, learners seem to be able to

exchange information solely through use of semantic- and pragmatic-based strategies combined with their background knowledge. Such tasks, then, may do more to develop strategic than linguistic competency” (pp. 125–126). This weakness of communicative tasks has led L2 researchers to recognize a need for incorporating form-focused treatments into instruction, that is, a need for “devising methods of focusing on form without losing the values of communication tasks as realistic communicative motivators, and as opportunities to trigger acquisitional processes” (Skehan, 1996, p. 42).

The Present Study

Based on the results of previous focus-on-form and feedback studies, this study employed a communicative instructional technique termed interaction enhancement (IE), in which interaction was enhanced by means of feedback provided by a classroom teacher. More specifically, this technique aimed at leading learners to develop their interlanguage systems by providing enhanced interactional modifications (e.g., requests for repetition and recasts) that responded to the well-formedness of target forms during problem-solving tasks. The effect of IE was tested on the acquisition of the English article systems by university EFL students in this study (see the Method section for details).

The general research question addressed in this study was the following:

Research Question 1: Does IE, in which a teacher provides implicit negative feedback during an interactive problem-solving task, affect EFL learners’ restructuring of their interlanguage article systems?

To answer Research Question 1, this study examined whether each experimental group (receiving the IE treatments) surpassed the contrast group (participating in unenhanced interactions) in performance with English articles, as measured by immediate and delayed post-tests.

The following four research questions also motivated the study:

Research Question 2: Do two types of IE, differing in the manner of focus on form (implicit negative feedback during an interactive problem-solving task plus formal debriefing versus implicit negative feedback during an interactive problem-solving task plus meaning-focused debriefing), have different effects on EFL learners' restructuring of their interlanguage article systems?

This question concerns what type of form-focused treatment is necessary and sufficient for restructuring in L2 to be triggered. Though related research has indicated that both explicit and implicit types of feedback lead to L2 learning (e.g., Carroll & Swain, 1993), it is not known whether providing both types of feedback promotes L2 learning more effectively than providing only one type. This study examined whether explicit form-focused instruction (i.e., explicit grammar explanation) was necessary to help EFL learners who had received implicit form-focused instruction (i.e., implicit negative feedback to overgeneralization errors) restructure their interlanguage systems by comparing the impacts of two types of form-focused instruction: (1) instruction that provided both implicit negative feedback during a problem-solving task and later explicit grammar explanation (IE plus formal debriefing: IEF) and (2) instruction that provided implicit negative feedback during a problem-solving task without any explicit explanation of grammar (IE plus meaning-focused debriefing: IEM). To address Research Question 2, this study investigated whether the group receiving IE plus formal debriefing (IEF) surpassed the group receiving IE plus meaning-focused debriefing (IEM) in performance with English articles as measured by immediate and delayed post-tests.

Research Question 3: If IE, in which a teacher provides implicit negative feedback during an interactive problem-solving task, has an effect on EFL learners' restructuring of their interlanguage article systems, will the effect hold over the five-week post-test period?

This question concerns the longer-term effect of IE. L2 researchers hoping to claim a positive effect for a particular type of instruction must investigate whether the effect is lasting or “peripheral and fragile” (Krashen, 1992; Spada & Lightbown, 1993). This study aimed at identifying what types of instructional treatment exerted long-lasting influence on EFL learners’ restructuring of their interlanguage article systems. Differences between immediate and delayed post-test scores for the experimental groups receiving the IE treatments (IEF and IEM) were examined to address Research Question 3.

Research Question 4: Does IE have differential effects on learner performance with English articles depending on the elicitation task used?

In this study, the impact of each treatment was examined with a test consisting of four different elicitation tasks (oral story description, oral picture description, written picture description, and grammaticality judgment tasks). Previous studies on L2 learning of articles have revealed that learners’ accuracy in article use greatly depends on the type of elicitation tasks. Tarone and Parrish (1988), for example, reported that their subjects produced English articles least accurately on the measure assumed to require attention to form (a written grammatical judgment task) and far more accurately on measures assumed to require less attention to form (oral interview and oral narrative tasks). They explained that the cohesiveness of the discourse in oral tasks favored increased accuracy as opposed to the grammar task. Considering these findings, this study used four tasks differing in discourse type (story description, picture description, and sentence-level grammaticality judgment) and mode (oral/written). To answer Research Question 4, this study examined whether each experimental group (receiving the IE treatments) performed differently with English articles on four different elicitation tasks.

Research Question 5: Does the difference in mode of participation (performers versus observers) in IE make a difference in learner performance with English articles?

There are two modes of learner participation in IE, which was devised as an instructional technique to be used in large classes: performers of interactive tasks and observers. The latter did not perform roles in interactive tasks but worked as consultants when their representatives (i.e., performers) needed assistance. Whether IE positively affected not only learners who directly participated in IE, but also learners who observed it, was a great concern in this study. In a study examining the relationship between negotiation and comprehension, Pica (1992) found that observers who merely saw and listened to other learners negotiate for meaning with an interlocutor (= teacher) in an information-gap task comprehended the negotiated input as accurately as the negotiators themselves. However, as already discussed, Mackey (1999) found that watching interaction had limited effects on L2 development of question formation. This study aimed at obtaining more data on this issue by investigating whether mere observation of negotiation helped learners produce L2 more accurately. To address Research Question 5, this study examined whether there were significant differences in performance with English articles between performers and observers in the experimental groups receiving the IE.

Method

In this quasi-experimental study, the effects of two types of IE and the contrast treatment (non-enhanced interaction: NEI) were compared quantitatively. The participants were first-year Japanese college students enrolled in three intact EFL classes at a university in Japan. The students were randomly assigned to the classes by the university registration office. A total of 114 students participated in this study and the data of 91 participants were analyzed.² Two types of the IE treatment were provided to two separate experimental groups: Experimental Group 1 (IEF, $n = 31$) and Experimental Group 2 (IEM, $n = 30$). The contrast treatment was given to a contrast group (NEI, $n = 30$). A one-way analysis of variance (ANOVA) was run to determine if there were

any statistically significant differences among the three groups' mean scores on the pretest measuring ability to use English articles. No significant difference among the groups was revealed ($F(2/88) = 0.903, ns$). Because the selection of the participants was made by using three intact classes rather than by random sampling, caution in interpreting the findings is needed; the three classes were initially equal in terms of their performance with English articles, but they might have differed in other linguistic and affective factors.

The instructional treatments were provided during three training sessions, each of which lasted approximately 30 minutes. All the three treatments were administered during weekly 90-minute EFL classes, which I taught. One pretest and two post-tests consisting of oral/written production tasks and a grammaticality judgment task were given in order to examine both short- and long-term effects of the instruction, with an interval of five weeks between the two post-tests.³ As the present researcher served as the instructor in training sessions, this study may have had such problems as "researcher expectancy" (Beck & Eubank, 1991; Brown, 1988). Though I was very careful during training sessions to avoid influencing learner performance along the lines of my predictions for the experiment, it is not possible for me to claim that the expectancy effect had absolutely no impact on the data, because no objective evaluation of the effect was made. The data, therefore, must be interpreted with consideration of the possible influence of researcher expectancy.

Instructional Focus

The effect of IE was tested on the acquisition of the English article system. The article system was chosen because many ESL and EFL learners, even advanced learners, use English articles inaccurately (e.g., overuse articles in nonobligatory contexts). This learning difficulty arises because article use depends on a variety of linguistic and pragmatic factors, such as specificity and presupposition of referents and interlocutor relationships (see Pica, 1985;

Master, 1990). Though article errors rarely cause miscomprehension, because articles have low communicative value, it is still important for ESL learners to overcome their problems with articles. Master (1995), for example, points out that attention to the article system is important because article errors, along with errors with prepositions and subject–verb agreement, may leave the impression that the user (= the learner) has inadequate control of the language, especially in academic writing.

Throughout the study the instructional focus was on article use in indefinite contexts, specifically the function of the indefinite article (a/an) to denote that the referent of the accompanying noun carries new information. No direct instruction was given on the use of the indefinite plural (zero article), nor on definite article use.⁴ This study, therefore, investigated whether instruction on indefinite article use affected performance not only on the targeted indefinite article but also on the untargeted definite article, which has a binary relation with the indefinite article.

The students had received prior instruction on English articles at junior and senior high schools. However, their performance on English articles, as indicated by pretest data, was poor; accuracy rates for use of the indefinite article and the definite article in the pretest were 30.3% and 40.7%, respectively (see Table 2 for details).

Pretesting Measures

The participants were pretested on their command of the English article system with the English Article Diagnostic Test (EADT), which was developed for this study. The EADT was designed to measure learner performance with English articles in a valid manner. As learner performance with articles varies greatly depending on the elicitation task used, it should be measured with different elicitation tasks. The EADT consisted of four different tasks, including both oral and written modes: an oral story description (OS) task, an oral picture description (OP) task, a written picture description (WP) task, and a grammaticality

judgment (GJ) task (Appendix B). Data collection was conducted in audiovisual rooms where all students could individually record their performances simultaneously. In the OS task, the students were shown two short (approximately 90 seconds) silent scenes taken from two American movies and asked to describe what was going on in the scenes. The scenes were played twice silently so that the students had to depend solely on visual cues to describe the scenes. The students were given 30 seconds to rehearse what to say before the recording. They then had 90 seconds to describe each scene (speaking into a microphone attached to their headsets). The students were encouraged to try to depict the scenes in as much detail as possible.

In the OP task the students were given a four-page test packet with one situation (four pictures) on each page and asked to describe the situations orally. The pictures provided the students with contexts in which they had to use appropriate articles to refer to new, current, and known referents (see Appendix A). The students had 30 seconds to prepare for the oral description of each situation, using several content words assigned next to each picture. Description of the first situation was not included in the statistical analyses, because it was used as a warm-up practice. All the descriptions elicited in these oral tasks were tape-recorded and analyzed.

The WP task was used to measure the students' ability to produce English articles in the written mode. The students were presented with two sets of four pictures depicting several people and animals. The WP task had the same format as that of the OP task, but the pictures in the WP task were not the same as those used in the OP task. Students were directed to describe the situations by writing down sentences on an answer sheet.

The students were also asked to judge the grammaticality of 16 sentences that contained 36 NPs (see Appendix B). Article use was ungrammatical in 16 of the NPs. Half of the incorrect items required an indefinite article, and the other half required a definite article. The task also included distracters with ungrammatical verb forms. The students were directed to decide whether the

underlined parts of each sentence were grammatical or ungrammatical. They were also asked to correct all incorrect, ungrammatical parts.

Instructional Treatments

IE was developed as an attempt to incorporate form-focused instruction into communicative language teaching by borrowing from Di Pietro's (1987) use of scenarios in strategic interaction. The strategic interaction approach uses scenarios to create contexts in which learners are led to use the target language naturally. A scenario, by Di Pietro's (1987) definition, is "a strategic interplay of roles functioning to fulfill personal agendas within a shared context" (p. 41). Di Pietro also wrote, "By fostering an interactive tension through scenarios, the teacher leads the students to discover aspects of the target language needed to resolve the tension" (Di Pietro, 1987, p. 15). In other words, scenarios are the frames that lead learners to engage in realistic discourse in the target language. The underlying psychological basis of strategic interaction is Vygotsky's philosophy, which claims that "individuals develop thinking processes through dialog with other individuals" (from Di Pietro, 1987, p. 4). This emphasis on the dynamics of human interactions also underlies IE.

The scenarios used in this study have two goals. One is to guide L2 learners to use the target language in realistic situations. The other is to present learners obligatory contexts for a particular linguistic form. That is, the scenarios were used in this study to both lead learners to interact with each other in the L2 and guide them to accurately produce a particular target form. These contexts set by scenarios help teachers predict when learners will produce the form and make errors with the form. Teachers, therefore, can provide well-prepared and organized feedback. This proactive feature contrasts with reactive focus-on-form instruction, which "places considerable requirements on the teacher's 'on-line' capacity to notice and assess the need for the intervention, and instantly to devise consistent FonF (focus on form) interven-

tions for learner errors, even as he or she must attend to other pedagogical problems” (Doughty & Williams, 1998, p. 206).

For this study, I devised three sets of scenarios. Each scenario had Roles A and B. Both roles were set to create contexts in which learners were obliged to use the target form (the indefinite article) to solve a problem provided by the scenario.

Another feature of Di Pietro’s strategic interaction incorporated in IE is the three-phase structure of instruction (i.e., a rehearsal phase, a performance phase, and a debriefing phase). This three-phase structure helps L2 teachers organize interactive tasks effectively and incorporate focus-on-form treatments into communicative tasks in an organized manner. Each strategic interaction session consisted of three phases: (1) a rehearsal phase, (2) a performance phase, and (3) a debriefing phase.

The rehearsal phase. The first phase, a rehearsal phase, was identical for the experimental groups and the contrast group. First, the instructor (= I) gave each learner a sheet (Role A or Role B; see Appendix C) describing a scenario to be performed. Students were directed to form pairs with their neighbors. Learners then worked in pairs for approximately 10 minutes, preparing to perform the assigned role. Only one role was presented to each pair. The instructor emphasized that this was practice for the use of the target language in a realistic situation, thus hoping to focus the learners’ central concern on communication. Target functions of each scenario were emphasized by means of a written list of useful expressions. The list contained expressions that were useful in fulfilling such functions as “requesting,” “asking questions,” “complaining,” “apologizing,” and so on. During the rehearsal phase, the instructor walked around the classroom to help each pair prepare for the performance.

The performance phase. After the rehearsal phase, the scenarios were performed in class. Roles A and B were all performed by Teacher–Student (T–S) pairs. For each T–S interaction, the instructor randomly nominated one student representative, who was then asked to play a role in interaction with the instructor (e.g., Role A for the instructor and Role B for a student representative). In each

class, three separate T-S interactions were performed with three different representatives in the first training session. During the second session, four T-S interactions were conducted with four student representatives in each class. In the third session, three T-S interactions were performed. In each class, a total of 10 different representatives participated in 10 separate strategic interactions over the course of three training sessions. Every performance was videorecorded and transcribed for qualitative analyses (see *Analyses of Interaction*).

As suggested by Di Pietro (1987), the representatives in strategic interaction were encouraged to consult with their partners when they had difficulty in performing their roles. Also, the instructor tried to maintain interaction by providing interactional modifications such as confirmation checks and clarification requests when communication breakdowns occurred.

For Experimental Groups 1 and 2, the instructor enhanced interaction in order to guide the learners to produce output and modify it when it was ill formed. This procedure distinguishes the present instruction from Di Pietro's strategic interaction; the accuracy of communicatively redundant forms is not given priority for correction in Di Pietro's strategic interaction.

The instructor provided the students with form-focused feedback in the forms of request for repetition and recast in response to all errors with the indefinite article in obligatory contexts and some errors with tense-aspect forms. Other grammatical errors were ignored. Previous studies reporting positive effects of feedback (e.g., Doughty & Varela, 1998; Mackey & Philp, 1998; Takashima, 1995) also employed specific instructional targets. The students, however, were not directly informed of what the linguistic focus of the instruction was.

In this study, "request for repetition" is defined as "a question or a statement with rising intonation that asks for further clarification of the learner utterance" (Doughty, 1994, p. 102).⁵ "Recast" is operationalized as "the teacher's reformulation of all or part of student's utterance, minus the error" (Lyster & Ranta, 1997, p. 46). Such feedback, referred to as "negative input enhancement" by

Sharwood Smith (1993), flags given forms as incorrect, thus signaling to the learner that they have violated the target norms (Sharwood Smith, 1993, p.177). For example, a request for repetition given to a learner in response to his or her incorrect output is designed to help the learner notice a mismatch between his or her interlanguage rule and the target rule. Also a recast is given in this focus-on-form treatment when a request for repetition does not lead the learner to notice the mismatch (i.e., to modify incorrect output). Providing these interactional modifications is here termed “input enhancement,” because learners rarely receive such types of negative feedback in naturalistic interactions. The same requests for repetition are also designed to work as “output enhancement” aimed at guiding the learner to modify their interlanguage rule and produce modified output. Along with these feedback treatments, obligatory contexts for a target form provided by scenarios also push learners to produce output. “Input enhancement” (Gass, 1997; Sharwood Smith, 1993; White, Spada, Lightbown, & Ranta, 1991) and “output enhancement” (Swain, 1985, 1995; Takashima, 1995) are thus interwoven in a systematic manner in the IE treatments to lead L2 learners to focus on form.

The following example illustrates how interaction is actually enhanced in modified strategic interaction:

(Student looks at a drawing depicting a big rat running around in a room)

Teacher: And any other problem?	<output enhancement>
Student : . . . I saw rat.	<incorrect output>
Teacher: You saw what?	<request for repetition (input/output enhancement)>
Student: A rat.	<successful modification>
Teacher: Uh-huh, you saw a rat in your room.	<repetition (input enhancement)>
That's terrible.	<topic continuation>

In this example, the student is guided to produce the target form, the indefinite article, by a scenario which sets an obligatory context for the form (i.e., output enhancement). As the student's output is incorrect in this example, the teacher provides a request

for repetition which may lead to the student's correct modification of the interlanguage grammar with the indefinite article. As already discussed, the request for repetition has a dual function here; it works as enhanced input to attach a "flag" to an incorrect use of the NP (input enhancement) and as a facilitator that guides the student to produce modified output (output enhancement). After hearing the student's modified output, the teacher repeats the correct form (input enhancement).

When the learner does not correctly modify output after receiving requests for repetition twice, the instructor provides a corrective recast, which presents the correct grammatical form using as much of the student's sentence as possible. The teacher then continues the communication. This IE procedure is illustrated in Figure 1.

The theoretical underpinnings of IE are cognitive models of L2 acquisition (e.g., DeKeyser, 1998; R. Ellis, 1997; Gass, 1988; McLaughlin, 1987, 1990; Schmidt, 1990; Skehan, 1998). It is the prediction of this study that teacher feedback provided in IE guides the learner to notice the mismatch between his or her current interlanguage system and the target system (e.g., Gass, 1997; Pica, 1992; Schmidt, 1990; Swain, 1985, 1995). Noticing the mismatch promotes "cognitive comparison" of the interlanguage system with the target system (R. Ellis, 1997; Nelson, 1987; Tomasello & Herron, 1989). This cognitive comparison should lead the learner to modify his or her current interlanguage system and produce modified output to test the hypothesis (e.g., Gass, 1988, 1997; Pica, 1992; Swain, 1985, 1995; Swain & Lapkin, 1995). The drive for hypothesis modification is assumed to be the learner's intention to continue the ongoing communication. When the learner produces modified output based on the modified hypothesis, the teacher evaluates the output by providing feedback (i.e., teacher repetitions), so that the learner can confirm that his or her hypothesis is valid. Alternatively, when the initial teacher feedback does not lead the learner to notice an error nor modify his or her output, then the teacher gives another request for repetition. If this does not succeed in eliciting modified output, the teacher

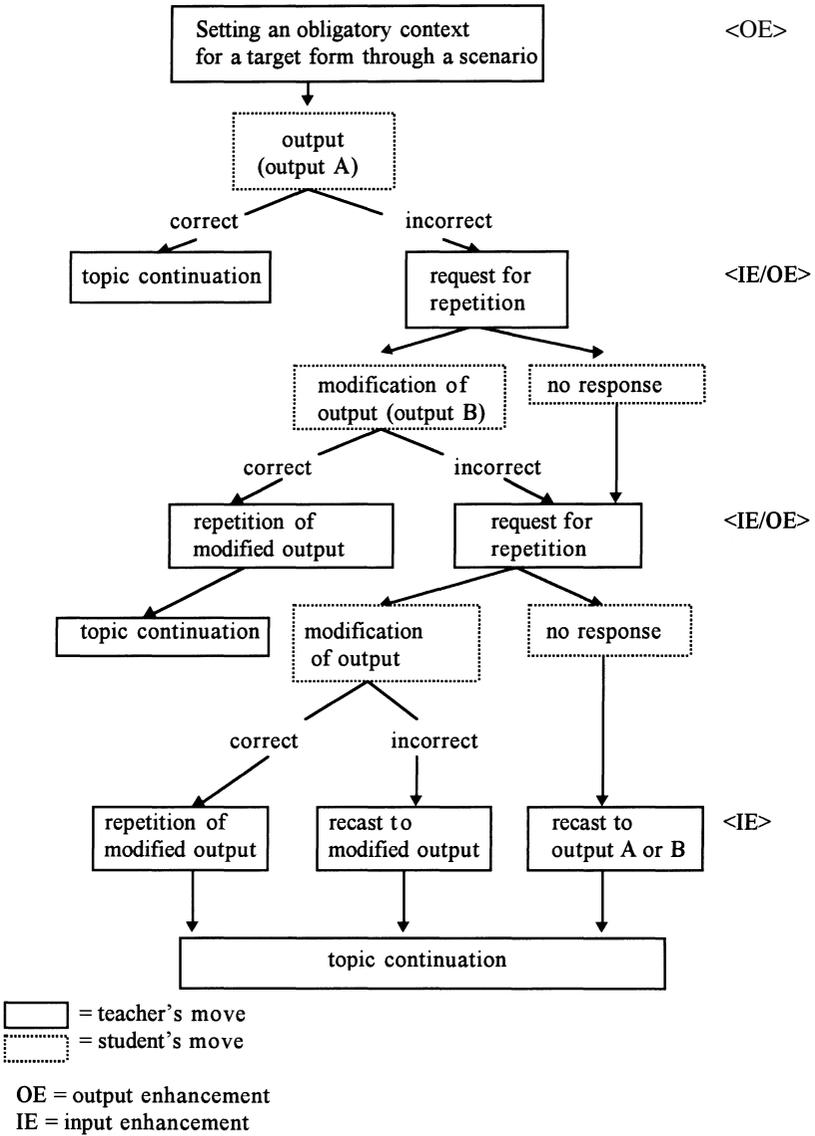


Figure 1. Flow of interaction enhancement during the performance phase

gives a recast and continues the conversation. With this IE procedure, the teacher can promote activation of such cognitive processes as noticing, cognitive comparison, and hypothesis modification without seriously disturbing the flow of communication.

This study also predicted that the activation of these cognitive processes would lead to restructuring of interlanguage systems. Restructuring is a transitional change that occurs when learners modify or reorganize their cognitive internal representations (Cheng, 1985; Karmiloff-Smith, 1986, 1992; McLaughlin, 1987, 1990).⁶ According to cognitive models of restructuring, overgeneralization occurs when learners regulate linguistic rules through only top-down (rule-driven) processing (Karmiloff-Smith, 1986, 1992; McLaughlin, 1987). IE was, therefore, devised as an instructional technique that guided learners to modify rules making use of both top-down (rule-driven) and bottom-up (data-driven) processing, through activating cognitive processes involved in the restructuring of interlanguage systems. In other words, IE aimed at leading learners to (1) notice that their internal rule-driven systems did not match the external linguistic data through input enhancement, and (2) adjust the mismatch through output enhancement.

Though previous effect-of-instruction studies have paid great attention to individual conditions that promote such cognitive processes as noticing the mismatch, cognitive comparison, hypothesis testing and modification, we do not yet know how to facilitate these processes in real L2 and foreign language classrooms in a feasible and effective manner. IE is a candidate for such an optimal combination.

In summary, during IE, the instructor enhanced both input and output by (1) setting obligatory contexts for the target form (output enhancement), (2) providing implicit negative feedback (input enhancement), (3) guiding learners to modify their output through feedback (output enhancement), and (4) evaluating learners' modified output with feedback (input enhancement). It was predicted that these treatments would activate cognitive pro-

cesses involved in restructuring of L2 learners' interlanguage systems, such as noticing, comparison, hypothesis modification, and hypothesis testing.

In this study, the students in the contrast group also participated in strategic interaction. The interaction for this group, however, was not enhanced. That is, the interactional modifications were provided by "normally" responding to the comprehensibility of the students' utterances. That is, the teacher provided interactional modifications to negotiate for meaning only when he did not understand what the student performer meant. No corrective recasts in response to errors with articles were given to the contrast group.

The debriefing phase. After performing the scenarios, the instructor reviewed student performance in class. This was termed the "debriefing phase" by Di Pietro (1987). In this study, debriefing was administered in the learners' first language (Japanese) and audiorecorded for qualitative analysis. Experimental Group 1 (the IEF group) received formal debriefing, whereas Experimental Group 2 (the IEM group) and the contrast group were provided with meaning-focused debriefing. The formal debriefing that Experimental Group 1 received was given based on accuracy of target form use. During the formal debriefing in the first and second sessions of IE, the teacher first explained why Japanese learners of English tend to make errors with articles by focusing on two issues: (1) lack of an article system in Japanese; (2) low saliency of articles. The teacher also pointed out the important role that the indefinite article plays in communication, emphasizing two points: (1) the indefinite article "a" indicates that the referent of the accompanying noun carries information new to the hearer; (2) the function of the indefinite article is almost equivalent to one of the functions of the Japanese post-position "ga" in that both forms mark new information. During the formal debriefing in the third session of IE, the teacher explained the function of the indefinite article in classifying nouns based on Master (1990). That is, the students in Experimental Group 1 were told to add "a" to X if they want to mention any Y which belongs to the group of X (i.e.,

classification). The function of the definite article, identification, was not presented to the students.

Experimental Group 2 and the contrast group received meaning-focused debriefings after the performance phase of strategic interaction. Meaning-focused debriefing was based on how successfully the intended communication was carried out. The teacher made comments on the students' performance in terms of accuracy in communicating messages, not accuracy of the target forms. The teacher emphasized whether the performers were successful in fulfilling such functions as giving opinions and disagreeing during the performance phase of strategic interaction.

Summary of instructional treatments. This study operationalized IE as follows: an instructional treatment in which a teacher pushes L2 learners to produce output and provides them with interactional modifications (e.g., requests for repetition and recasts) in order to lead them to notice a mismatch between their interlanguage grammar and the target language grammar, as well as to lead them to modify the incorrect output within the framework of strategic interaction.

IE has the following characteristics that are not present in other feedback treatments published to date:

1. IE is provided during an interactive communicative task using a scenario that sets a realistic problem to be solved using the L2. IE is, therefore, an instructional treatment, which has the strong feature of focus on form (i.e., placing the primary focus on communication with occasional form-focused treatments);
2. The scenario is devised to create not only a problem that elicits learner use of an L2, but also obligatory contexts for a particular form. Teachers can thus predict when learners are likely to make errors with the target form and can react to the errors in a prepared and systematic manner;⁷
3. IE has a clear and specific framework of instruction (e.g., the three-phase structure, the use of scenarios, and rigorous

instructional procedures), with which the teacher can lead learners to active classroom interaction and provide learners with form-focused treatments in an organized manner, even in a classroom with a large number of students.

In this study, two types of IE, which differed in the focus of the debriefing phases, were provided to two separate experimental groups. Experimental Group 1 (the IEF group; $n = 31$) received the IEF treatment (IE plus formal debriefing), involving implicit negative feedback and explicit grammar explanation. Experimental Group 2 (the IEM group; $n = 30$) received the IEM treatment (IE plus meaning-focused debriefing), involving implicit negative feedback without any explicit grammar instruction. The contrast group (the NEI group; $n = 30$) received the NEI treatment (non-enhanced interaction plus meaning-focused debriefing), involving no formal instruction. This assignment of different treatments to the three groups is summarized in Figure 2.

It must be noted here that the three instructional treatments (IEF, IEM, NEI) were not single variables but composite variables. For instance, when I compared the IEF treatment with the NEI treatment, I tested the value of the combination of IE and formal debriefing in comparison with the combination of non-enhanced interaction and meaning-focused debriefing. I was aware that comparing three treatments consisting of composite variables limited interpretation of the results of this study, because it would be impossible to attribute possible effects to individual variables. However, instructional types combining two variables were tested in this study because I wanted to investigate whether particular combinations of different types of formal instruction would be more effective for L2 learning than other combinations, because L2 teachers are more likely to mix several instructional options rather than exclusively using one type of formal instruction. That is, the three instructional treatments were considered to be compatible with instructional techniques typically used by L2 teachers.

	Performance Phase	Debriefing Phase
IEF Group (n = 31)	<Interaction Enhancement> problem-solving tasks (meaning-focused) implicit negative feedback (form-focused)	formal debriefing: explicit grammar explanation (form-focused)
IEM Group (n = 30)	<Interaction Enhancement> problem-solving tasks (meaning-focused) implicit negative feedback (form-focused)	meaning-focused debriefing (meaning-focused)
NEI Group (n = 30)	<Non-enhanced Interaction> problem-solving tasks (meaning-focused) meaning-oriented feedback (meaning-focused)	meaning-focused debriefing (meaning-focused)

IEF: Interaction Enhancement plus Formal Debriefing

IEM: Interaction Enhancement plus Meaning-Focused Debriefing

NEI: Non-enhanced Interaction plus Meaning-Focused Debriefing

Figure 2. Instructional treatments in the performance and debriefing phases

Post-Testing Measures

Two post-tests, an immediate post-test and a delayed post-test, were conducted with an interval of five weeks between the tests. To preclude a practice effect, the GJ task, the OP task, and

the WP task of the pretest were slightly modified for the immediate post-test, but without changing their complexity. Item sequences were shuffled and several content words were replaced with others equally familiar to the students. For the OS task, the same elicitation material was used in all three tests.

Scoring Procedure

Data gathered with the EADT were scored to quantify learner restructuring of interlanguage. First, participants who made few errors on the pretest were eliminated from the study. Eighty percent accuracy was set as the cutoff point, and the data of seven participants who scored 80 and above (the total score = 100) on the EADT were eliminated.

For each EADT task (OS, OP, WP, and GJ tasks), target-like use (TLU) scores were calculated. The TLU analysis was employed to measure EFL learners' restructuring of their interlanguage article system because it quantifies learner performance by taking overuse of the target form into account. TLU scores were obtained according to guidelines from Pica (1984). Articles were first scored for correct use in obligatory contexts. This score then became the numerator of a ratio whose denominator was the sum of the number of obligatory contexts for articles and the number of nonobligatory contexts in which articles were supplied inappropriately.⁸

The total score of the EADT was obtained by adding the TLU scores of the four tasks and dividing the sum by four. The highest possible EADT score was 100. Total scores for the indefinite article and for the definite article were obtained separately and referred to as EADT-INDEF and EADT-DEF, respectively. Total scores (EADT-TOTAL) combining EADT-INDEF scores and EADT-DEF scores were also obtained.

The reliability of each task for indefinite and definite article uses in the pretest was estimated by using the Kuder-Richardson 20 (KR-20) and Kuder-Richardson 21 (KR-21), and interscorer reliability. The KR-20 and KR-21 values for the WP task were .922

and .908, respectively. The KR-20 and KR-21 values for the OP task were .767 and .767, respectively. The KR-20 and KR-21 values for the GJ task were .820 and .752, respectively. The interscorer reliability of the OS task was .976. These values indicate that the measures testing the students' ability to use articles were considerably reliable.

To address Research Questions 1 through 3, data obtained through the EADT-INDEF and EADT-DEF were submitted to two separate repeated measures analyses of variance (ANOVAs) using one between-subject and one within-subject factorial design. In the repeated measures ANOVAs, Instruction (IEF, IEM, NEI) was the between-subject factor, whereas Test (pretest, immediate post-test, delayed post-test) was the within-subject factor. The first ANOVA was performed on the indefinite article data obtained from the EADT-INDEF and the second on the definite article data from the EADT-DEF. To address Research Question 4, score gains (Gain = immediate post-test score minus pretest score) on EADT-TOTAL for each task were submitted to an ANOVA using one between-subject factor (Instruction) and one within-subject factor (Task-Type). For Research Question 5, the total score gains of performers and those of observers were compared by using a repeated measures ANOVA with two between-subject factors (Instruction, Participation-Type) and one within-subject factor (Gain).

Results

Analyses of Interaction

First, quality and quantity of the interactions between the teacher and each representative during the performance and debriefing phases were analyzed to examine whether the two experimental groups received identical IE treatment during the performance phase. All the interactions performed during the performance phases of strategic interaction were recorded with an

8-mm videorecorder. I transcribed all the interactions, identified interactional modifications, and coded them into three categories: (1) teacher requests for repetition, (2) teacher requests for repetition that led learners to modify output correctly, and (3) teacher corrective recasts. Interrater reliability for coding was obtained by having another non-native EFL instructor identify and code interactional modifications in the transcription. The reliability was 1.00. Table 1 shows the frequencies of the interactional modifications provided to each experimental group during the performance phases of the three training sessions. A chi-square test revealed no statistically significant difference ($\chi^2 = 0.10, df = 1, ns$) between the two groups in the frequencies of the interactional modifications. This suggests that the two experimental groups received practically identical IE treatment in terms of the quality and quantity of interactional modifications.

Table 1

Frequencies of interactional modifications during interaction enhancement

Session	Experimental Group 1 (IEF)				Experimental Group 2 (IEM)			
	1	2	3	Total	1	2	3	Total
Student representatives (<i>N</i>)	3	4	3	10	3	4	3	10
Interactional modifications								
Teacher requests for repetition	2	3	4	9	2	5	2	9
Teacher requests for repetition that led learners to modify output correctly	1	1	2	4	2	0	1	3
Teacher corrective recasts	1	2	2	5	2	3	1	6

Another intervening variable that could affect results was exposure to English articles during instruction. Exposure to the target form during the performance and debriefing phases of strategic interaction was analyzed according to the observation data from the training sessions (i.e., audio- and videorecordings of the performance and debriefing phases). Frequencies of definite and indefinite article production by the instructor and student representatives were determined, and a chi-square test was performed to examine whether there was any significant difference in the total frequencies of correct English articles (produced by the instructor and student representatives) among the three groups (IEF = 226, IEM = 236, NEI = 205). The result revealed no significant difference ($\chi^2 = 2.25$, $df = 2$, *ns*). The three groups thus received generally equal exposure to the target form during the performance and debriefing phases.⁹

Analyses of the Effect of Instruction

Table 2 indicates the means and standard deviations of TLU scores on the EADT-INDEF and the EADT-DEF. Figures 3 and 4 graphically display the mean total scores of the EADT-INDEF and the EADT-DEF, respectively.

Indefinite article data. As shown in Table 3, the results of the repeated measures ANOVA for the TLU scores on the EADT-INDEF revealed a significant main effect for Instruction, a significant main effect for Test, and a significant two-way interaction between Instruction and Test. As the ANOVA revealed a significant two-way interaction effect for Instruction \times Test, the significant main effects found in this analysis (i.e., Instruction, Test) were interpreted in light of this interaction. This is because the significant interaction effect indicates that the two factors influenced the results interdependently. The results of the analysis of the Instruction \times Test interaction revealed that the simple main effects for both Instruction on Post-test 1 and Instruction on Post-test 2 were significant. The simple main effect for Instruction

Table 2

Descriptive statistics for the total TLU scores from the English Article Diagnostic Test

[Highest possible score = 100]

	<u>Experimental Group 1 (IEF)</u>					
	EADT-INDEF			EADT-DEF		
	PR	PS1	PS2	PR	PS1	PS2
<i>n</i>	31	31	31	31	31	31
<i>M</i>	24.90	66.06	60.81	39.65	64.13	65.74
<i>SD</i>	13.40	18.69	24.52	16.16	24.35	21.80

	<u>Experimental Group 2 (IEM)</u>					
	EADT-INDEF			EADT-DEF		
	PR	PS1	PS2	PR	PS1	PS2
<i>n</i>	30	30	30	30	30	30
<i>M</i>	30.87	44.57	49.50	40.63	48.27	54.13
<i>SD</i>	16.52	17.42	20.48	17.27	18.97	20.53

	<u>Contrast Group (NEI)</u>					
	EADT-INDEF			EADT-DEF		
	PR	PS1	PS2	PR	PS1	PS2
<i>n</i>	30	30	30	30	30	30
<i>M</i>	31.30	34.90	40.23	41.77	45.33	47.97
<i>SD</i>	12.19	15.81	16.58	9.00	13.61	15.57

Note. PR = Pretest, PS1 = Post-test 1, PS2 = Post-test 2
 EADT-INDEF = English Article Diagnostic Test (Indefinite Article Data)
 EADT-DEF = English Article Diagnostic Test (Definite Article Data)

on Pretest was not significant. These results suggest that for the indefinite article data, there were no significant differences in the pretest scores among the three groups (i.e., no significant pre-instructional difference), whereas the three groups performed differently on each post-test. The strength of association between

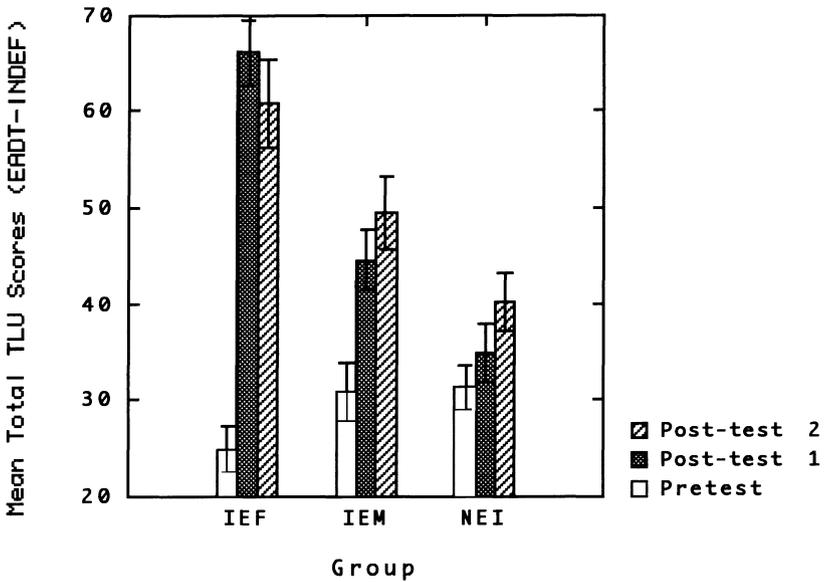


Figure 3. Mean total TLU scores (SE) from the pre- and post-tests for the English Article Diagnostic Test (Indefinite Article)

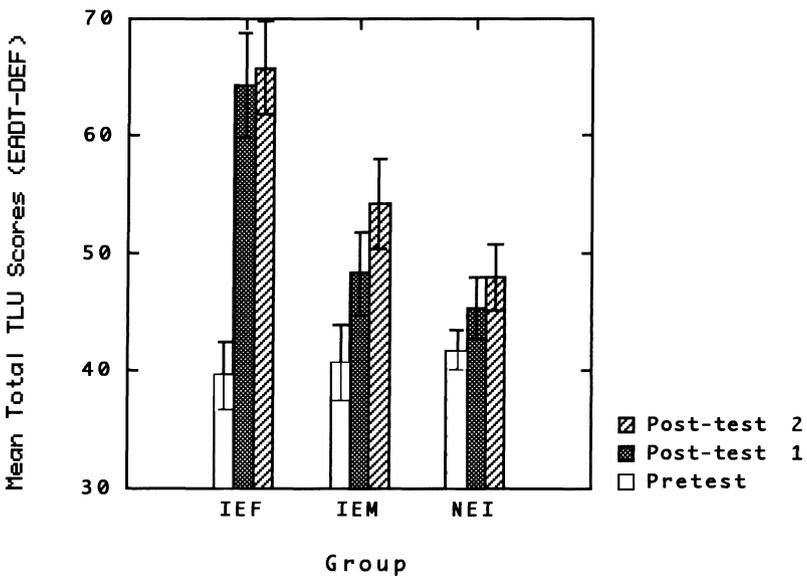


Figure 4. Mean total TLU scores (SE) from the pre- and post-tests for the English Article Diagnostic Test (Definite Article)

Instruction and the post-test scores was estimated with eta-squared. Results indicated that Instruction accounted for 36% of the variance in the Post-test 1 scores and 14% of the Post-test 2 scores.

Multiple comparisons among groups were performed using least square differences (LSD) to determine which groups were significantly different from each other. As summarized in Table 4, the results of between-group comparisons for the indefinite article in Post-test 1 revealed (1) that both experimental groups, which received IE, outperformed the contrast group, which received unenhanced interaction ($MSe = 311.77$, $LSD = 9.67$, $p < .05$) and (2) that the IEF group outperformed the IEM group ($MSe = 311.77$,

Table 3

Summary of results of repeated measures ANOVAs for the EADT

Source	df	F	
		INDEF	DEF
Between subjects			
Instruction (I)	2	7.33**	4.31*
S within-group error	88	(719.72)	(762.15)
Within subjects			
Test (T)	2	100.50**	47.26**
I × T	4	24.23**	9.36**
T × S within-group error	176	(125.25)	(123.80)
Simple main effects in Instruction × Test Interaction			
I in Pretest	2	1.87 <i>ns</i>	0.15 <i>ns</i>
S within-group error	88	(207.01)	(221.24)
I in Post-test 1	2	24.75**	7.86**
S within-group error	88	(311.77)	(394.64)
I in Post-test 2	2	7.18**	6.27**
S within-group error	88	(448.45)	(393.87)
T in IEF	2	121.50**	52.38**
T in IEM	2	22.57**	11.22**
T in NEI	2	4.89**	2.37†
T × S within-group error	176	(125.25)	(123.80)

Note. Values enclosed in parentheses represent mean square errors. INDEF = indefinite article data; DEF = definite article data; S = subjects.

† $p < .10$. * $p < .05$. ** $p < .01$.

LSD = 9.67, $p < .05$). The results of between-group comparisons in Post-test 2 revealed (1) that the IEF group also outperformed the IEM and NEI groups five weeks after instruction ($MSe = 448.45$, LSD = 10.88, $p < .05$), but (2) that the IEM group did not perform better than the NEI group in the delayed post-test.

Effect sizes for these between-group comparisons were calculated to examine the practical significance of the between-group differences. The effect sizes reported in Table 4 indicated that the between-group differences reported to be statistically significant were significant at a practical level, too, because they all represented at least medium size effects (above .50).¹⁰ The effect size for the comparison between the IEM group and the NEI group in the delayed post-test indicated that the between-group difference had a medium size effect (ES = 0.80). This suggests that the difference in the Post-test 2 scores between the IEM group and the NEI group was practically significant, though the difference was not statistically significant. Effect sizes for the between-group comparisons in the Pretest revealed that the mean for the IEF group was significantly lower than those for the other two groups.

Table 4

Summary of the results of between-group comparisons (LSD)

	Indefinite Article Data (ES)			Definite Article Data (ES)		
Pretest	IEF = IEM	(-0.39)	IEF = IEM	(-0.05)		
	IEF = NEI	(-0.50)	IEF = NEI	(-0.85)		
	IEM = NEI	(-0.03)	IEM = NEI	(-0.09)		
Post-test 1	IEF > IEM *	(1.19)	IEF > IEM *	(0.72)		
	IEF > NEI *	(1.80)	IEF > NEI *	(0.95)		
	IEM > NEI *	(0.58)	IEM = NEI	(0.18)		
Post-test 2	IEF > IEM *	(0.50)	IEF > IEM *	(0.55)		
	IEF > NEI *	(0.98)	IEF > NEI *	(0.93)		
	IEM = NEI	(0.50)	IEM = NEI	(0.34)		

Note. ES = Effect Sizes.

* $p < .05$.

The results of the analyses of simple main effect for Test revealed that there were significant simple main effects for Test in all three groups (see Table 3). Results of eta-squared indicated that Test accounted 58%, 20%, and 5% of the variance in the IEF group, the IEM group, and the NEI group, respectively. Multiple comparisons among tests using LSD for the indefinite article data were made on the simple main effects for Test in order to determine which test results were significantly different from each other for the three groups receiving three different treatments. As summarized in Table 5, the results of between-test comparisons for the IEF group revealed (1) that the IEF group did significantly better on the two post-tests than on the Pretest for the indefinite article ($MSe = 125.25$, $LSD = 5.69$, $p < .05$), and (2) that there was no significant difference between Post-test 1 scores and Post-test 2 scores, suggesting that their gains did not significantly decrease over the five-week period.

Table 5

Summary of the results of between-test comparisons (LSD)

	Indefinite Article Data (ES)	Definite Article Data (ES)
IEF		
Post-test 1 > Pretest *	(2.53)	Post-test 1 > Pretest * (1.18)
Post-test 2 > Pretest *	(2.09)	Post-test 2 > Pretest * (1.35)
Post-test 2 = Post-test 1	(0.24)	Post-test 2 = Post-test 1 (0.07)
IEM		
Post-test 1 > Pretest *	(0.81)	Post-test 1 > Pretest * (0.42)
Post-test 2 > Pretest *	(1.00)	Post-test 2 > Pretest * (0.71)
Post-test 2 = Post-test 1	(0.26)	Post-test 2 > Post-test 1 * (0.29)
NEI		
Post-test 1 = Pretest	(0.26)	Post-test 1 = Pretest (0.30)
Post-test 2 > Pretest *	(0.61)	Post-test 2 > Pretest * (0.49)
Post-test 2 = Post-test 1	(0.33)	Post-test 2 = Post-test 1 (0.18)

Note. ES = Effect Sizes.

* $p < .05$.

Between-test comparisons for the IEM group revealed (1) that the IEM group did significantly better on the two post-tests than on the Pretest for the indefinite article ($MSe = 125.25$, $LSD = 5.69$, $p < .05$), and (2) that there was no significant difference between Post-test 1 scores and Post-test 2 scores, suggesting that the gains measured in the first post-test lasted for five weeks.

Between-test comparisons for the NEI group revealed (1) that the NEI group performed better on Post-test 2 than on the Pretest ($MSe = 125.25$, $LSD = 5.69$, $p < .05$) and (2) that there were no significant differences between Pretest scores and Post-test 1 scores, or between Post-test 1 scores and Post-test 2 scores. These results indicate that the NEI group gradually improved their performance on the test for the indefinite article during the training and testing sessions, which lasted more than two months.

Effect sizes for between-test comparisons for each group indicated that all the between-group differences reported to be statistically significant were significant at a practical level, too (Table 5). The between-test comparisons that were not statistically significant (Post-test 1 \times Post-test 2 for the three groups) had small size effects, suggesting that the differences were not likely to be visible but "not so small as to be trivial" (Cohen, 1992, p. 156). In other words, all three groups performed somewhat better on the delayed post-test than on the immediate post-test. Effect-size analysis also revealed that the difference between the pretest scores and the immediate post-test scores for the NEI group had a small size effect ($ES = 0.26$); the NEI group performed better to some degree in the immediate post-test than in the pretest.

Definite article data. As shown in Table 3, the results of the repeated-measures ANOVA for the TLU scores on the EADT-DEF suggest that for the definite article data, there were no significant differences in the pretest scores among the three groups (i.e., no significant preinstructional difference), whereas the three groups performed differently on each post-test. Results of eta-squared estimating the strength of association indicated that Instruction accounted for 15% of the variance in learner performance on Post-test 1 and 12% on Post-test 2.

After the analysis of the significant two-way interaction effect for Instruction \times Test, multiple comparisons (LSD) between groups for the definite article in Post-test 1 were performed. As summarized in Table 4 (q.v.), results revealed (1) that the IEF group outperformed not only the contrast group, but also the IEM group ($MSe = 394.64$, $LSD = 10.20$, $p < .05$); and (2) that the IEM group did not outperform the contrast group. Between-group comparisons for the definite article in the delayed post-test revealed (1) that the IEF group outperformed the IEM and NEI groups five weeks after the instruction ($MSe = 393.87$, $LSD = 10.19$, $p < .05$), and (2) that the IEM group did not outperform the NEI group in the delayed post-test.

The results of the analysis of simple main effect for Test revealed that there were significant simple main effects for Test in the two experimental groups (see Table 3). No significant simple main effect for Test was found for the NEI group, though the effect was marginally significant ($p < .10$). Eta-squared results show that Test accounted 37%, 11%, and 3% of the variance in the IEF group, the IEM group, and the NEI group, respectively. These results suggest that the two experimental groups performed differently in the three tests, whereas the NEI group's performance on the three tests did not differ significantly for the definite article, though there was a trend toward a significant difference.

Effect sizes for the between-group comparisons for the definite article data indicated that the statistically significant differences were also significant at a practical level (Table 4). The effect size for the difference between the IEM group and the NEI group in Post-test 2 ($ES = 0.34$) suggested that there was a small but not trivial difference between the two groups. Effect sizes for the pretest data revealed that the mean for the IEF group was significantly lower than that for the NEI group.

As shown in Table 5, between-tests comparisons for the definite article in the IEF group revealed (1) that the IEF group did significantly better on the two post-tests than on the Pretest for the definite article ($MSe = 123.80$, $LSD = 5.66$, $p < .05$), and (2) that there was no significant difference between Post-test 1 scores

and Post-test 2 scores, suggesting that the gains they obtained did not fade during the five-week interval between the two post-tests.

Between-test comparisons for the IEM group indicated (1) that the IEM group did significantly better on the two post-tests than on the Pretest for the definite article ($MSe = 123.80$, $LSD = 5.66$, $p < .05$), and (2) that the IEM group performed better on Post-test 2 than on Post-test 1 ($MSe = 123.80$, $LSD = 5.66$, $p < .05$), suggesting that the gains increased over the five-week interval between the two post-tests.

Between-test comparisons for the NEI group revealed (1) that the group performed better on Post-test 2 than on the Pretest ($MSe = 123.80$, $LSD = 5.66$, $p < .05$) and (2) that there were no significant differences between Pretest scores and Post-test 1 scores, or between Post-test 1 scores and Post-test 2 scores. These results suggest that the NEI group gradually improved their performance with the definite article during the training and testing sessions, which, as was the case for the indefinite article, lasted more than two months.

As reported in Table 5, the effect size analysis for the between-test comparisons for the definite article data supported the results of the statistical analysis, except that the difference between the Pretest scores and the Post-test 1 scores for the NEI group, which was not statistically significant, had a small size effect ($ES = 0.30$).

Task type. This study also investigated the effects of task types, because the test used to measure the students' performance with articles consisted of four different tasks: OS, OP, WP, and GJ. In order to determine if task type had any influence on learner performance with English articles, gain scores obtained by the three groups were submitted to an ANOVA using one between-subject factor (Instruction: IEF, IEM, NEI) and one within-subject factor (Task Type: OS, OP, WP, and GJ). Table 6 displays means and standard deviations of score gains for four tasks on the EADT. The ANOVA and multiple comparisons of Instruction (Groups) for each task revealed (1) that the IE treatments had greater effects on the students' performance on all tasks than the control treatment (except for the IEM treatment in WP and GJ tasks) and

(2) that the IEF treatment had greater effects on the students' performance on all tasks than the IEM treatment (results of multiple comparisons of Instruction for each task type: for OS, IEF > IEM > NEI, $MSe = 444.22$, $LSD = 10.83$, $p < .05$; for OP, IEF > IEM > NEI, $MSe = 357.65$, $LSD = 9.71$, $p < .05$; for WP, IEF > IEM = NEI, $MSe = 502.06$, $LSD = 11.51$, $p < .05$; for GJ, IEF > IEM = NEI, $MSe = 207.11$, $LSD = 7.39$, $p < .05$). These results suggest that the IEF treatment had positive effects on performance on both oral and written tasks, whereas the IEM treatment had effects only on oral tasks.

The results of the multiple comparisons of Task Type for each group also revealed that there was a general tendency toward IE treatments (IEF and IEM) having greater effects on performance on oral tasks than on written tasks.

Participation type. This study also examined whether differ-

Table 6

Means and standard deviations of gain scores for four tasks on the EADT

Task	<i>M</i> (Gain Score)	<i>SD</i>
[IEF $n = 31$]		
Oral story description	44.58	25.31
Oral picture description	31.71	21.55
Written picture description	29.87	27.90
Grammaticality judgment	26.45	18.04
[IEM $n = 30$]		
Oral story description	15.70	21.55
Oral picture description	17.60	18.20
Written picture description	3.60	18.01
Grammaticality judgment	8.77	12.29
[NEI $n = 30$]		
Oral story description	4.13	14.88
Oral picture description	7.60	16.55
Written picture description	1.20	19.85
Grammaticality judgment	1.83	11.85

Note. Gain Score = immediate post-test score minus pretest score.

ence in participation type (performer versus observer) in strategic interaction made a difference in score gains on the EADT. Table 7 displays the descriptive statistics for scores on the EADT for performers and observers. The total gains in TLU scores of the three groups were submitted to a repeated measures ANOVA using two between-subject factors (Instruction; Participation-Type) and one within-subject factor (Score Gain). The results of the ANOVA revealed no significant main effect for Participation Type ($F(1/85) = 0.12, ns$), and no significant interaction effects

Table 7

Descriptive statistics for total scores on the EADT for performers and observers

[Highest possible score = 100]

	Pretest	Post-test 1	Post-test 2
Experimental Group 1 (IEF)			
[Performers, $n = 10$]			
<i>M</i>	35.11	74.56	74.00
<i>SD</i>	13.50	16.00	20.89
[Observers, $n = 21$]			
<i>M</i>	31.73	62.55	59.59
<i>SD</i>	12.92	21.53	22.05
Experimental Group 2 (IEM)			
[Performers, $n = 10$]			
<i>M</i>	41.30	49.30	53.20
<i>SD</i>	13.40	13.61	16.58
[Observers, $n = 20$]			
<i>M</i>	33.50	46.45	51.80
<i>SD</i>	16.40	18.67	20.84
Contrast Group (NEI)			
[Performers, $n = 10$]			
<i>M</i>	34.83	38.17	41.33
<i>SD</i>	2.64	10.40	19.36
[Observers, $n = 20$]			
<i>M</i>	37.17	41.08	45.08
<i>SD</i>	9.79	13.93	14.56

(Instruction \times Participation-Type: $F(2/85) = 1.97$, *ns*; Participation-Type \times Score Gain: $F(1/85) = 0.00$, *ns*; Gain \times Instruction \times Participation-Type: $F(2/85) = 0.11$, *ns*). This indicates that difference in participation type had no significant influence on L2 learning of English articles and that the treatments provided in this study were beneficial for all participants regardless of participation type.

Summary of Results

Major results of the study are summarized below:

1. IE had positive effects, which lasted for at least five weeks, on L2 learning of English articles;
2. IE plus formal debriefing (IEF) had a greater effect than IE plus meaning-focused debriefing (IEM);
3. IE treatments had greater effects on learner performance with articles than the contrast treatment on all elicitation tasks (except for the IEM treatment in written tasks), but the strength of the effects varied with the type of task (the instruction had greater effects on oral tasks than on written tasks);
4. learners who observed IE improved their performance with the English article system as much as learners who directly participated in it.

Discussion

My first research question was: Does IE affect EFL learners' restructuring of their interlanguage article systems? Results, especially those from between-test comparisons, indicate that the IEF group receiving implicit negative feedback plus explicit formal debriefing and the IEM group receiving implicit negative feedback plus meaning-focused debriefing performed significantly better than the contrast group (NEI group) with both the indefinite and the definite articles on the post-tests. IEF and IEM treatments

thus both appear to have been effective in helping the EFL learners restructure their interlanguage article system, though the effect of the IEM treatment was more limited than that of the IEF treatment. Therefore, the answer to Research Question 1 was yes, IE positively affected EFL learners' restructuring of their interlanguage article systems.

These results lend strong support to previous claims for the efficacy of focus on form (Doughty & Williams, 1998; Lightbown & Spada, 1990; Long, 1996; Long & Robinson, 1998; Skehan, 1996; Spada & Lightbown, 1993). In particular, these results provide some evidence for Lightbown and Spada's (1990) claim that "accuracy, fluency, and overall communicative skills are probably best developed through instruction that is primarily meaning-based but in which guidance is provided through timely form-focused activities and correction in context" (p. 443). As this study looked at only accuracy, further research investigating the effect of IE on fluency and overall communicative skills is in order.

The results also provide support for the argument that corrective feedback is beneficial to L2 learners in modifying their overgeneralized interlanguage grammars (e.g., Carroll & Swain, 1993; Doughty & Varela, 1998; Pica, 1992; Takashima, 1995; Tomasello & Herron, 1989). Here, the IEM treatment, providing implicit negative feedback without any explicit grammar explanation during problem-solving tasks, promoted better accuracy in article use than the NEI treatment in which no form-focused feedback was provided—counterevidence to the claim that language learners do not notice and use form-focused feedback (e.g., Krashen, 1984).

It must also be noted that IE was effective when it targeted specific forms. Previous feedback studies have revealed that negative feedback, such as recasts, is beneficial to L2 learning when it specifically targets particular forms (e.g., Doughty & Varela, 1998), whereas feedback is less effective when it is given without focusing on any specific form (e.g., Lyster & Ranta, 1997). The results of this study suggest that instruction in which teachers provide

negative feedback during interactive tasks can positively affect L2 learning when it has a particular linguistic focus.

Though results of this study suggest positive effects of IE, one caution for interpretation is raised: a practice effect may have affected the students' performance on post-tests. That is, it might have been possible that the learners improved their scores on the post-tests simply because they took the test more than once. To obtain further data for examination of the practice effect, a follow-up study was undertaken. Twenty-one new students comparable to the former groups took the EADT test three times without any intervening instruction. The scores on the three tests were submitted to a multivariate analysis of variance (MANOVA) with three variables (Article-Type, Task-Type, Test). The MANOVA revealed that no significant difference was found among the three test scores ($F(2/40) = .10, ns$). The result suggests that these students did not exhibit any significant practice effect. This provides me with reasonable grounds for claiming that the gains in post-test scores that the students in the main study obtained were not attributable to practice effects.

That the contrast group, which received unenhanced interaction, did not significantly improve their scores on the immediate post-test is also significant in showing that purely meaning-oriented interaction without any kind of form-focused treatment (i.e., no IE treatment) did not lead the learners to change their interlanguage grammar as successfully as the IE treatments did. This is consistent with the claim that an interactive task plus a small amount of form-focused treatment is more beneficial for learners in restructuring their interlanguage systems than is an interactive task without any focus on form (Loschky & Bley-Vroman, 1993; Skehan, 1996). The NEI group, however, improved their performance to a small degree in the immediate post-test, as indicated by the effect-size analysis, and also performed significantly better in the delayed post-test. This result suggests that meaning-oriented interaction, in which the learners were presented with a number of obligatory contexts for the target form

and exposed to a large number of correct uses of the target form, may have had some effects on L2 learning.

My second research question was: Do two types of IE, differing in the manner of focus on form (IEF versus IEM), have different effects on EFL learners' restructuring of their interlanguage article systems? Results indicate that the IEF group outperformed the IEM group on all post-tests. This suggests that IE plus formal debriefing was more beneficial for L2 learning of English articles than IE plus meaning-oriented debriefing.¹¹ Therefore, the answer to Research Question 2 was yes, two types of IE, different from each other in the manner of focus on form, had different effects on L2 learning. This finding is consistent with Carroll and Swain's (1993) findings on the effects of explicit metalinguistic instruction in comparison with implicit feedback treatments. They reported that a group getting explicit metalinguistic feedback outperformed other groups receiving explicit utterance rejection, modeling/implicit negative feedback, and indirect metalinguistic feedback treatments (Carroll & Swain, 1993, p. 366). They concluded that "claims that learning about the language is useless in promoting learning of the language may be incorrect" (p. 372). The results of this study provide support for Carroll and Swain's argument for the positive role of explicit grammar instruction and for other theorists' claims that explicit linguistic instruction is useful in L2 learning (DeKeyser, 1995; R. Ellis, 1997; Green & Hecht, 1992; Robinson, 1996; Rosa & O'Neil, 1999; Sorace, 1985; Terrell, 1991).

The positive effect of explicit grammar instruction on L2 learning of the English article system is of great importance, because the article system is considered to involve functionally and semantically complex rules (Doughty & Williams, 1998; Hulstijn, 1995; Pica, 1985). Previous studies have not provided clear-cut support for an advantage of explicit instruction over implicit instruction in L2 acquisition of complex rules, though it has been relatively clearly shown that explicit instruction for simple rules is better than implicit instruction (see Doughty & Williams, 1998; Long & Robinson, 1998; for reviews of related studies, see

DeKeyser, 1995, and Robinson, 1996 for empirical data). From a pedagogical perspective, Hulstijn (1995) proposed that for complex rules teachers should provide their learners with a number of well-chosen and well-organized examples without relying on grammar explanations. The results of the present study, however, suggest that the learning of complex rules can be facilitated by explicit instruction when it is provided along with implicit instruction, as in IE. This provides learners with numerous examples and implicit feedback. Similar results were obtained by N. Ellis (1993), who examined the effect of three different types of instruction on the learning of highly complex rules of Welsh morphology by native speakers of English. He found that explicit instruction in which rules were blended with relevant examples had an advantage over both implicit instruction that provided examples alone and explicit instruction that presented rules alone. Doughty and Williams (1998) summarized the findings of these studies: "The combination of rules plus carefully considered examples of the rules . . . appears to be the most effective learning condition for complex rules, at least insofar as has been revealed by effects of nonclassroom, experimental learning on knowledge . . ." (p. 233). The results of this study suggest that the combination of rules and examples is beneficial for L2 learners in a real-classroom setting, too.

As the between-group comparisons did not reveal a significant positive effect of the IEM treatment for the definite article, it appears appropriate to claim that the combination of an implicit form-focused treatment and an explicit form-focused treatment (i.e., the IEF treatment) had a greater transfer effect on the untargeted form (the definite article) than instruction providing an implicit form-focused treatment only (i.e., the IEM treatment). This result is another support for the significant role of explicit grammar instruction in L2 acquisition.

My third research question was: If IE has an effect on EFL learners' restructuring of their interlanguage article system, will the effect hold over the post-test period? Results of the between-test comparisons indicate that the effects of IE endured over the

five-week interval between the two post-tests (see Table 5).¹² Therefore, the answer to Research Question 3 was yes, the effect of IE held over at least five weeks. This finding leads us to assume that instruction that appropriately incorporates form-focused treatments into communication-oriented language teaching can have a lasting positive effect on L2 acquisition. In other words, it appears that a key to longer-term effect of instruction is the manner in which form-focused instruction and meaning-focused instruction are balanced. More specifically, the results of this study suggest that lasting instructional effects can be obtained through providing learners with opportunities to use the target form in a meaning-oriented task in combination with appropriate form-focused treatments that aim at strengthening the connection between form and function.

My fourth research question was: Does IE have differential effects on learner performance with English articles depending on the elicitation task used? Results indicate that IE treatments had greater effects on the students' performance on all tasks than the contrast treatment (except for the IEM treatment in WP and GJ tasks) and that the IEF treatment had greater effects on the students' performance on all tasks than the IEM treatment. In other words, this result suggests that the effect of the IEM treatment was more limited than that of the IEF treatment. This result, therefore, is further support for the superiority of IE plus formal debriefing over IE plus meaning-focused debriefing. It might be possible to assume that this finding lends support to a prediction of skill acquisition theory that declarative knowledge (i.e., explicit linguistic knowledge) is generative. Johnson (1996), for example, claims that learners who have acquired declarative knowledge on a linguistic form have a generative capacity to use that knowledge in whatever circumstance presents itself. The superiority of the IEF treatment to the IEM treatment might be due to this generativity of declarative knowledge.

The data also indicate that there was a general tendency toward IE treatments (IEF and IEM) having greater effects on performance on oral tasks than on written tasks. The most

plausible explanation for this difference would be that the completely oral instruction had a direct training effect on the oral tasks.

That the students receiving IE raised their scores on oral tasks has an important implication for L2 teachers. The results of this study indicate that a carefully designed interactive task can help learners improve performance on oral tasks that require L2 learners' automatic processing of linguistic knowledge.

This evidence of task variance also adds support to the claim that L2 learners' performance must be investigated not only through a single measure but with a test that consists of diverse elicitation tasks, because learner performance is significantly influenced by the type of elicitation tasks (Tarone & Parrish, 1988). The answer to Research Question 4 was yes, IE had differential effects on learner performance on different elicitation tasks.

My fifth research question was: Does the difference in participation type (performers versus observers) in IE make a difference in learner performance with English articles? Analyses revealed that the difference in participation type in IE made no difference in score gains. The results indicate that the treatments provided in this study were beneficial for all participants regardless of participation type. This finding has an important pedagogical implication for foreign language teachers who teach in large-size classes where teachers can give feedback to only a limited number of students (e.g., teaching EFL in Japanese high schools). It suggests that the instructional treatments provided to a small number of students were beneficial, even for the students who did not receive the instruction directly. The answer to Research Question 5 was no, IE was beneficial both to learners who performed roles during the performance phase and to those who observed the performance.

Finally, it must be noted that this study has several methodological limitations, including the use of intact classes rather than randomly selected participants, the possible researcher expectancy, and the limited number of the participants. These problems limit the generalizability of the findings reported in this

study, and follow-up research must be conducted to determine whether IE is effective under different conditions.

Pedagogical Implications

The results of this study reaffirm focus-on-form researchers' claims that L2 instruction in which form-focused treatments are integrated into meaning-oriented interactive tasks in a timely fashion facilitates L2 learning. In particular, this study suggests that IE treatments that systematically combine such instructional techniques as output enhancement, input enhancement, problem-solving tasks, and explicit grammar instruction can be beneficial for guiding EFL learners to restructure their interlanguage systems. This finding is in sharp contrast with the position that learner errors should not be corrected (e.g., Krashen, 1984, 1992). Along with the claim that meaning-oriented communicative tasks, which have obtained great popularity, may not contribute sufficiently to grammatical competence (Loschky & Bley-Vroman, 1993), the data presented here encourage L2 teachers to incorporate corrective treatments that lead learners' attention to linguistic forms into their communicative language teaching in an appropriate and timely manner based on cognitive theories of L2 acquisition.

This study also suggests that explicit grammar instruction plays an important role in helping L2 learners develop their interlanguage when it is systematically incorporated into an interactive problem-solving task. The data obtained here indicate that brief and focused explicit grammar instruction can exert a positive effect on L2 learning if it accompanies an interactive communicative task that makes form–function connections more salient.

There are two other major pedagogical implications arising from this study: (1) having even a limited number of students participate in focus-on-form tasks can be beneficial to the entire class; (2) focus-on-form treatments involving both implicit and

explicit formal instruction can help learners improve their performance on not only oral tasks but also written tasks.

The results of this study also confirm the claim that guiding learners to focus on form within meaning-focused instruction based on cognitive theories of L2 acquisition is definitely profitable.

Revised version accepted 01 June 2000

Notes

¹In this study the term “L2” is used as a superordinate of “second language” and “foreign language.” Where there is a need to distinguish these terms, each construct is specified.

²The data of 16 participants who missed any two of the six testing and training sessions and another 7 students who made few article errors in the pretest were excluded from statistical analyses (see the *Scoring Procedure* section).

³As no consensus on the definition of long-term effect has yet been attained in L2 acquisition research, it appears that conducting several post-tests over a long period of time is the most valid way to examine whether effects of a particular training on L2 learning are long lasting. In this study, however, administering two post-tests with a five-week interval was dictated by scheduling constraints.

During the five-week interval, the students in all three groups attended regular EFL courses and academic courses in English literature and linguistics. They were exposed to input that included English articles, but they did not receive any instruction specifically focused on English articles during this interval.

⁴This choice of form originally derived from projection theory and discourse markedness theory in regard to noun phrases. Projection theory predicts that instruction incorporating marked data potentially generalizes beyond the marked context to other contexts that are implicationally associated with the marked data (Doughty, 1991; Zobl, 1983). Discourse markedness theory claims that the indefinite NPs are more marked than definite NPs in terms of discourse constraints (Chaudron & Parker, 1990). Based on these two theories I assumed at the onset of this study that instruction focusing on the more marked indefinite article should have a concomitant positive effect on errors with the less marked definite article. This assumption, however, received criticisms that projection theory does not hold true with the English article system, because the indefinite and the definite articles in English have a closed binary relation, such that change in one form almost necessarily results in a change in another, having nothing to do with markedness. Taking this view into account, I restrict discussion to whether instruction on one form

(the indefinite article) affected performance on the other, binarily related form (the definite article).

⁵“Request for repetition” is also referred to as “clarification request.”

⁶The well-known U-shaped developmental pattern, in which learners first use a linguistic rule correctly, then overgeneralize it, and finally constrain it properly, is considered to be a type of restructuring.

⁷Though scenarios are generally valuable for setting obligatory contexts for particular linguistic forms, this feature is not strongly seen in this study, because the target form is the indefinite article, which is obligatory in any task.

⁸For the GJ task, the students were directed not only to judge the grammaticality of underlined parts but to correct all incorrect parts. They overused both the indefinite article and the definite article in nonobligatory contexts. TLU scores for the GJ task were obtained by dividing the number of correctly supplied articles in obligatory contexts by the sum of the number of obligatory contexts and the number of inappropriately supplied articles in nonobligatory contexts.

⁹Needless to say, the participants were exposed to English articles outside the class sessions involved in this experiment (e.g., other EFL classes, etc.). Though such exposure also constitutes possible intervening variables, no attempt was made to control such exposure, for educational and ethical reasons.

¹⁰Cohen (1992) practically defined medium effect size as “an effect likely to be visible to the naked eye of a careful observer” (p. 156).

¹¹Analysis of the qualitative changes in errors made by the three groups reveals that the IEM treatment had an excessive, or negative, impact, which led some learners to overuse a/an in definite contexts, while there was no significant increase in the number of overgeneralization errors with the indefinite article among the IEF group. This is another support for the claim that the IEF treatment worked better than the IEM treatment. The explicit grammar explanation the IEF group received appears to have helped prevent the learners from overgeneralizing the target form. See Muranoi (1996, 2000) for detailed analyses of changes in errors due to the IE treatments.

¹²These results must be interpreted somewhat cautiously, however, given the interval between the pretest and post-tests.

References

- Aljaafreh, A., & Lantolf, J. (1994). Negative feedback as regulation and second language learning in the Zone of Proximal Development. *Modern Language Journal*, 78, 465–483.

- Beck, M., & Eubank, L. (1991). Acquisition theory and experimental design: A critique of Tomasello and Herron. *Studies in Second Language Acquisition*, 13, 73–76.
- Brown, J. D. (1988). *Understanding research in second language learning*. New York: Cambridge University Press.
- Carroll, S., & Swain, M. (1993). Explicit and implicit negative feedback: An empirical study of the learning of linguistic generalizations. *Studies in Second Language Acquisition*, 15, 357–386.
- Chaudron, C., & Parker, K. (1990). Discourse markedness and structural markedness: The acquisition of English noun phrases. *Studies in Second Language Acquisition*, 12, 43–64.
- Cheng, P. (1985). Restructuring versus automaticity: Alternative accounts of skill acquisition. *Psychological Review*, 92, 412–423.
- Cohen, J. (1992). A power primer. *Psychological Bulletin*, 112, 155–159.
- DeKeyser, R. (1995). Learning L2 grammar rules: An experiment with a miniature linguistic system. *Studies in Second Language Acquisition*, 17, 379–410.
- 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.
- Di Pietro, R. (1987). *Strategic interaction: Learning languages through scenarios*. New York: Cambridge University Press.
- Doughty, C. (1991). Second language instruction does make a difference: Evidence from an empirical study of SL relativization. *Studies in Second Language Acquisition*, 13, 431–469.
- Doughty, C. (1994). Fine-tuning of feedback by competent speakers to language learners. In J. Alatis (Ed.), *Georgetown University Roundtable on Languages and Linguistics 1993* (pp. 96–108). Washington, DC: Georgetown 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.
- 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, N. (1993). Rules and instances in foreign language learning: Interactions of explicit and implicit knowledge. *European Journal of Cognitive Psychology*, 5, 289–318.
- Ellis, R. (1997). *SLA research and language teaching*. Oxford: Oxford University Press.

- Ellis, R., & He, X. (1999). The role of modified input and output in the incidental acquisition of word meanings. *Studies in Second Language Acquisition, 21*, 285–301.
- Farrar, M. (1992). Negative evidence and grammatical morpheme acquisition. *Developmental Psychology, 28*, 90–98.
- Fujimura, T., Kato, Y., Ahmed, M., & Leong, M. (Eds.). (1996). *Proceedings of the 7th conference on second language research in Japan*. Yamato-machi, Niigata: Language Program, International University of Japan.
- Gass, S. (1988). Integrating research areas: A framework for second language studies. *Applied Linguistics, 9*, 198–217.
- Gass, S. (1997). *Input, interaction, and the second language learner*. Mahwah, NJ: Lawrence Erlbaum.
- Green, P., & Hecht, K. (1992). Implicit and explicit grammar: An empirical study. *Applied Linguistics, 13*, 168–184.
- Hulstijn, J. (1995). Not all grammar rules are equal: Giving grammar instruction its proper place in foreign language teaching. In R. Schmidt (Ed.), *Attention & awareness in foreign language learning* (pp. 359–386). Honolulu, HI: University of Hawai'i Press.
- Johnson, K. (1996). *Language teaching & skill learning*. Oxford: Blackwell.
- Karmiloff-Smith, A. (1986). Stage/structure versus phase/process in modeling linguistic and cognitive development. In I. Levin (Ed.), *Stage and structure: Reopening the debate* (pp. 164–190). Norwood, NJ: Ablex.
- Karmiloff-Smith, A. (1992). *Beyond modularity: A developmental perspective on cognitive science*. Cambridge, MA: The MIT Press.
- Kowal, M., & Swain, M. (1994). Using collaborative language production tasks to promote students' language awareness. *Language Awareness, 3*, 73–93.
- Krashen, S. (1984). *The input hypothesis: Issues and implications*. New York: Longman.
- Krashen, S. (1992). Formal grammar instruction: An educator comments. . . . *TESOL Quarterly, 26*, 409–411.
- Lightbown, P., & Spada, N. (1990). Focus-on-form and corrective feedback in communicative language teaching: Effects on second language learning. *Studies in Second Language Acquisition, 12*, 429–448.
- Long, M. (1991). Focus on form: A design feature in language teaching methodology. In K. De Bot, R. Ginsberg, & C. Kramsch (Eds.), *Foreign language research in cross-cultural perspective* (pp. 39–52). Philadelphia: John Benjamins.
- Long, M. (1996). The role of the linguistic environment in second language acquisition. In W. C. Ritchie & T. K. Bhatia (Eds.), *Handbook of second language acquisition* (pp. 413–468). New York: Academic Press.

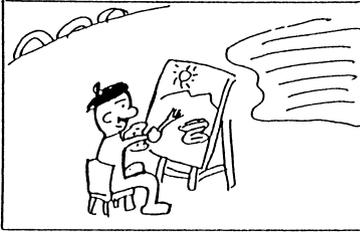
- Long, M., Inagaki, S., & Ortega, L. (1998). The role of implicit negative feedback in SLA: Models and recasts in Japanese and Spanish. *Modern Language Journal*, 82, 357–371.
- Long, M., & Robinson, P. (1998). Focus on form: Theory, research, and practice. In C. Doughty & J. Williams (Eds.), *Focus on form in classroom second language acquisition* (pp. 15–41). New York: Cambridge University Press.
- Loschky, L., & Bley-Vroman, R. (1993). Grammar and task-based methodology. In G. Crookes and S. Gass (Eds.), *Tasks and language learning: Integrating theory and practice* (pp. 123–167). Clevedon, Avon: Multilingual Matters.
- Lyster, R. (1998). Recasts, repetition and ambiguity in L2 classroom discourse. *Studies in Second Language Acquisition*, 20, 50–81.
- 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. (1999). Input, interaction, and second language development: An empirical study of question formation in ESL. *Studies in Second Language Acquisition*, 21, 557–587.
- Mackey, A., & Philp, J. (1998). Conversational interaction and second language development: Recasts, responses, and red herrings? *Modern Language Journal*, 82, 338–356.
- Master, P. (1990). Teaching the English articles as a binary system. *TESOL Quarterly*, 24, 461–478.
- Master, P. (1995). Consciousness raising and article pedagogy. In D. Belcher & G. Brain (Eds.), *Academic writing in a second language* (pp. 183–204). Norwood, NJ: Ablex.
- McLaughlin, B. (1987). *Theories of second-language learning*. London: Edward Arnold.
- McLaughlin, B. (1990). Restructuring. *Applied Linguistics*, 11, 113–128.
- Muranoi, H. (1996). *Effects of interaction enhancement on restructuring of interlanguage grammar: A cognitive approach to foreign language instruction*. Unpublished doctoral dissertation, Georgetown University, Washington, DC.
- Muranoi, H. (2000). *Focus on form: Implicit and explicit form-focused instruction incorporated into a communicative task*. Manuscript in preparation.
- Nelson, K. (1987). Some observations from the perspective of the rare event cognitive comparison theory of language acquisition. In K. Nelson & A. Kleeck (Eds.), *Children's language* (Vol. 6, pp. 289–331). Hillsdale, NJ: Lawrence Erlbaum.
- Pica, T. (1984). Methods of morpheme quantification: Their effect on the interpretation of second language data. *Studies in Second Language Acquisition*, 6, 69–78.

- Pica, T. (1985). Instruction on second-language acquisition. *Applied Linguistics*, 6, 214–222.
- Pica, T. (1992). The textual outcomes of native speaker–non-native speaker negotiation: What do they reveal about second language learning? In C. Kramsch & S. McConnell-Ginet (Eds.), *Text and context: Cross-disciplinary perspectives on language study* (pp. 198–237). Cambridge, MA: Heath.
- Pica, T., Holliday, L., Lewis, N., & Morgenthaler, M. (1989). Comprehensible output as an outcome of linguistic demands on the learner. *Studies in Second Language Acquisition*, 11, 63–90.
- Robinson, P. (1996). Learning simple and complex second language rules under implicit, incidental, rule-search, and instructed conditions. *Studies in Second Language Acquisition*, 18, 27–67.
- Rosa, E., & O'Neil, M. (1999). Explicitness, intake, and the issue of awareness: Another piece to the puzzle. *Studies in Second Language Acquisition*, 21, 511–556.
- Schmidt, R. (1990). The role of consciousness in second language learning. *Applied Linguistics*, 11, 129–158.
- Sharwood Smith, M. (1993). Input enhancement in instructed SLA: Theoretical bases. *Studies in Second Language Acquisition*, 15, 165–180.
- Skehan, P. (1996). A framework for the implementation of task-based instruction. *Applied Linguistics*, 17, 38–62.
- Skehan, P. (1998). *A cognitive approach to language learning*. Oxford: Oxford University Press.
- Sorace, A. (1985). Metalinguistic knowledge and language use in acquisition-poor environments. *Applied Linguistics*, 6, 239–254.
- Spada, N., & Lightbown, P. (1993). Instruction and the development of questions in L2 classrooms. *Studies in Second Language Acquisition*, 15, 205–224.
- 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). Cambridge, MA: Newbury House.
- Swain, M. (1995). Three functions of output in second language learning. In G. Cook & B. Seidlhofer (Eds.), *Principle & practice in applied linguistics: Studies in honor of H. G. Widdowson* (pp. 125–144). Oxford: Oxford University Press.
- Swain, M. (1998). Focus on form through conscious reflection. In C. Doughty & J. Williams (Eds.), *Focus on form in classroom second language acquisition* (pp. 64–81). New York: Cambridge University Press.
- Swain, M., & Lapkin, S. (1995). Problems in output and the cognitive processes they generate: A step towards second language learning. *Applied Linguistics*, 16, 371–391.

- Takashima, H. (1995). *A study of focused feedback, or output enhancement, in promoting accuracy in communicative activities*. Unpublished doctoral dissertation, Temple University, Philadelphia.
- Tarone, E., & Parrish, B. (1988). Task-related variation in interlanguage: The case of articles. *Language Learning*, 38, 21–44.
- Terrell, T. (1991). The role of grammar instruction in a communicative approach. *The Modern Language Journal*, 75, 52–63.
- Tomasello, M., & Herron, C. (1989). Feedback for language transfer errors: The garden path technique. *Studies in Second Language Acquisition*, 11, 385–395.
- White, L., Spada, N., Lightbown, P., & Ranta, L. (1991). Input enhancement and L2 question formation. *Applied Linguistics*, 12, 416–432.
- Zobl, H. (1983). Markedness and the projection problem. *Language Learning*, 33, 292–313.

Appendix A: Example of Pictures Used in the Oral Picture Description Task

1



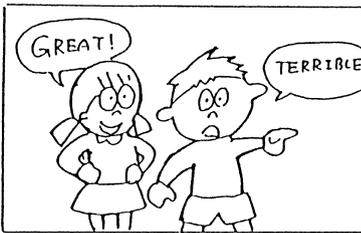
old man / paint / picture / park

2



boy / girl / come / look at / picture

3



boy / say / terrible
girl / say / great

4



old man / hit / boy / bat

Appendix B: Grammaticality Judgment Task (Post-test 1)

(Direction)

Please read each sentence and decide whether the underlined parts of each sentence are grammatical or ungrammatical. Correct all incorrect, ungrammatical parts. If you find correct, grammatical parts, circle them. (All words are spelled correctly.)

次の文を読み、各文の下線部が正しいかどうか判断して下さい。正しくないと思うものは、全て訂正して下さい。正しいと思うものは下線部を丸で囲んで下さい。(スペルの誤りはありません。)

Example: Why are you looking ^{at} me? -- wrong

Why are you (looking for) John? -- correct

1. I went to the French restaurant¹ in Tokyo. A restaurant² is famous for its seafood³.
2. There was very kind doctor⁴ in my home town.
3. John is the student⁵ of biology⁶ at Kyoto University⁷.
4. A young woman⁸ and a tall man⁹ were talking¹⁰ outside my house. I think a young woman¹¹ was Chinese and tall man¹² was American.
5. I saw a very interesting movie¹³ last night. The name¹⁴ of movie¹⁵ is *Jurassic Park*.
6. When you turn onto Pine Avenue, you will see two houses¹⁶: a blue one¹⁷ and a yellow one¹⁸. I live in a blue house¹⁹.
7. Yesterday I see²⁰ police officer²¹ chasing your dog.
8. There may be a more direct route²².
9. Yoko is studying²³ Spanish very hard²⁴.
10. Women²⁵ are sometimes the better learners of English²⁶ than men²⁷.
11. Steve found²⁸ volunteer job²⁹ with an agency³⁰. An agency³¹ was a business³² that helped other businesses.
12. If you want to buy a new car³³, consider buy³⁴ a small one³⁵. Small car³⁶ costs you less.
13. There is the Japanese student³⁷ in financial trouble.
14. This morning I read³⁸ a magazine³⁹ and a newspaper⁴⁰, but I don't know where a newspaper⁴¹ is.
15. Do you think brothers⁴² of Elvis Priestly are still alive⁴³?
16. Mary has lived⁴⁴ in Japan all her life⁴⁵.

Appendix C: Examples of Scenarios Used in Strategic Interaction

Scenario 1

Role A (employee at a trading company): Your boss tells you that your department has ¥1,000,000 which can be spent for any purpose for your department. You think that there are many things that must be replaced immediately in your office. Look carefully at the following picture and tell your boss what you think your department needs the most.

ロールA（貿易会社の社員）：あなたの上司があなたの部（職場）でどのような用途に使ってもいい予算が百万円余っているとっています。あなた自身は職場のなかで買い換えるべきものがたくさんあると思っています。次の絵を見て上司に何が一番必要か伝えてください。



Role B (department-chief of a trading company): Your department has ¥1,000,000 which can be spent for any purpose for your department. You think providing the employees with a recreation trip to a hot spring might be a good way to spend the money (There are 20 employees in your department. You can go to any famous hot spring spot by air and stay in a first-class Japanese inn with the money). Ask someone in your department what s/he thinks of this idea and decide how to spend the money.

ロールB（貿易会社の部長）：あなたの部にどのような用途に使ってもいい予算が百万円あります。あなたは社員旅行で温泉にでも行ってそのお金を使うのがいいと思っています（あなたの部には20人の社員がいます。百万円あればどの温泉へでも飛行機で行って一流の旅館に泊まります）。社員のだれかにこのアイデアについてどう思うか聞いて、使い道を決めてください。