SSLA, **27**, 79–103. Printed in the United States of America.
DOI: 10.1017/S0272263105050047

IDENTIFYING THE IMPACT OF NEGATIVE FEEDBACK AND LEARNERS' RESPONSES ON ESL QUESTION DEVELOPMENT

Kim McDonough
University of Illinois at Urbana-Champaign

Swain's (1985, 1995, 2000) output hypothesis states that language production is facilitative of second language (L2) learning. An important component of the output hypothesis involves *pushing* learners to produce appropriate, accurate, and complex language (Swain, 1993), which may occur when interlocutors provide learners with negative feedback (Gass, 1997, 2003; Long, 1996; Mackey, in press; Pica, 1994; Swain & Lapkin, 1995). When learners modify their previous utterances in response to negative feedback, learning opportunities are created by both the provision of negative feedback and the production of modified output. Consequently, it is difficult to determine how these interactional features—alone or in combination—positively impact L2 development. The current study examines the impact of negative feedback and learners' responses on English as a second language (ESL) question development, which is operationalized as stage advancement in Pienemann and Johnston's developmental sequence for ESL question formation (Pienemann & Johnston, 1987; Pienemann, Johnston, & Brindley, 1988). Thai English as a foreign language (EFL) learners ($n = 60$) carried out a series of communicative tasks with native English speakers in four conditions that provided different negative feedback and modified output opportunities and also completed four oral production tests over an 8-week period. Analysis of the treatment data identified the amount of modified output involving developmentally advanced question forms produced by the learners, and analysis of the test data revealed whether the learners' stage

I am grateful to Alison Mackey for her insightful comments on this paper and on the dissertation research on which it is based. I also thank Rhonda Oliver, Jeff Connor-Linton, Jennifer Leeman, Jenefer Philp, Ana-Maria Nuevo, and the anonymous *SSLA* reviewers for their valuable comments. Any errors, of course, are my own.

Address correspondence to: Kim McDonough, University of Illinois, Urbana-Champaign, 3070 Foreign Languages Building, MC-172, 707 S. Mathews Avenue, Urbana, IL 61801; e-mail: mcdonokr@uiuc.edu.

© 2005 Cambridge University Press 0272-2631/05 $12.00

assignment changed over time. Logistic regression indicated that the only significant predictor of ESL question development was the production of modified output involving developmentally advanced question forms in response to negative feedback.

The output hypothesis emerged in the mid 1980s from Swain's (1985) observation that learners in Canadian immersion classrooms failed to achieve nativelike proficiency for some aspects of language despite considerable exposure to comprehensible input. Swain argued that the comprehensible input provided to immersion learners might not be sufficient "to ensure that the outcome will be native-like performance" (p. 236). She speculated that the immersion learners did not achieve nativelike grammatical competence because they had few opportunities to produce the target language. These observations led Swain to formulate the output hypothesis, which states that "output that extends the linguistic repertoire of the learner as he or she attempts to create precisely and appropriately the meaning desired" is facilitative of second language (L2) learning (p. 252).

Swain (1993) later argued that an important component of the output hypothesis is *pushed* output. Speaking or writing alone may not be sufficient to facilitate some aspects of L2 learning because it is possible to successfully convey meaning despite the use of ungrammatical or pragmatically inappropriate forms. Swain suggested that learners "need to be pushed to make use of their resources; they need to have their linguistic abilities stretched to their fullest; they need to reflect on their output and consider ways of modifying it to enhance comprehensibility, appropriateness, and accuracy" (p. 160). Learners may be pushed to stretch their linguistic resources in a variety of contexts, such as during collaborative language production tasks (Fotos, 1993; Kowal & Swain, 1994; Storch, 2001; Swain, 1998), tasks that include planning time (Foster & Skehan, 1996) or posttask activities (Skehan & Foster, 1997), and interaction with negative feedback (Linnell, 1995; Nobuyoshi & Ellis, 1993).

Proponents of the interaction hypothesis of L2 acquisition (Gass, 2003; Long, 1996; Mackey, in press; Pica, 1994) have argued that interaction that pushes learners to modify their output in response to an interlocutor's negative feedback may facilitate L2 development. This type of interaction brings together input features (e.g., negative feedback), internal learner capacities (e.g., attention), and language output. Empirical studies have demonstrated the positive effect of interaction on L2 development for several linguistic forms, including adverb placement in Spanish (Long, Inagaki, & Ortega, 1998), noun-adjective gender agreement in Spanish (Leeman, 2003), past tense in English (Han, 2002; Mackey, 2000), past tense and subjunctive in French (Ayoun, 2001), and verbal morphology in Japanese (Iwashita, 2003). Several studies carried out within this framework (Mackey, 1997, 1999, 2000; Mackey & Oliver, 2002; Mackey & Philp, 1998; Silver, 2000) have investigated the effect of interaction on English

as a second language (ESL) question development, which was operationalized as movement to a higher stage in Pienemann and Johnston's (1987; Pienemann, Johnston, & Brindley, 1988) developmental sequence for ESL question formation (see Appendix A). As summarized by Mackey (1999), defining development very narrowly in terms of stage movement in this sequence has been motivated by the following reasons: (a) questions are complex structures that are readily elicited and likely to be affected by interaction, (b) different question forms are present at all stages of learning, (c) empirical support for the developmental stages is relatively robust, and (d) developmental stages allow researchers to assess and control learners' readiness to acquire certain forms.

In summary, empirical research within the interaction hypothesis framework has demonstrated that interaction that pushes learners to stretch their linguistic resources through negative feedback and opportunities to modify their output in response to feedback may facilitate L2 development of some linguistic forms. Additionally, several studies have shown positive effects for ESL question development. However, these studies were not designed to explore the individual and combined effects of these interactional features. The findings indicated that the combination of negative feedback and subsequent modified output facilitated ESL question development but did not determine whether each interactional feature was predictive of the learners' advancement to a higher stage. The current study builds on this research by exploring whether the negative feedback or the resulting modified output—or both—are predictive of ESL question development. The following sections summarize the arguments that have been advanced to support the claim that these interactional features are beneficial for L2 development.

THE CONTRIBUTION OF NEGATIVE FEEDBACK

Negative feedback through interaction may contribute to L2 development by informing learners about the comprehensibility of their utterances (Long, 1996; Schachter, 1986, 1991) and by raising their awareness of language (Ellis, 1991). Negative feedback, which may occur with varying degrees of explicitness or implicitness, may draw learners' attention to the language forms they have produced and help them to detect gaps or holes in their L2 knowledge or to notice specific linguistic forms in the subsequent input (Gass, 1997, 2003; Long; Pica, 1994; Schmidt, 1995, 2001; Swain & Lapkin, 1995). Some types of implicit negative feedback—such as recasts—may be beneficial because they provide positive evidence that is salient to learners (Leeman, 2003). In general, negative feedback may serve L2 development by encouraging learners to attend to features of the input that otherwise may have remained undetected (Long).

One important question is whether learners recognize that negative feedback provides information about the acceptability of their language (Birdsong, 1989; Carroll, 1995; Leeman, 2003). An interlocutor provides negative feedback for a variety of reasons, such as failure to hear what the learner said or

failure to understand the meaning of what a learner said. If learners recognize that negative feedback contains information about the acceptability of their utterance, they must identify which aspect of that utterance was unacceptable. This has been discussed in the literature on negative evidence as the problem of blame assignment (Carroll; Pinker, 1989), which occurs when negative feedback does not pinpoint the problem. Empirical studies that have investigated learners' perceptions of implicit negative feedback (Mackey, Gass, & McDonough, 2000; Morris & Tarone, 2003; Roberts, 1995) have suggested that learners are unlikely to perceive some types of implicit negative feedback—such as recasts that reformulate learners' nontargetlike morphosyntax—as negative feedback. Nevertheless, it is possible that learners benefit from negative feedback even if they do not perceive the linguistic target.

THE CONTRIBUTION OF MODIFIED OUTPUT

The potential contribution of modified output to L2 development can be understood through reference to Levelt's model of speech production (Bock, 1995; Bock & Levelt, 1994; Levelt, 1989), which accounts for the generation of fluent speech through processing in three components of speech production: the message, the grammatical, and the phonological components. The aspect of this model particularly relevant for understanding the potential role of modified output in ESL question development is the monitor, which oversees the entire speech production system. The monitor initiates self-repair prior to articulation or redirects speech that has already been produced. When learners recognize that an utterance is deficient, they have two general options for repair. First, they can send the original output of the message component back to the grammatical component for new functional or positional processing, or both. Second, they can create a different or supplemental message in the message component and send the new message to the grammatical component. Thus, when learners modify their output, they either generate a new message or reprocess their original message, both of which trigger additional grammatical encoding (Izumi, 2003).

When learners produce a particular structure, they may repeatedly use that same structure in subsequent utterances, a phenomenon Mackey (1999) referred to as *clustering*. If learners produce a more complex or accurate form in their modified output, they may be more likely to produce that form in their subsequent utterances. An example of subsequent production of a modified form is illustrated in (1) (taken from the current study). In this task, the learner produced a stage 3 question while identifying the differences between two versions of a scene depicting astronauts and aliens on another planet. In response to the learner's stage 3 question, the interlocutor requested clarification (*huh?*). The learner modified her output by inserting an auxiliary verb (*are*), which resulted in a stage 5 question. After reformulating her utterance in this way, she then produced several more stage 5 questions that included the same auxiliary verb.

(1)	Learner:	*what two guys doing?*	← Stage 3 question
	Native speaker (NS):	*huh?*	← Clarification request
	Learner:	*what are two guys doing?*	← Stage 5 question
	NS:	*signing a contract*	
	Learner:	*what contract?*	
	NS:	*they want to start a business together*	
	Learner:	*when are they going to do the business?*	← Stage 5 question
	NS:	*uh in a couple days*	
	Learner:	*is he happy now because he look a little bit bored*	
	NS:	*I think he's happy I think he's a bit worried*	
	Learner:	*uh and what are they going to use the gun for?*	←Stage 5 question

In sum, producing modified output may contribute to L2 development by strengthening knowledge representations that learners already have stored (Nobuyoshi & Ellis, 1993) and by encouraging automatic retrieval of linguistic forms (de Bot, 1996).

PURPOSE OF THE STUDY

As outlined in the previous sections, researchers have advanced arguments for the beneficial role of negative feedback and modified output in L2 development. Previous empirical studies have shown that interaction that provides negative feedback and opportunities for learners to modify their output in response to feedback has facilitated L2 development for several linguistic forms, including ESL question development. However, the frequent co-occurrence of negative feedback and modified output makes it difficult to determine whether these interactional features positively contribute to development alone or in combination. The purpose of the present study is to build on previous research by identifying whether both negative feedback and learners' responses to that feedback are predictive of ESL question development. The following research question was formulated: Are negative feedback and modified output produced in response to negative feedback significant predictors of ESL question development? Given the claim that negative feedback and modified output play a positive role in L2 development, the expectation was that both interactional features would be predictive of ESL question development.

METHOD

Participants

English as a Foreign Language (EFL) Learners. The EFL participants were 109 students enrolled in the English department at a large public university in northern Thailand. Forty learners were excluded from the analysis because they had missed treatments or tests ($n = 18$) or had participated in a treatment con-

dition not reported here (n = 22).[1] Additionally, only learners who were classified at the same developmental stage in Pienemann and Johnston's developmental sequence for ESL question formation (1987; Pienemann et al., 1988) on the basis of their pretest performance were included in the study. Because the majority of the learners were classified as stage 4, those at stage 3 or stage 5 (n = 9) were removed from the analysis. The resulting participant pool consisted of 60 EFL learners, 50 women and 10 men, who were all NSs of Thai. Their ages ranged from 17 to 21 years, with an average of 18.5 years. Their amount of previous English study ranged from 8 to 14 years, with an average of 9.8 years, but few learners had ever lived in a country where English was spoken as a native language or used English as a medium for communication while traveling in other Asian countries. They reported infrequent use of English outside class and limited exposure to English through mass media, such as movies and the Internet. Information about the learners' general proficiency level was not available because they had never taken standardized tests—such as the Test of English as a Foreign Language (TOEFL) or the Test of English for International Communication (TOEIC)—or any university placement tests. Consequently, to ensure that the participants represented the same population, only learners with the same initial stage assignment for question formation (stage 4) were included in the study.

NS Interactors. The NS interactors, four women and one man, were NSs of American and British English who were working as lecturers in the English department at the same university as the EFL learners. None of the EFL learners were enrolled in courses taught by the NS interactors in the semester that data was collected. The NS interactors were trained to provide negative feedback while carrying out the treatment tasks with the EFL learners. The training consisted of the following: (a) discussing the purpose and background of the research project, (b) reading a written description of each treatment task, (c) listening to audiotapes and reading transcripts of the tasks being carried out by a native English speaker and Thai students, and (d) piloting the treatment conditions while carrying out tasks with learners from a similar population. The NS interactors were neither provided with scripts nor instructed to strictly direct the conversation with each participant. Instead, they were asked to participate in collaborative dialogue with the learners and to respond to the learners' nontargetlike question forms as needed and when contextually appropriate.

Design

The study employed a pretest-posttest design to identify the impact of negative feedback and modified output produced in response to negative feedback on EFL learners' question development. One of the independent variables under investigation—modified output—was learner generated. As a result, it was not possible to determine a priori which learners would produce modified output or to isolate modified output as an independent variable in the

treatment conditions. Instead, the treatment conditions manipulated variables that affected the learners' opportunities to produce modified output in response to negative feedback.

The first variable manipulated in the treatment conditions was implicit negative feedback in the form of open-ended clarification requests (e.g., *pardon?*, *what?*, or *huh?*). Clarification requests were selected because previous research has indicated that they create more opportunities for learners to produce modified output than forms of negative feedback that provide a reformulation, such as recasts (Anton, 1999; Linnell, 1995; Lyster & Ranta, 1997; Pica, 1988; Pica, Holliday, Lewis, & Morgenthaler, 1989). The second variable manipulated in the treatment conditions was the salience of the learners' nontargetlike forms. As described previously, learners may have difficulty recognizing that implicit negative feedback contains information about the grammaticality of their utterances as well as identifying the linguistic target of feedback. If they do not perceive the linguistic target of the negative feedback, they may be unlikely to modify the relevant forms. Therefore, the salience of the learners' nontargetlike questions was enhanced through repetition. Repetition was used to enhance the salience of the learners' nontargetlike questions because researchers have suggested that it is an effective technique for drawing learners' attention to specific linguistic features (Chaudron, 1986; Doughty & Varela, 1998; Lyster & Ranta; Panova & Lyster, 2002; Samuda, 2001; Sharwood Smith, 1991). The four treatment conditions, which are described in the following section, represented different combinations of these two variables in order to increase the likelihood that learners would produce varying levels of modified output in response to negative feedback. The 60 participants were equally divided among the four treatment conditions through random assignment.

Enhanced Opportunity to Modify (n = 15). To draw the learners' attention to the problematic feature of their previous utterance, the NS interactors responded to learners' nontargetlike questions by enhancing the salience of those forms through repetition with stress and rising intonation. Immediately following the repetition, the NS interactors requested clarification by way of an open-ended clarification request (e.g., *sorry?* or *what?*) and paused to allow the learners an opportunity to modify their output. The enhanced opportunity to modify treatment condition is illustrated in (2).

(2) Learner: *what angel doing in this situation?*
NS: *what angel doing? Huh?*
Learner: *what is angel doing?*

Opportunity to Modify (n = 15). The NS interactors responded to learners' nontargetlike questions by requesting clarification using an open-ended clarification request (e.g., *pardon?* or *huh?*). They did not provide any information to help the learners identify the problematic features of their utterances. They paused after the clarification requests so that the learners had opportunities to modify their output. The opportunity to modify treatment condition is illustrated in (3).

(3) Learner: *what happen for the boat?*
NS: *what?*
Learner: *what's wrong with the boat?*

Feedback without Opportunity to Modify (n = 15). The NS interactors responded to learners' nontargetlike questions by enhancing the salience of those forms through stress and rising intonation, thereby highlighting the problematic features of their utterances. However, immediately following the repetition, the NS interactors continued talking so that the learners did not have an opportunity to acknowledge the feedback or modify their output.[2] This is illustrated in (4).

(4) Learner: *what we do with it?*
NS: *what we do? Uh let's see well we could talk about the purpose if you want*

No Feedback (n = 15). The NS interactors did not provide any implicit negative feedback when the learners produced nontargetlike questions. In the event that a breakdown in the communication of meaning occurred, the NS interactors were instructed to feign understanding. The no feedback condition is illustrated in (5).

(5) Learner: *where you going the last holiday?*
NS: *to Laos*

Materials

Treatment Materials. The treatment materials consisted of information-exchange and information-gap activities adapted from commercial textbooks and resource books to elicit a variety of question types. Three sets of treatment materials were created; each set consisted of two communicative activities that elicited questions (see Appendix B). During each treatment session, the learners were given a learning journal, a blank form with spaces for them to write any comments about what they were learning during the interaction (see Appendix C). The form provided space for learners to write down items in three categories (pronunciation, vocabulary, and grammar) and had another space for any other aspects of language or communication. The learning journals, modeled on a format preferred by L2 learners in a prior study (Mackey, McDonough, Fujii, & Tatsumi, 2001), were used as an indication of whether the learners reported any attention to question forms during the treatment sessions.

Testing Materials. The testing materials were communicative tasks that the learners carried out individually in a language laboratory (see Appendix D). Each test contained a warm-up activity (questions about the learners' recent activities) and two activities that elicited questions (story completion and brainstorming interview questions). The oral production tests differed from the treatment tasks to avoid problems associated with the use of identical treatment and testing materials (Krashen, 1998; Truscott, 1999). The use of

different treatment and testing materials sets a more conservative standard for evidence of development because it requires learners to generalize any changes that occurred as a result of the treatment tasks to the new contexts presented through the testing tasks.

A series of pilot studies was conducted with Thai EFL learners (n = 35) from a comparable population to assess the effectiveness of the treatment and test materials. The treatment tasks and oral production tests successfully elicited a variety of question types, and the learning journals successfully elicited comments in all four categories.

Procedure

The learners participated in three treatment sessions and completed four oral production tests over an 8-week period. Pretests were completed during week 1, and posttests were completed in weeks 2, 5, and 8. The oral production tests were administered in a language laboratory using a tape-mediated format (a prerecorded audiotape that gave instructions and controlled the amount of time per activity) while learners were seated at individual carrels equipped with boom microphones. The tests took 20 minutes to complete. The learners completed questionnaires during weeks 2 and 9. The questionnaires were used to explore whether the learners had identified the purpose of the research study. A variety of distracter items were included to minimize the possibility that the questionnaires would draw the learners' attention to question forms.

Learners were randomly assigned to the treatment conditions, in which they carried out communicative tasks with a NS interactor in three sessions. Each treatment session provided approximately 10 minutes of interaction for activities that elicited questions. All the treatment activities elicited an equivalent number of possible contexts for the target forms. During the treatment sessions, the NSs provided implicit negative feedback to learners in the feedback conditions in response to nontargetlike questions. Given that excessive feedback may lead to learner irritation (Aston, 1986), the NS interactors were instructed not to provide feedback in response to every nontargetlike question form. Additionally, the NS interactors were instructed to continue with the task if the learners who were given opportunities to modify their output following feedback did not modify their question forms. Finally, all learners completed the learning journals either while carrying out the activities, or at the end of each session, or both.

Analysis

Treatment Task Performance. The audiotapes of the interaction between the learners and the NS interactors during the treatment tasks were transcribed by the researcher and a paid research assistant. The interaction was examined for evidence of modified output, operationalized narrowly as learners' responses to an interlocutor's negative feedback involving a reformulation of a question form. Complete and partial repetitions of the learners'

original utterances were not coded as modified output because they did not involve reformulation. Each instance of modified output was classified according to the developmental stage represented in the reformulation. Because the learners began the study at stage 4, only modified output involving stage 5 questions was considered as having potential to drive question development. An example of modified output with a stage 5 question is provided in (6).

(6) Learner: *what Mark put in the glass?* ← Stage 3 question
NS: *what Mark put? Sorry?*
Learner: *what is Mark put in the glass?* ← Stage 5 question

The total amount of modified output involving stage 5 questions produced by each learner was calculated. An independent scorer coded 20% of the treatment data for the occurrence of modified output involving question forms and their stage assignments. Agreement between the coding of the researcher and the independent scorer was 98%.

The learning journals were analyzed for any evidence that the learners had attended to question forms during the treatment tasks. Evidence of attention to question forms was operationalized as any written comment about questions. A binary distinction was made that separated the learners who provided commentary about questions in their journals from those who did not. The types of comments that were classified as showing evidence of attention to questions included *I learn how to ask questions* and *I can ask the question in English*. An independent scorer coded all of the learning journal data, and agreement with the researcher was 100%.

Oral Production Tests. The audiotapes of the oral production tests were transcribed by the researcher, and questions were coded for stages in the developmental sequence for ESL question formation. The following types of questions were removed from the data: (a) incomplete questions, such as *how about Phuket?* and *what time?*; (b) echo questions; (c) multiple exemplars of the same question on the same task; and (d) formulaic chunks, such as *where do you come from?* and *do you like* + object? Following Mackey (1999, 2000), the current study also required the presence of at least two different higher level questions in different tasks. However, the current study operationalized development more conservatively by requiring the presence of two questions from a higher stage on all three posttests. A more conservative measure of development was used to reduce the possibility of Type 1 error. The independent scorer coded 20% of the test data for question stage assignment, and agreement with the researcher was 96%. Alpha was set at .05 for all statistical tests.

RESULTS

Treatment Task Data

Table 1 summarizes of the amount of negative feedback, the modified output involving stage 5 questions, and the reported attention to question forms for

Table 1. Treatment task performance by group

Treatment group	Negative feedback			Modified output[a]			Attention[b]	
	Sum	Mdn	IQR	Sum	Mdn	IQR	Yes	No
Enhanced opportunity to modify	99	5.00	4.00	20	2.00	2.00	3	12
Opportunity to modify	93	5.00	5.00	12	.00	1.00	3	12
Feedback without opportunity to modify	72	4.00	2.00	0	0	0	0	15
No feedback	0	0	0	0	0	0	3	12

Note. In each treatment group, $n = 15$.
[a]This includes only modified output involving stage 5 questions.
[b]This includes any comments about questions noted in the learning journals.

each treatment group. Learners in the enhanced opportunity to modify group received the most negative feedback (sum = 99), followed by the opportunity to modify group (sum = 93), and the feedback without opportunity to modify group (sum = 72). As expected, learners in the no feedback group did not receive any negative feedback. A Kruskal-Wallis test (a nonparametric ANOVA)[3] indicated that the difference in the amount of negative feedback provided to the three groups that received feedback was not significant, $H(2, 45) = 2.42$, $p = .30$. Learners in the enhanced opportunity to modify group produced more modified output involving stage 5 questions (sum = 20) than the learners in the opportunity to modify group (sum = 12). However, a Mann-Whitney test (a nonparametric *t*-test) indicated that the difference in the amount of modified output produced by the two opportunity groups was not significant, $Z(1, 30) = -1.63$, $p = .14$. As expected, learners in the no opportunity to modify group and the no feedback group did not produce any modified output involving stage 5 questions in response to negative feedback. Finally, an examination of the learning journals revealed that nine learners reported attention to question forms, with three learners each in the enhanced opportunity to modify, opportunity to modify, and no feedback groups. No learners in the feedback without opportunity to modify group commented on question forms in their journals.

Test Data

The number of learners in each treatment group who advanced from stage 4 to stage 5 in the developmental sequence for ESL questions was calculated. As shown in Table 2, 9 of 15 learners in the enhanced opportunity to modify group advanced to stage 5. Five out of 15 learners in the opportunity to modify group advanced to stage 5. In the remaining two groups, only 2 of 15 learners advanced to stage 5. Although it is common in the literature on SLA to use chi-square or Fisher's exact tests to assess the relationship between treatment group membership and developmental outcomes, these tests have two

Table 2. Question development by group

Stages	Enhanced opportunity to modify	Opportunity to modify	Feedback without opportunity to modify	No feedback
4 → 5	9	5	2	2
No change	6	10	13	13

Note. In each treatment group, $n = 15$.

important shortcomings: first, they cannot identify which independent variables best predict the observed distribution, and, second, they cannot provide information about possible interactions between independent variables (Hatch & Lazaraton, 1991; Saito, 1999). Additionally, group frequency counts can obscure important individual differences, such as the amount of modified output produced by each learner. Consequently, logistic regression was used to identify whether negative feedback, modified output, or both were predictive of ESL question development.

The goal of logistic regression is to find the most appropriate model to describe the relationship between an outcome, a dependent variable, and a set of predictors that act as the independent variables (Hosmer & Lemeshow, 1989). Unlike linear regression, which requires a numeric dependent variable, logistic regression is appropriate when the dependent variable is categorical. The independent variables included in a logistic regression model are selected on the basis of (a) theoretically motivated reasons to predict a relationship between the independent and dependent variable, and (b) univariate statistical analysis of each independent variable to assess the strength of its relationship with the dependent variable (Bernard, 1995; Cox & Snell, 1989; Hatch & Lazaraton, 1991; Hosmer & Lemeshow; Pampel, 2000; Young & Yandell, 1999). The independent variables initially considered for the model were (a) the interactional features manipulated in the treatment conditions, which included clarification requests and enhanced salience of nontargetlike forms; (b) the amount of modified output involving stage 5 questions produced in response to negative feedback; and (c) learners' reported attention to question forms during the treatment tasks. Pearson correlation coefficients for the dependent and independent variables were calculated to determine the strength of relationships. As shown in Table 3, ESL question development was significantly correlated with modified output and clarification requests only.

On the basis of the results of the univariate statistical analysis, modified output and clarification requests were included in the logistic regression model using an enter selection method. The results of the logistic regression indicated that the model was significant, $\chi^2(2, 60) = 30.78$, $p < .05$. Goodness of fit measures showed that the model explained 57% of the pseudovariance, Nagelkerke $R^2 = .567$, which estimates the variance in the dependent variable

Table 3. Correlation coefficients

Variables	ESL question development	Modified output	Clarification request	Enhanced NTL forms	Attention reported
ESL question development	1.00	.67*	.36*	.15	−.07
Modified output	.67*	1.00	.51*	.13	.14
Clarification request	.36*	.51*	1.00	.00	.14
Enhanced NTL forms	.15	.13	.00	1.00	−.14
Attention reported	−.07	.14	.14	−.14	1.00

*$p < .05$.

accounted for by the model. The model successfully predicted 88% of group membership (i.e., developed or did not develop), which is more accurate than simply assigning membership based on the largest category of the dependent variable (70%). The independent variables in the model were analyzed to determine the strength of their relationship to ESL question development. Tests of significance showed that modified output was the only significant predictor of ESL question development, $\beta(1, 60) = 2.38$, $p < .05$, and that the effect of modified output on the odds of development was $(\text{Exp})\beta = 10.79$. In other words, increasing the production of modified output increased the odds of development by a multiple of 10.79. In summary, although I expected that both negative feedback and modified output would be significant predictors of ESL question development, the logistic regression model revealed that only modified output produced in response to negative feedback was significant.

DISCUSSION

The logistic regression model identified modified output involving developmentally advanced questions as the only significant predictor of ESL question development. Although the negative feedback variables were not predictive of ESL question development, clarification requests were positively correlated with both ESL question development and modified output. This suggests that clarification requests play an indirect rather than direct role in ESL question development by facilitating the production of modified output. An examination of data from the learners in the enhanced opportunity to modify group illustrates the relationships among clarification requests, modified output, and ESL question development. As shown in (7), one subject, Sasithorn,[4] produced primarily stage 3 and 4 questions on the pretest but produced several stage 5 questions on the posttests.

(7) a. Pretest: Stage 3 and 4 questions
Do you like your reading?
How your story is interesting?
What is the thing that he carry?
Is he drink wine?

b. Posttests: Stage 5 questions
Why do you want to give a lecture?
Who is he want to meet?
What would you like to eat?
What are they waiting for?

When Sasithorn produced stage 3 questions during the treatment sessions, shown in (8a–c), the NS interlocutors repeated her nontargetlike forms and requested clarification. Sasithorn responded to these clarification requests by inserting auxiliary verbs, which resulted in stage 5 questions. For example, in the first treatment session (8a), she produced the stage 3 question *where it have a good view?* When the NS interlocutor requested clarification (*sorry?*), Sasithorn modified her output by inserting the auxiliary verb *does*.

(8) a. Treatment 1
Sasithorn: *where it have a good view?*
NS: *where it have? Sorry?*
Sasithorn: *where does it have a good view?* ← Stage 5 question
b. Treatment 2
Sasithorn: *why she made an injection?*
NS: *why she made? Huh?*
Sasithorn: *why did she make an injection?* ← Stage 5 question
c. Treatment 3
Sasithorn: *when you work in Australia?*
NS: *when you work? Sorry?*
Sasithorn: *how long have you worked there?* ← Stage 5 question

In the second treatment session (8b), Sasithorn again modified a stage 3 question (*why she made an injection?*) by supplying the auxiliary *did*. Similarly, during the third treatment session (8c), Sasithorn produced the stage 3 question *when you work in Australia?* When asked to clarify, she modified her output, producing the stage 5 question *how long have you worked there?* Sasithorn's production of modified output involving stage 5 questions during the treatment sessions was predictive of her advancement to stage 5, and the NS interactors' clarification requests gave her opportunities for reformulation.

In contrast, Jindarat, also in the enhanced opportunity group, produced modified output involving stage 3 and 4 questions during the treatment sessions but did not advance to stage 5 on the posttests. During the treatment sessions, Jindarat predominantly employed three strategies in response to the NS interactor's feedback: she deleted auxiliary verbs (9a), changed contracted auxiliary forms to the uncontracted forms (9b), and substituted different main verbs (9c). Importantly, none of these reformulations involved stage 5 questions.

(9) a. Treatment 1
Jindarat: *where are she come from?*
NS: *where are? what?*
Jindarat: *where she come from?* ← Stage 3 question
b. Treatment 2
Jindarat: *where's he's going?*
NS: *where's he's? what?*
Jindarat: *where is he's going to?* ← Stage 3 question
c. Treatment 3
Jindarat: *has it have four legs?*
NS: *has it have? huh?*
Jindarat: *has it got four legs?* ← Stage 4 question

Jindarat continued producing stage 3 and 4 questions on the posttests, many of which were characterized by the omission of auxiliary verbs, the absence of subject-verb inversion, or the presence of a copied auxiliary verb. Although Jindarat had opportunities to respond to the clarification requests, she did not modify her output in a way that resulted in stage 5 questions and did not advance to stage 5.

As the preceding examples illustrated, several choices are available to learners when they respond to clarification requests. Some learners, like Sasithorn, responded to clarification requests by reformulating their previous utterances in ways that resulted in more developmentally advanced questions, a strategy which was predictive of development. However, other learners did not reformulate any of the nontargetlike features of their question forms in response to clarification requests. Instead, they repeated their original utterances, substituted different lexical items, or reformulated other nontargetlike features of their utterances. In (10), Karn employed the third response pattern mentioned. Karn asked the question *what price of apple in pack?* while carrying out a treatment task, but he did not respond to the interlocutor's clarification request by modifying the nontargetlike aspects of his question. Instead, he modified his utterance by adding a plural morpheme (*apples*) and expanding a prepositional phrase (*in pack*) into a relative clause (*you put in pack*).

(10) Karn: *what price of apple in pack?*
NS: *sorry?*
Karn: *what price of apples you put in pack?*

Swain (2000) has pointed out that "it seems essential in research to test what learners *actually do*, not what the researcher assumes instructions and task demands will lead learners to focus on" (p. 80, emphasis in original). The current study addressed this issue by narrowly operationalizing development in terms of a specific linguistic form (i.e., questions) and by differentiating among learners based on whether they produced modified output involving that form. However, this approach necessarily ignores learners like Karn, who chose to modify other aspects of their utterances rather than question forms. An alternative approach is to identify the linguistic features that each learner modifies while carrying out the treatment tasks and creating learner-specific posttests that elicit those forms (Swain, 1998, 2000; Swain & Lapkin, 1998). For example, learner-specific posttests could be used to explore whether Karn's production of modified output involving plurals and relative clauses had an impact on his subsequent production of those forms. Although a case-by-case analysis was beyond the scope of this study, it could be incorporated in future studies in order to determine whether the production of modified output predicts the development of a variety of linguistic forms.

Although the current study explored the specific impact of both negative feedback and modified output produced in response to negative feedback on ESL question development, the learners in the no feedback group also participated in an interactive context that could have conceivably pushed them to

modify their output. In this study, these learners were not provided with negative feedback, so they did not have opportunities to modify their output in response to feedback. However, they were not prevented from producing self-initiated modified output while carrying out the treatment tasks. As shown in (11), Pornsiri, a learner in the no feedback group, independently modified her question forms when asking the NS interlocutor about his encounter with a circus tiger.

(11)	Pornsiri:	*who are you—uh where you—where did you was piss on by a tiger?*	← Stage 3 question
	NS:	*I was in my hometown back in America and we went to go see some animals at a circus*	
	Pornsiri:	*at the zoo?*	
	NS:	*no it wasn't the zoo it was a circus that had stopped in my town*	
	Pornsiri:	*and who uh who were—who was you with?*	← Stage 4 question
	NS:	*my family*	
	Pornsiri:	*and uh do you hurt—did you hurt?*	← Stage 3 question
	NS:	*no*	
	Pornsiri:	*you were almost attacked by a tiger?*	
	NS:	*no it didn't attack me it just peed on me*	

Even though this interactional context did not provide Pornsiri with any negative feedback, she still produced modified output involving question forms. However, her reformulations did not involve developmentally advanced question forms as defined in this study, and she did not advance to stage 5. Overall, the learners in the no feedback group rarely modified their question forms, and their modifications did not involve developmentally advanced questions. Further research is necessary to identify the interactional contexts that push learners to produce modified output in the absence of negative feedback and to determine whether self-initiated modified output is also predictive of ESL question development.

CONCLUSION

This study investigated whether negative feedback and modified output produced in response to that feedback were significant predictors of ESL question development. The findings indicated that the production of modified output involving developmentally advanced question forms was the only significant predictor. Additionally, negative feedback in the form of clarification requests may indirectly contribute to question development by creating opportunities for learners to modify their output. Thus, this study provides empirical support for the output hypothesis (Swain, 1985, 1993, 1995) and strengthens claims for an association between modified output and ESL question development (Mackey, 1997). To test hypothesized relationships between negative feedback, modified output, and L2 development, it was necessary to operationalize development very narrowly. Consequently, the findings are not generalizable

to other linguistic features. A case-by-case analysis, as described in the previous section, would allow researchers to identify the individual and combined contributions of specific interactional features to the development of a wide range of linguistic structures. This approach would also complement Swain's (2000) reconceptualization of output in terms of the role of external speech in facilitating the appropriation of strategic processes and linguistic knowledge through collaborative construction.

The current study was situated in a foreign language context in which all the learners shared the same first language (L1) as well as amount and type of prior exposure to English. Although this context strengthened the internal validity of the study, it posed threats to external validity because the findings may not be generalizable to learners in other contexts or learners from different L1 backgrounds. There is some evidence that cultural background influences the learning strategies chosen (O'Malley, Chamot, Stewner-Manzanares, Russo, & Kupper, 1985; Politzer & McGroarty, 1985; Shonerd, 1994). Furthermore, learners in a foreign language context whose primary exposure to the target language has been in formal educational settings with explicit grammar teaching may find grammatical errors more salient than L2 speakers in a naturalistic environment (Bardovi-Harlig & Dörnyei, 1998). As such, foreign language learners may allocate attentional resources and make decisions about repair differently than learners in a L2 context (Kormos, 2000). As a result, replication studies involving EFL learners from a wide variety of cultural and linguistic backgrounds as well as studies with ESL learners are needed.

Although the results suggested that modified output involving developmentally advanced forms was predictive of ESL question development, they do not warrant the conclusion that modified output is necessary for development. Learners who did not modify their question forms during the treatment sessions may eventually advance to a higher stage, which would suggest that producing modified output facilitates the rate at which learners advance through developmental stages. However, this possibility is speculative and cannot be confirmed without longitudinal or more long-term cross-sectional studies. Although evidence of question development persisted over an 8-week period in the current study, in future studies, additional delayed posttesting would offer greater understanding of the longer term persistence of the learners' gains as well as determine whether the learners who did not immediately benefit from the treatment caught up with their peers.

The inclusion of measures of individual differences may benefit future studies that investigate the relationship between modified output and L2 development. Recent research has demonstrated that individual differences in proficiency (Iwashita, 2001), L2 working memory capacities (Mackey, Philp, Egi, Fujii, & Tatsumi, 2002), and speaking style (Kormos, 1999) may affect learners' abilities to make use of the learning opportunities available through interaction. Research in these areas may help explain why some learners respond to negative feedback by modifying their nontargetlike forms, whereas others respond by repeating their original utterances. This research may also help account for variation in an individual learner's reaction to interactional fea-

tures in different situations. Additionally, more fine-grained measures of attention could be used to investigate the claim that negative feedback promotes attention to linguistic form. The current study did not identify an association between negative feedback and attention to question forms, but it was not designed to test this claim rigorously.

A more qualitative analysis of the learning journal data may have provided greater insight into the learners' allocation of attention during the collaborative tasks. Recent research has used a variety of offline and online measures to investigate the role of attention in L2 learning (e.g., stimulated recall and think-aloud protocols), and future studies could incorporate these measures to systematically test whether different types of negative feedback promote attention to form.

The current study focused exclusively on negative feedback in the form of clarification requests and enhanced salience of nontargetlike forms because the former has been shown to elicit modified output and the latter is believed to help learners detect the problematic features of their utterances. As mentioned previously, other types of negative feedback (e.g., recasts) may be less likely to push learners to modify their output. Because recasts provide learners with reformulations of problematic forms, they may not have opportunities to respond or may respond by simply acknowledging the reformulation (Mackey, Oliver, & Leeman, 2003; Oliver, 1995). Consequently, future research could identify whether modified output in response to recasts is also predictive of ESL question development or whether recasts alone predict developmental outcomes.

Clearly, many remaining issues warrant continued empirical efforts. In particular, it will be important for future research to identify which interactional features are facilitative of L2 development, to explore the cognitive processes that mediate environmental and learner internal factors, and ultimately to explain how certain environmental or internal factors play a role in facilitating or inhibiting development. The findings of this study indicate that the production of modified output in response to clarification requests is predictive of ESL question development, and the next challenge is to determine why it is beneficial.

(*Received 6 July 2004*)

NOTES

1. The design of the full study (McDonough, 2001) included a treatment condition that provided learners with negative feedback in the form of a general statement that the interlocutor had not understood. This group was excluded from the present analysis to ensure that the variables in the logistic regression model were independent.

2. Pilot tests indicated that it was often necessary for the NS interlocutors to change the topic immediately following the negative feedback to try to ensure that the learners (a) did not interpret the repetition as positive feedback and (b) did not attempt to modify their previous utterances.

3. Nonparametric statistics were used because the treatment task data was not normally distributed and the variance was not equal.

4. All names have been changed.

REFERENCES

Anton, M. (1999). The discourse of a learner-centered classroom: Sociocultural perspectives on teacher-learner interaction in the second-language classroom. *Modern Language Journal*, *83*, 303–318.

Aston, G. (1986). Trouble-shooting in interaction with learners: The more the merrier? *Applied Linguistics*, *7*, 128–143.

Ayoun, D. (2001). The role of negative and positive feedback in the second language acquisition of the passé composé and imparfait. *Modern Language Journal*, *85*, 226–243.

Bardovi-Harlig, K., & Dörnyei, Z. (1998). Do language learners recognize pragmatic violations? Pragmatic versus grammatical awareness in instructed L2 learning. *TESOL Quarterly*, *32*, 232–262.

Bernard, H. (1995). *Research methods in anthropology: Qualitative and quantitative approaches*. Walnut Creek, CA: Altamira.

Birdsong, D. (1989). *Metalinguistic performance and interlinguistic competence*. New York: Springer Verlag.

Bock, K. (1995). Sentence production: From mind to mouth. In J. Miller & P. Eimas (Eds.), *Speech, language, and communication* (pp. 181–216). San Diego, CA: Academic Press.

Bock, K., & Levelt, W. (1994). Language production: Grammatical encoding. In M. Gernsbacher (Ed.), *Handbook of psycholinguistics* (pp. 945–985). San Diego, CA: Academic Press.

Carroll, S. (1995). The irrelevance of verbal feedback to language learning. In L. Eubank, L. Selinker, & M. Sharwood Smith (Eds.), *The current state of interlanguage: Studies in honor of William E. Rutherford* (pp. 73–88). Amsterdam: Benjamins.

Chaudron, C. (1986). Teachers' priorities in correcting learners' errors in French immersion classes. In R. R. Day (Ed.), *"Talking to learn": Conversation in second language acquisition* (pp. 327–351). Rowley, MA: Newbury House.

Cox, D., & Snell, E. (1989). *Analysis of binary data*. London: Chapman & Hall.

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

Doughty, C., & Varela, E. (1998). Communicative focus on form. In C. Doughty & J. Williams (Eds.), *Focus on form in classroom SLA* (pp. 129–154). New York: Cambridge University Press.

Ellis, R. (1991). The interaction hypothesis: A critical evaluation. In E. Sadtano (Ed.), *Language acquisition and the second/foreign language classroom* (pp. 179–211). Singapore: Regional English Language Centre.

Fotos, S. (1993). Consciousness raising and noticing through focus on form: Grammar task performance versus formal instruction. *Applied Linguistics*, *4*, 385–407.

Foster, P., & Skehan, P. (1996). The influence of planning on performance in task-based learning. *Studies in Second Language Acquisition*, *18*, 299–324.

Gass, S. (1997). *Input, interaction, and the second language learner*. Mahwah, NJ: Erlbaum.

Gass, S. (2003). Input and interaction. In C. Doughty & M. H. Long (Eds.), *Handbook of second language acquisition* (pp. 224–255). Oxford: Blackwell.

Han, Z. (2002). A study of the impact of recasts on tense consistency in L2 output. *TESOL Quarterly*, *36*, 543–572.

Hatch, E., & Lazaraton, A. (1991). *The research manual: Design and statistics for applied linguistics*. Boston: Heinle & Heinle.

Hosmer, D., & Lemeshow, S. (1989). *Applied logistic regression*. New York: Wiley.

Iwashita, N. (2001). The effect of learner proficiency on interactional moves and modified output in nonnative-nonnative interaction in Japanese as a foreign language. *System*, *29*, 267–287.

Iwashita, N. (2003). Negative feedback and positive evidence in task-based interaction. *Studies in Second Language Acquisition*, *25*, 1–36.

Izumi, S. (2003). Comprehension and production processes in second language learning: In search of the psycholinguistic rationale of the output hypothesis. *Applied Linguistics*, *24*, 168–196.

Kormos, J. (1999). The effect of speaker variables on the self-correction behaviour of L2 learners. *System*, *27*, 207–221.

Kormos, J. (2000). The role of attention in monitoring second language speech production. *Language Learning*, *50*, 343–384.

Kowal, M., & Swain, M. (1994). Using collaborative language production tasks to promote students' language awareness. *Language Awareness*, *3*, 73–93.

Krashen, S. (1998). Comprehensible output? *System*, *26*, 175–182.

Leeman, J. (2003). Recasts and second language development: Beyond negative evidence. *Studies in Second Language Acquisition*, *25*, 37–63.

Levelt, W. (1989). *Speaking: From intention to articulation*. Cambridge, MA: MIT Press.

Linnell, J. (1995). Can negotiation provide a context for learning syntax in a second language? *Working Papers in Educational Linguistics*, *11*, 83–103.
Long, M. H. (1996). The role of the linguistic environment in second language acquisition. In W. Ritchie & T. K. Bhatia (Eds.), *Handbook of language acquisition: Vol. 2. Second language acquisition* (pp. 413–468). San Diego, CA: Academic Press.
Long, M. H., Inagaki, S., & Ortega, L. (1998). The role of implicit negative feedback in SLA: Models and recasts in Japanese and Spanish. *Modern Language Journal*, *82*, 357–371.
Lyster, R., & Ranta, L. (1997). Corrective feedback and learner uptake. *Studies in Second Language Acquisition*, *19*, 37–66.
Mackey, A. (1997). *Interactional modifications and the development of questions in ESL*. Unpublished manuscript, Georgetown University, Washington, DC.
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. (2000, March). *Interactional feedback on L2 morpho-syntax: Learners' perceptions and developmental outcomes*. Paper presented at the American Association of Applied Linguistics Conference, Vancouver, British Columbia, Canada.
Mackey, A. (in press). Interaction and second language development: Perspectives from SLA research. In R. DeKeyser (Ed.), *Practice in second language learning: Perspectives from linguistics and psychology*. New York: Cambridge University Press.
Mackey, A., Gass, S., & McDonough, K. (2000). How do learners perceive implicit negative feedback? *Studies in Second Language Acquisition*, *22*, 471–497.
Mackey, A., McDonough, K., Fujii, A., & Tatsumi, T. (2001). Investigating learners' reports about the L2 classroom. *IRAL*, *39*, 285–307.
Mackey, A., & Oliver, R. (2002). Interactional feedback and children's L2 development. *System*, *30*, 459–477.
Mackey, A., Oliver, R., & Leeman, J. (2003). Interactional input and the incorporation of feedback: An exploration of NS-NNS and NNS-NNS adult and child dyads. *Language Learning*, *53*, 35–66.
Mackey, A., & Philp, J. (1998). Conversational interaction and second language development: Recasts, responses, and red herrings? *Modern Language Journal*, *82*, 338–356.
Mackey, A., Philp, J., Egi, T., Fujii, A., & Tatsumi, T. (2002). Individual differences in working memory, noticing of interactional feedback, and L2 development. In P. Robinson & P. Skehan (Eds.), *Individual differences in L2 learning* (pp. 181–209). Amsterdam: Benjamins.
McDonough, K. (2001). *Exploring the relationship between modified output and L2 learning*. Unpublished doctoral dissertation, Georgetown University, Washington, DC.
Morris, F., & Tarone, E. (2003). Impact of classroom dynamics on the effectiveness of recasts in second language acquisition. *Language Learning*, *53*, 325–368.
Oliver, R. (1995). Negative feedback in child NS-NNS conversation. *Studies in Second Language Acquisition*, *17*, 459–481.
Nobuyoshi, J., & Ellis, R. (1993). Focused communication tasks and second language acquisition. *ELT Journal*, *47*, 203–210.
O'Malley, J., Chamot, A., Stewner-Manzanares, R., Russo, R., & Kupper, L. (1985). Learning strategy applications with students of English as a second language. *TESOL Quarterly*, *19*, 185–296.
Pampel, F. (2000). *Logistic regression: A primer*. London: Sage.
Panova, I., & Lyster, R. (2002). Patterns of corrective feedback and uptake in an adult ESL classroom. *TESOL Quarterly*, *36*, 573–595.
Philp, J. (2003). Constraints on "noticing the gap": Nonnative speakers' noticing of recasts in NS-NNS interaction. *Studies in Second Language Acquisition*, *25*, 99–126.
Pica, T. (1988). Interlanguage adjustments as an outcome of NS-NNS negotiated interaction. *Language Learning*, *38*, 45–73.
Pica, T. (1994). Research on negotiation: What does it reveal about second-language learning conditions, processes, and outcomes? *Language Learning*, *44*, 493–527.
Pica, T., Holliday, L., Lewis, N., & Morgenthaler, L. (1989). Comprehensible output as an outcome of linguistic demands on the learner. *Studies in Second Language Acquisition*, *11*, 63–90.
Pica, T., Kanagy, R., & Falodun, J. (1993). Choosing and using communication tasks for second language instruction and research. In G. Crookes & S. Gass (Eds.), *Tasks and second language learning: Integrating theory and practice* (pp. 9–34). Clevedon, UK: Multilingual Matters.
Pienemann, M., & Johnston, M. (1987). Factors influencing the development of language proficiency. In D. Nunan (Ed.), *Applying second language acquisition research* (pp. 45–141). Adelaide: National Curriculum Resource Centre, AMEP.

Pienemann, M., Johnston, M., & Brindley, G. (1988). Constructing an acquisition-based procedure for second language assessment. *Studies in Second Language Acquisition*, *10*, 217–243.
Pinker, S. (1989). *Learnability and cognition: The acquisition of argument structure*. Cambridge, MA: MIT Press.
Politzer, R., & McGroarty, M. (1985). An exploratory study of learning behaviors and their relationship to gains in linguistic and communicative competence. *TESOL Quarterly*, *19*, 103–123.
Roberts, M. (1995). Awareness and the efficacy of error correction. In R. Schmidt (Ed.), *Attention and awareness in foreign language learning* (Tech. Rep. No. 9, pp. 163–182). Honolulu: University of Hawai'i Press.
Saito, H. (1999). Dependence and interaction in frequency data analysis in SLA research. *Studies in Second Language Acquisition*, *21*, 453–475.
Samuda, V. (2001). Guiding relationships between form and meaning during task performance: The role of the teacher. In M. Bygate, P. Skehan, & M. Swain (Eds.), *Researching pedagogic tasks: Second language teaching and testing* (pp. 119–140). London: Longman.
Schachter, J. (1986). Three approaches to the study of input. *Language Learning, 36*, 211–225.
Schachter, J. (1991). Corrective feedback in historical perspective. *Second Language Research*, *7*, 89–102.
Schmidt, R. (1995). Consciousness and foreign language learning: A tutorial on the role of attention and awareness in learning. In R. Schmidt (Ed.), *Attention and awareness in foreign language learning* (Tech. Rep. No. 9, pp. 1–63). Honolulu: University of Hawai'i Press.
Schmidt, R. (2001). Attention. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 3–32). New York: Cambridge University Press.
Sharwood Smith, M. (1991). Speaking to many minds: On the relevance of different types of language information for the L2 learner. *Second Language Research*, *7*, 118–132.
Shonerd, H. (1994). Repair in spontaneous speech: A window on second language development. In V. John-Steiner, C. Panofsky, & L. Smith (Eds.), *Sociocultural approaches to language and literacy* (pp. 82–108). New York: Cambridge University Press.
Silver, R. (2000). Input, output and negotiation: Conditions for second language development. In B. Swierzbin, F. Morris, M. Anderson, C. Klee, & E. Tarone (Eds.), *Social and cognitive factors in second language acquisition: Selected proceedings of the 1999 Second Language Research Forum* (pp. 345–371). Somerville, MA: Cascadilla Press.
Skehan, P., & Foster, P. (1997). Task type and task processing conditions as influences on foreign language performance. *Language Teaching Research*, *1*, 185–211.
Spada, N., & Lightbown, P. (1993). Instruction and the development of questions in L2 classrooms. *Studies in Second Language Acquisition*, *15*, 205–224.
Spada, N., & Lightbown, P. (1999). Instruction, first language influence, and developmental readiness in second language acquisition. *Modern Language Journal*, *83*, 1–22.
Storch, N. (2001). Comparing ESL learners' attention to grammar on three different classroom tasks. *RELC Journal: A journal of language teaching and research in Southeast Asia*, *21*, 104–124.
Swain, M. (1985). Communicative competence: Some roles of comprehensible input and comprehensible output in its development. In S. Gass & C. Madden (Eds.), *Input in second language acquisition* (pp. 235–253). Rowley, MA: Newbury House.
Swain, M. (1993). The output hypothesis: Just speaking and writing aren't enough. *Canadian Modern Language Review*, *50*, 158–164.
Swain, M. (1995). Three functions of output in second language learning. In G. Cook & B. Seidlhofer (Eds.), *Principle and practice in applied linguistics: Studies in honour of H. G. Widdowson* (pp. 125–144). Oxford: Oxford University Press.
Swain, M. (1998). Focus on form through conscious reflection. In C. Doughty & J. Williams (Eds.), *Focus on form in classroom SLA* (pp. 64–81). New York: Cambridge University Press.
Swain, M. (2000). The output hypothesis and beyond: Mediating acquisition through collaborative dialogue. In J. Lantolf (Ed.), *Sociocultural theory and second language learning* (pp. 97–114). Oxford: Oxford 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*, 371–191.
Swain, M., & Lapkin, S. (1998). Interaction and second language learning: Two adolescent French immersion students working together. *Modern Language Journal*, *82*, 320–337.
Truscott, J. (1999). What's wrong with oral grammar correction. *Canadian Modern Language Review*, *55*, 437–456.
Young, R., & Yandell, B. (1999). Top-down versus bottom-up analyses of interlanguage data: A reply to Saito. *Studies in Second Language Acquisition*, *21*, 477–488.

APPENDIX A

Developmental stages in ESL question formation with data from the current study (stages adapted from Pienemann & Johnston, 1987, and Pienemann et al., 1988, following a range of previous research: Mackey, 1997, 1999, 2000; Mackey & Oliver, 2002; Mackey & Philp, 1998; Philp, 2003; Silver, 2000; Spada & Lightbown, 1993, 1999).

Table A1. Examples of developmental stages and question forms

Stages	Constructions	Examples
3: Fronting	*Do* + SVO?	*Does he like Chiang Mai University?* **Does he came here with his friends?* *Do you think it's suitable for him?* **Does a man and a woman married?*
	Wh + (*be*/*do*) SVO?	**Why does her party was very idiotic?* **Why she do that?* **What is Mark is going to do?* **Why the woman are waiting?*
	Be + SVO?	**Are these three people get involved with the plan?* **Is they are friends?*
4: Pseudo-inversion; *yes*/*no* inversion	(*Wh*) + copula + S	*Where are they now?* *Who is the girl wearing a yellow shirt?* *How hard is English to understand?* *Is it the department store?*
	Aux/modal + SV	**Could you gave me some suggestion?* *Have you been in trouble on social relationships because of the effect of technology?*
5: Aux second	*Wh* + aux/*do*	**Where is he come from?* *Why is the girl looking at the man?* *What will happen in the future if we still use technology in the wrong way?* *How can this technology affect social relationships?* **What does they do?*
6: Cancelled inversion; negative questions; tag questions	Cancelled inversion	*May I ask you how English is used in your country?* *Would you mind telling me where you are staying right now?*
	Negative questions	*Why didn't he go to see the doctor about his hair?* **Hasn't he marry Susan already?*
	Tag questions	*That is Bob's farm, isn't it?* **He wants to divorce Mary, didn't he?*

APPENDIX B

Table B1. Treatment activities

Name	Description	Type	Exchange of information	Direction of information
Puzzle stories	Solving a mystery by asking questions about what happened	Information gap	Required	One way
What are they talking about?	Figuring out what happened by asking questions	Information gap	Required	One way
Who's telling the truth?	Determining if an interlocutor is telling the truth or exaggerating	Information gap	Required	One way
Picture difference	Identifying the differences in two versions of a picture	Information exchange	Required	Two way

Note. The treatment and testing tasks were described following the framework for classifying tasks put forward in Pica, Kanagy, and Falodun (1993).

APPENDIX C

Pronunciation? → → → → → →
Vocabulary? → → → → → →
Grammar? → → → → → →
Anything else? → → → → → →

Figure C1. Learning journal completed by learners in response to the question *did you learn anything today about. . . .*

APPENDIX D

Table D1. Test items

Name	Examples	Description	Type	Exchange of information	Direction of information
Story completion	Explosive revenge; poisonous perfume; ranch romance; steal from the rich	Discovering a story by asking questions about a series of pictures	Information gap	Required	One way
Topic discussion	Language learning; conversation strategies; world English; future life	Expressing opinions about topics and issues related to course themes	Information gap	Optional	One way
Interview preparation	Learning styles; learning strategies; English in advertising; computers/technology	Brainstorming questions to be used in oral interviews	Information gap	Optional	One way