

# Roles of Output and Feedback for L2 Learners' Noticing

**Hideki Sakai**

*Shinshu University*

This study explores the roles of output and feedback in L2 learners' noticing and repair in later production. Sixteen Japanese learners of English, assigned to one of two conditions (model or recast), performed communicative tasks and afterwards reported what they noticed about errors or linguistic problems in a retrospective interview. Results showed that although the learners noticed linguistic problems by producing the target language, they had difficulty in attending to and incorporating the subsequent input for later production. Also the results indicated that recasts were more effective to prompt the noticing of errors than were models, and that noticing a *gap* through feedback led to a relatively higher number of repairs on the second trial than was the case of noticing a *hole* through output.

本研究は、アウトプットとフィードバックの役割を、気づきと学習効果（2度目の発話時における修正）の点から調べた。日本人英語学習者 16 名が、フィードバックとしてモデルが与えられたグループと言い直し (recasts) が与えられたグループに分けられた。学習者は、コミュニケーション・タスクの後で気づきについて言語報告した。主な結果は次の通りである。学習者は目標言語を表出することによって言えないことに気づいたが、その後でフィードバックによって与えられたインプットに注目して取り入れ、2度目の発話時に表出することは難しかった。また、言い直しはモデルよりも気づきを促進する点で効果的であった。さらに、フィードバックによってギャップに気づくこと (noticing a gap) の方がアウトプットによって言えないことに気づくこと (noticing a hole) よりも2度目の発話時の修正につながることを示された。

### *Noticing and Second Language Acquisition*

The issue of noticing has drawn much attention in second language acquisition (SLA) research (e.g., Ellis, 1994; Gass, 1988, 1997; Long, 1991, 1996; Long & Robinson, 1998; Robinson, 1995; Schmidt, 1990, 1993, 1995; Schmidt & Frota, 1986; Skehan, 1998; Swain, 1985, 1998; Tomlin & Villa, 1994; Truscott, 1998). In his noticing hypothesis, Schmidt (1990, 1993, 1995; Schmidt & Frota, 1986) shed light on focal attention, or noticing, as a necessary and sufficient condition for input to become intake in SLA, by claiming "intake is that part of the input that the learner notices" (1990, p. 139). According to Schmidt (1993), second language (L2) learners need to not only comprehend the input but also notice "whatever features of the input are relevant for the target system" (p. 209).

An attempt has been made to categorize noticing based on what is noticed. Swain (1998, p. 66) pointed out that there can be different types of noticing such as: (a) noticing a form in the input, (b) noticing one's interlanguage deficiencies (or holes), and (c) noticing the gap between the interlanguage and the target language. The first type of noticing is when, while listening to or reading the target language input, the learner simply attends to the formal aspects of the target language in the input. Input frequency, saliency of the input, and external manipulations such as input enhancement, may influence this first type of noticing. In the second type, noticing the interlanguage deficiencies is when learners may notice that they cannot say what they want to say precisely in the target language. This type of noticing is also referred to as noticing "holes" (Doughty & Williams, 1998, p. 228; Swain, 1998, p. 66). Thus, the term *holes* has been adopted in this paper. Swain argued that one major function of production in the target language is to facilitate this type of noticing (1993, 1995). Finally, the third type of noticing is when learners may notice that their current interlanguage is different from the target language. Feedback provided during interactions may help learners notice their errors, that is, notice the gap (Long, 1996). The present study focuses on output and feedback because both are considered to promote different types of noticing, as will be seen in the following sections.

### *Output and Noticing*

In her output hypothesis, Swain (1985, 1993, 1995) proposed that producing the target language contributes to SLA via several functions: (a) the automaticity function, (b) the hypothesis-testing function, (c) the

metalinguistic function, and (d) the noticing/triggering function (see also de Bot, 1996).

The last function of output is the most important to this study. Swain (1995) stated that “the activity of producing the target language may prompt second language learners to consciously recognize some of their linguistic problems; it may bring to their attention something they need to discover about their L2” (p. 126). Consequently, output enhances the likelihood of processing the subsequent input (Swain, 1993, p. 160). Swain and her colleague (Swain, 1998; Swain & Lapkin, 1998) showed that the activity of producing the target language helped L2 learners to notice their linguistic problems in the target language.

The question arises then: Does noticing the linguistic problems promote the noticing of relevant forms in subsequent input and result in learning? Izumi, Bigelow, Fujiwara, and Fearnow (1999) and Izumi and Bigelow (2000) attempted to answer this question, but found no empirical support for this assumption. They made two groups (an experimental group and a control group) and gave two types of writing tasks to both groups: text reconstruction tasks and guided essay-writing tasks. In the text reconstruction tasks, the experimental group read texts for the succeeding writing tasks; on the other hand, the control group read the same texts only for comprehension. In the guided essay-writing tasks, the experimental group performed writing tasks for given topics before the model input while the control group performed writing tasks for unrelated topics. Both groups wrote essays after the model input. The researchers examined the learners’ noticing while reading, the use of the target forms (the past hypothetical conditional) in their writing tasks, and the performance in the multiple-choice recognition tests and the picture-cued production tests. Results showed that the writing activity did not better enhance noticing of the target forms in the model input provided after the writing activity in the experimental group than in the control group. These studies did not lend support for Swain’s (1993, 1995) argument that noticing holes facilitates noticing the relevant forms in the subsequent input.

It should be noted that it was not clear whether their learners noticed their linguistic problems at the first production (for other methodological issues, see Whitlow, 2001; Izumi & Bigelow, 2001). As Izumi and Bigelow (2000) admitted, “not all learners necessarily found their IL [interlanguage] grammar to be problematic during production, which in turn affected their attention to the grammatical form when they were exposed to the input” (p. 271). Schmidt (1993) cautions that “the prob-

lem with this external approach is that the treatment may not have the intended effect” (pp. 218-219). Izumi and his colleagues assumed that writing tasks had caused their learners to notice their linguistic problems, but in fact they might have failed to notice their linguistic problems as to the target structure. Thus, it remains unclear how or even whether output functions as a facilitator of noticing.

### *Feedback and Noticing*

In this section, I will focus on feedback and its effect on noticing. In his updated version of the interaction hypothesis, Long (1996) proposed that the negative feedback provided during negotiation for meaning may facilitate SLA because it effectively draws L2 learners’ attention to the linguistic forms in the course of interaction. Recasts are such feedback, defined as reformulation of L2 learners’ erroneous utterances by the interlocutor<sup>1</sup>, maintaining the meanings which L2 learners intend to convey (Long, 1996, p. 434; Lyster, 1998a, p. 58).

Recasts occur most frequently in NS-NNS interactions (Oliver, 1995; Van den Branden, 1997) and in classrooms (Doughty, 1994; Lyster, 1998a, 1998b; Lyster & Ranta, 1997; Roberts, 1995). Also, several studies have attempted to show that recasts are effective for language learning<sup>2</sup> (Doughty & Varela, 1998; Long, Inagaki, & Ortega, 1998; Mackey & Philp, 1998; for a review, see Nicholas, Lightbown, & Spada, 2001). Recasts are considered to be effective for SLA in that they possess the dual function of providing positive evidence and negative evidence at the same time, which may prompt L2 learners’ noticing because it makes cognitive comparison easier (Long, 1996, p. 434). Here is an example from the present study.

(1) From the recast group

Hikaru:        *His mother angry.*

Researcher:   *His mother is angry?* [recast]

The learner, Hikaru<sup>3</sup>, produced an utterance that included an error related to the “Be” verb *is*. The recast was provided in the next turn in the form of a confirmation check.<sup>4</sup> This feedback was a reformulation of the erroneous utterance and maintained the proposition intended by the learner. In this example, the recast indicated that (a) the omission of

the “Be” verb was wrong in English (negative evidence), and that (b) the correct form to be inserted was *is* (positive evidence).

As Ellis (1994, p. 96) pointed out, an external comparison appearing in the recast exchange may not always result in a cognitive comparison. The question then arises as to whether L2 learners actually notice the corrective nature of recasts. In other words, do recasts promote L2 learners’ noticing of the gap? The first evidence for this, though indirect, is that L2 learners are more likely to imitate recasts than they are to repeat a non-corrective repetition (Long, 1996, pp. 435-439). For example, Doughty (1994) found that university students learning French as a foreign language imitated 21.5% of the teacher’s recasts while they repeated 2.3% of the non-corrective repetition. Because recasts and non-corrective repetition have common functional distributions and are both repetitive in nature (Lyster, 1998a), L2 learners are considered to be responding to the corrective nature of recasts.<sup>5</sup>

Another approach to the question is to examine L2 learners’ noticing through introspection<sup>6</sup> (Mackey, Gass, & McDonough, 2000). Mackey et al. (2000) utilized stimulated recall to explore how L2 learners perceive feedback in interactions. The participants for their study were 10 ESL learners and seven learners of Italian as a foreign language. The learners performed a communicative task, which was videotaped for a later recall session. They were then asked to report verbally what they had thought during the interaction. During the task, the feedback in the form of recasts and negotiation was naturally provided. Results of the analysis of the feedback and the verbal reports showed that morphosyntactic feedback was generally provided in the form of recasts and was not perceived as being about morphosyntax. In contrast, phonological and lexical feedback, generally in the form of negotiation and combination (negotiation and recast), was perceived more accurately by the learners. The researchers went on to state that “the window of opportunity for these learners to notice grammar in interaction may have been relatively small” (p. 488). Their study is of great interest because their findings suggest the possibility that L2 learners would not notice the target of the feedback as intended by the feedback provider. However, only a small number of studies have attempted to explore the noticing issue in order to investigate the roles of feedback. Further studies need to be carried out.

### Purpose of the Study

The main purpose of this study was to examine the verbal reports and to clarify how L2 learners notice errors or linguistic problems in their production. The research questions posited for this study are as follows:

1. Does output promote L2 learners' noticing of their linguistic problems?
2. Do recasts and models help L2 learners notice their errors? If so, which feedback type is more effective?
3. Do modeling and recasting lead to any different learning effects? In other words, does the time between the utterance and the feedback make any difference in learning?
4. Do noticing the gap and noticing the hole lead to differential learning effects?

Two conditions (recasts and models) were set for this study. The term *model* in the present study refers to a sample of the target language provided after an interval of a few minutes, whose meanings reflect what the learner intended to convey. So, a model does not immediately follow the learner's ill-formed utterance; a recast does. In other words, a model is different from a recast in terms of its time of occurrence in relation to the learner's preceding utterance. The effects of recasts were compared with those of models to examine whether the elapsed time between the utterance and the feedback would make any difference. When learners notice holes in their production, both recasts and models will provide positive evidence to which they may readily pay attention, resulting in language learning. In this case, the difference between the model condition and the recast condition lies in how long learners can stay motivated to attend to the new input in order to receive their necessary information about the target language. When learners notice gaps through feedback, both recasts and models work as negative feedback. This provides negative evidence: information as to what is not allowed in the target language. In this case, the two conditions differ in that the learners in the model condition have to keep their original utterances in mind for a longer time than the learners in the recast condition do in order to make the cognitive comparison.

The model condition in this study was similar to the experimental condition in Izumi and Bigelow's (2000; Izumi et al., 1999) studies. The

learners in their experimental condition were involved in writing and then were provided with ready-made reading materials as the model input. The researchers pointed out that the degree of the discrepancy between the learner's production and the model input may influence the difficulty of noticing the gap (2000, p. 263). In contrast, for the present study, each learner in the model condition received the tailor-made model input, which made it easier to directly compare the output and the model input.

As described above, this study was designed taking the following points into account:

1. Two feedback types were provided: models and recasts. Recasts immediately followed the learners' utterances whereas models were provided a few minutes later.
2. To make models equivalent to recasts, models were created on the basis of the learners' utterances.

## Method

### *Participants*

Sixteen first-year university students (9 women and 7 men) volunteered to participate in this study. The participants were attending a required English class in a Japanese national university. They were Japanese learners of English as a foreign language (EFL); that is, all of them had the same L1 background (the Japanese language) and came from a similar learning environment (school subjects). The students were randomly assigned to one of the two groups: the model group ( $n = 8$ ) and the recast group ( $n = 8$ ). The model group included 5 women and 3 men (mean age = 18.4 years); the recast group included 4 women and 4 men (mean age = 18.5 years). The average years of English study was 6.06 for the model group and 6.75 for the recast group, ranging from 6 to 8 years; that is, most of the students had started to study English as a school subject when they were in the seventh grade, the first year of junior high school.

### *Procedure*

Each student took part individually in the experimental session<sup>8</sup> in June 2001. As can be seen in Table 1, the session consisted of three phases: Phase 1, Phase 2, and a retrospective interview. The total session

took approximately 30 minutes. The instruction and the retrospective interview were carried out in Japanese (L1).

**Table 1. Experimental Sequence**

	Model Group	Recast Group
	Explanation & Practice	Explanation & Practice
Phase 1	1. Picture Description Task - no feedback 2. Picture Identification Task	1. Picture Description Task - recasts/repetition
Phase 2	3. Picture Description Task	2. Picture Description Task
Phase 3	4. Retrospective Interview	3. Retrospective Interview

One by one, each student sat at a table across from me (a nonnative speaker of English whose L1 is Japanese). The session was videotaped and recorded with two tape recorders (one for the data collection and one for the interview). In Phase 1, the student was handed six cards. Each card had two similar pictures (Picture A and Picture B) on it (see Appendix). The student was told to select one of the two pictures and describe it in English so that I could identify that specific picture. Afterwards, the student was asked to move on to the next card and to repeat the procedure.

For the model group, I provided nodding along with expressions such as “I see,” “OK,” and “Yes.” In Example (2) from this study, Momoko tried to describe one of the pictures on Card 2 (Picture B). I demonstrated my understanding of her utterances by saying, “OK?” “Umm,” and “OK.” After she produced three utterances, I identified the picture (“Number 2-B”). Following the picture description task, the student in the model group performed the picture identification task. As in Example (3), I described the six pictures and the student identified those pictures. Through this task, model input was provided. During the previous picture description task, I took notes on what the student had said about the pictures and based the model input on these utterances in order to provide the model containing necessary information for the student. Therefore, the tailor-made model input was different for each student.

## (2) Picture Description Task (the model group)

Momoko: *There are baby and woman.* (a)

Researcher: *OK?*

Momoko: *Baby ask ... what is that.* (b)

Researcher: *Umm.*

Momoko: *Woman say this is cat.* (c)

Researcher: *OK. Number 2-B.*

Momoko: *Yes.*

## (3) Picture Identification Task (the model group)

Researcher: *There are a baby and a woman.* (a)

*The baby asks what that is.* (b)

*The woman says that is a cat.* (c)

Momoko: *Number 2-B.*

Researcher: *OK.*

The recast group performed only the picture description task. As in Example (4), I repeated every grammatical utterance (non-corrective repetition) or reformulated every erroneous utterance (recast). Both types of feedback were provided in the form of a confirmation check. Non-corrective repetition was provided to make the amount of feedback comparable to that of the model condition.

## (4) Picture Description Task (the recast group)

Hikaru: *There is a boy and his mother.*

Researcher: *There are a boy and his mother? [recast]*

Hikaru: *Yes.*

Researcher: *OK.*

Hikaru: *The boy is reading a comic books on the sofa.*

Researcher: *Umm. The boy is reading a comic book on the sofa? [recast]*

- Hikaru: *Yes.*
- Researcher: *Umm.*
- Hikaru: *His mother angry.*
- Researcher: *His mother is angry?* [recast] *OK.*
- Hikaru: *Because he ... he doesn't study.*
- Researcher: *Because he doesn't study?* [non-corrective repetition]
- Hikaru: *Yes.*
- Researcher: *OK. Number 7-A.*
- Hikaru: *Yes.*

The description task was designed to elicit a variety of errors. Target structures were not determined for this study in order for the experimental task to be more natural. The task was considered to be communicative in that the student's main focus was primarily on the messages (focus on meaning) and because I did not know what the student was going to describe (information gap).

The two groups performed different tasks only in Phase 1. In Phase 2, the students were asked to describe the same six pictures that they had described in the picture description task of Phase 1. For example, because Momoko of Example (2) selected and described Picture B of Card 2 in Phase 1, she was asked to describe that picture (2-B) in English. In this second task, no feedback was provided. Because students described the same pictures in English in Phases 1 and 2, comparing the two phases was relatively easy. In the second trial, learning effects were examined by analyzing whether the students corrected the errors they had made in the first task. Then the retrospective interview followed.

The students were informed that the purpose of the study was to examine the processes of speaking. However, they were not provided with information about the sequence of the tasks in the session.

#### *Measurement of Noticing*

The operational definition of noticing used for this study was the "availability for verbal report" (Schmidt, 1990, p. 132; see also Leow, 1997).<sup>9</sup> As Robinson (1995) pointed out, measuring noticing is not an

easy task because (a) “the experience of noticing may be fleeting and thus difficult to recall” and (b) “one may be aware of, yet unable to verbalize or otherwise articulate the nature of that which one is aware of” (p. 299). For this study, I needed to clarify whether L2 learners’ noticing would be brought about by feedback or not. In the stimulated recall (Gass & Mackey, 2000), learners are asked about what they were thinking at the time of feedback (Mackey et al., 2000). However, the stimulated recall did not seem to give sufficient information about noticing. For example, Mackey et al. (2000) provided an example of feedback that was perceived as phonological. The learner pronounced the word *flowers* as [flurs], and the interlocutor made the recast, “Floors?” In this case, the recast did not reformulate what the learner intended to say, and thus the learner had to reach the proper pronunciation by himself or herself. In the recall session, the learner stated, “I was thinking that my pronounce, pronunciation is very horrible” (p. 486). The researchers argued that this recall indicated that the feedback target was perceived as phonological by the learner. However, it is unclear from this recall protocol whether or not the feedback brought about the learner’s noticing of the wrong pronunciation. The learner might have already noticed his or her problem with pronunciation at the moment of the original utterance. For the present study, I decided to ask guided questions in the retrospective interview<sup>10</sup> in order to clarify the noticing types.

As Gass and Mackey (2000) argued, during the retrospective interview, stimulus was provided to promote recall in order for the students to report verbally as much as possible. That is to say, during the retrospective interview, the tape-recorded interactions in the picture description task (Phase 1) were played back, with the pictures, in order to assist the students in verbally reporting their noticing about their linguistic problems or errors. The tape was stopped after each utterance, and the following questions were asked: (a) Are there any errors in your utterance?, (b) If so, when did you notice them?, and (c) What were you thinking at that time? When the students requested it, their recorded utterances were repeated to them.

### *Coding*

#### *Noticing Types*

In this study, noticing was put into four categories, mainly on the basis of Swain’s (1998) noticing types: (a) problem, (b) production, (c) model/recast, and (d) a lack of noticing. The first type of noticing, coded

as *problem*, refers to noticing a hole in the interlanguage system. In most cases, the learners reported that in English they could not express what they wanted to say. In Example (5), Miho intended to say that the mother wanted the boy to study, but she had difficulty expressing this in English.<sup>11</sup> Thus, this verbal report indicated that Miho had noticed her linguistic problem at the moment of production.

(5) Noticing type 1: Problem

Tape: *(Her mother ... want ... he ... he mother want ... he ... he study ... homework.)*

Researcher: *Are there any errors?*

Miho: *Well, let's see, I don't know.*

Researcher: *At this moment, what ... what did you think?*

Miho: *Well, let' see, really.... As usual.... Like this. I wanted to say that this mother wanted this boy to study .... Somehow, with a that-clause... Like this ... as usual. What should I have said? I wanted to say she wanted this boy to study, but .... While I was saying in English, I found myself meaning that it was the mother who wanted to study .... Umm. I thought I should say it in a different way, but I didn't know how to express the meanings.*

In Example (6), Saori reported that what she had said in the task was wrong and that she had noticed the errors at the moment of production. When the learners noticed their errors at the moment of production, those reports were categorized as *production*. In the example, the learner did not report what she should have said in the target language (*old woman*). Some learners reported as follows, "I should have said *old woman*." This noticing type (*production*) is different from the first noticing type (*problem*) in that the learners did not state what they had intended to express; rather, they pointed out the errors in their utterances. This noticing type is identified with noticing the gap between the interlanguage and the target language without the help of feedback.

## (6) Noticing type 2: Production

- Tape:            (*There are old girl and boy.*)
- Saori:            *There is something strange with the expres-  
sion "old girl."*
- Researcher:    *Yes. When did you notice it?*
- Saori:            *When I was saying it.*

The third noticing type, *model/recast*, refers to noticing the gap with the help of feedback. In Example (7), Hikaru noticed the error in her utterance and reported that she had not inserted the "Be" verb *is*. Additionally, she stated that she had noticed the error through the feedback (recast for her).

## (7) Noticing type 3: Model/recast

- Tape:            (*His mother angry.*)
- Researcher:    *What about this?*
- Hikaru:         *I didn't say "is."*
- Researcher:    *Yes. When did you notice that?*
- Hikaru:         *When I heard you saying.*

Finally, those errors that the learners did not notice or the errors which they noticed at the moment of the interview were categorized as *a lack of noticing* (the fourth noticing type). In Example (8), Hikaru did not notice her error in using the noun phrase *a comic books*.

## (8) Noticing type 4: A lack of noticing

- Tape:            (*The boy is reading a comic books on the  
sofa.*)
- Researcher:    *What about this?*
- Hikaru:         *Nothing particular.*
- Researcher:    *No errors?*
- Hikaru:         *No.*

### *Learning Effects*

Learning effects were measured by analyzing the learners' performance in the second picture description task. When learners correctly produced utterances which had been problematic in the first task, those parts were categorized as *repair*. The repair category includes (a) the incorporation of the features provided in the model or recast and (b) the correct reformulation of the original problematic part, though the form is not identical with what the model or recast provided. The problematic parts, which still remained erroneous, were coded as *other*. This category contains (a) the same errors, (b) different errors, and (c) avoidance. Avoidance is referred to as non-occurrence of the linguistic context in question. These categories developed for this study were based on Lyster and Ranta's (1997) uptake categories.

### *Analysis*

The learners produced a variety of errors, such as lexical errors, grammatical-morpheme errors, and syntactic errors (Sakai, 2002). The learners' pronunciation errors were not examined in this study because most of the learners' English was intelligible enough for me to understand probably because we shared the same L1 background and also because I did not intend to correct their pronunciation from the beginning.

The following analyses of the present study were based on *error points*, operationalized as the absolute number of errors identified by comparing the learner's utterance with the feedback.<sup>12</sup> The following example, Example (9), is part of Examples (2) and (3) described above. The model provided for the learner indicated several corrections of the learner utterance: (a) insertion of the definite article, (b) the 3rd-person present singular form, and (c) the word order in the subordinate clause. Thus, in the example, there were three error points.

#### (9) Error points (in the case of Momoko)

Learner Utterance: *Baby ask ... what is that.* [Three error points]

Model: *The baby asks what that is.*

I found verbal reports to vary among learners, even if they had received the same corrections through models or recasts. For example, some learners may make some comments only about the present 3rd-

person singular form whereas others may talk about all three corrections as in Example (9). In this study, therefore, the more detailed analyses were made on the basis of error points.

I coded all the data of the 16 learners. To assess inter-rater reliability for the coding, a subset (50%) of both the retrospective interview data and the performance data (Phase 2) was coded by another rater. For the coding of noticing types in the retrospective interview data, agreement reached 88.7% (258 out of 291); for the coding of learning effects in the performance data, agreement was 88.6% (272 out of 307).

### *Statistical Procedures*

Because the sample size was small and the data were not normally distributed, I selected nonparametric procedures (Hatch & Lazaraton, 1991, p. 270). Medians were used as a measure of central tendency. Also, instead of standard deviations, this study used interquartile ranges (IQRs), which indicate the ranges of the middle 50% of the data. The Wilcoxon rank-sum tests were performed to test whether the two groups differed in terms of individual scores of error points and frequencies of feedback. For the analyses of noticing types and effects of learning, the chi-square tests were performed on the groups' data. When an expected cell frequency was less than five, Fisher's exact test was used (Hatch & Lazaraton, 1991, p. 409). Due to the small number of tokens of data, these analyses were carried out by merging individual data into two groups, model and recast.

## **Results**

### *General Description*

First of all, the two groups were compared in terms of error points in the picture description task of Phase 1 and the frequencies of feedback (models or recasts). Although this study did not control these factors in advance, the following general description data suggest that the learners in the two groups made errors and received feedback to a similar extent. Therefore, the two groups were considered to be comparable to each other in terms of the proficiency levels of learners and the provision of feedback.

Individual scores of error points in the picture description task of Phase 1 are summarized in Table 2. At first glance, the model group ( $Mdn = 39.5$ ,  $IQR = 9.5$ ) seemed to make more errors than the recast group did ( $Mdn = 28.5$ ,  $IQR = 15.25$ ). However, the Wilcoxon rank-sum

test showed that there was no statistically significant difference between the two groups ( $p = .19$ ). Table 3 shows the numbers of models or recasts provided for each group. The results show that each student in both groups received about 20 of the models or recasts. The Wilcoxon rank-sum tests revealed no statistically significant differences between the groups ( $p = 1.00$  for the numbers of models/recasts).

**Table 2. Error Points for Each Participant  
(Picture Description Task of Phase 1)**

Model Group ( $n = 8$ )			Recast Group ( $n = 8$ )		
Participants	Sex	Error points	Participants	Sex	Error points
Yuzuko	female	51	Ayumi	female	43
Momoko	female	41	Saori	female	31
Makiko	female	41	Ryoko	female	26
Miho	female	38	Hikaru	female	17
Hanako	female	32	Akira	male	52
Taro	male	43	Noritake	male	38
Kyoji	male	32	Koji	male	24
Hideki	male	29	Ichiro	male	24

*Note.* All the names are fictitious.

**Table 3. Number of Models/Recasts**

	Model Group ( $n = 8$ )		Recast Group ( $n = 8$ )	
	<i>Mdn</i>	<i>IQR</i>	<i>Mdn</i>	<i>IQR</i>
Number of models/recasts	18	4.25	17	5.25

#### *Analysis of Verbal Report on Noticing*

The results of the analysis of verbal report are summarized in Table 4. The learners of the model group made a total of 307 errors. Of the errors, 26 were noticed as linguistic problems at the moment of pro-

duction (noticing the hole coded as problem), and accounted for 8.5%. Noticing the gap through models accounted for 3.9% (12 instances). On the other hand, a total of 255 errors were observed in the recast group. Noticing the hole coded as problem accounted for 7.5% (19 instances) while noticing the gap through recasts was 11.8% (30 instances). Both groups reported noticing holes in their interlanguage during production (8.5% and 7.5%): In other words, through production, the learners noticed what they had not been able to say in English. In both groups, most of the errors were not noticed by the learners (82.1% and 77.7%). One difference between the groups was that the recast group noticed errors through recasts more often than the model group noticed errors through models. A statistical analysis supported this.

The chi-square test revealed that there was a statistically significant difference in the distributions of noticing types between the two groups ( $\chi^2(3) = 13.83, p < .01$ ). A further analysis of the residuals revealed that a significant difference existed only between the noticing of the gap through recasts by the recast group and the noticing of the gap through models by the model group ( $p < .01$ ); no significant differences were found between the two groups in terms of problem, production, or a lack of noticing.

**Table 4. Frequencies of Each Noticing Type**

Noticing types	Model Group ( $n = 8$ )				Recast Group ( $n = 8$ )			
	Total	%	<i>Mdn</i>	<i>IQR</i>	Total	%	<i>Mdn</i>	<i>IQR</i>
Problem	26	8.5	2.0	4.25	19	7.5	2.5	3.00
Production	17	5.5	1.5	1.50	8	3.1	1.0	0.50
Model/recast	12	3.9	1.0	0.50	30	11.8	3.0	2.25
Lack of noticing	252	82.1	31.0	10.00	198	77.7	22.0	15.75
Total	307	100	39.5	9.50	255	100	28.5	15.25

#### *Noticing Types and Learning Effects*

Learning effects in relation to noticing types were measured as repairs in Phase 2. Table 5 indicates the repairs made by the model group. When the learners noticed their linguistic problems during their production in Phase 1, they repeated 57.7% of the errors (see the far left column under

the problem category). There were only two instances of incorporation from the model input. When they noticed their errors through models (see the middle column under the model category), they incorporated the target features from the input and repaired most of the errors (8 of the 12 errors, 66.7%). No instances were observed for the same error category. As for the production category, the results show a relatively high tendency to incorporate the information provided by the feedback (8 of the 17 errors, 47.1%).

**Table 5. Learning Effects in the Model Group as Measured by Performance in Phase 2**

Learning effects	Model Group				Total
	Problem	Production	Model	Lack of noticing	
Incorporation	2 (7.7%)	8 (47.1%)	8 (66.7%)	53 (21.0%)	71 (23.1%)
Self-repair	0 (0.0%)	1 (5.9%)	1 (8.3%)	7 (2.8%)	9 (2.9%)
Same error	15 (57.7%)	5 (29.4%)	0 (0.0%)	136 (54.0%)	156 (50.8%)
Different error	3 (11.5%)	3 (17.6%)	3 (25.0%)	27 (10.7%)	36 (11.7%)
Avoidance	6 (23.1%)	0 (0.0%)	0 (0.0%)	29 (11.5%)	35 (11.4%)
Total	26	17	12	252	307

The learning effects in the recast group are summarized in Table 6. When the learners noticed their linguistic problems by production, they produced the same errors consisting of 42.1% of the errors (see the far left column under the problem category). They incorporated the target features from the recasts, accounting for only 15.8% (three instances). When they noticed their errors through recasts (see the middle column of the recast category), 19 of the 30 errors were repaired (63.3%). They repeated four errors in Phase 2 (13.3%). The results of the Fisher's exact tests showed that there were no statistically significant differences in the distributions of repair in each noticing type between the two groups ( $p = .68$  for problem;  $p = .72$  for production;  $p = .34$  for model/recast;  $p = .88$  for lack of noticing).

**Table 6. Learning Effects in the Recast Group as Measured by Performance in Phase 2**

Learning effects	Recast Group				Total
	Problem	Production	Recast	Lack of noticing	
Incorporation	3 (15.8%)	4 (50.0%)	19 (63.3%)	36 (18.2%)	62 (24.3%)
Self-repair	1 (5.3%)	1 (12.5%)	0 (0.0%)	4 (2.0%)	6 (2.4%)
Same error	8 (42.1%)	3 (37.5%)	4 (13.3%)	115 (58.1%)	130 (51.0%)
Different error	3 (15.8%)	0 (0.0%)	6 (20.0%)	19 (9.6%)	28 (11.0%)
Avoidance	4 (21.1%)	0 (0.0%)	1 (3.3%)	24 (12.1%)	29 (11.4%)
Total	19	8	30	198	255

Because learning effects in each noticing type did not differ statistically between the two groups, I divided the noticing types into two groups as in Table 7 to clarify the differences of learning effects by noticing types. Noticing linguistic problems (the problem category) by the two groups was summed up in the left columns; noticing the gap through the models or recasts (the models/recasts category) by the two groups appeared in the right columns. Six of the 45 errors noticed as linguistic problems at the moment of production were repaired (13.3%); on the other hand, 28 of the 42 errors which had been noticed through the feedback were repaired (66.7%). The difference in repairs was statistically significant between the two noticing types ( $\chi^2(1) = 25.95, p < .01$ ). In other words, noticing a gap through models or recasts led to more repairs in the later production than noticing a hole did.

**Table 7. A Comparison of Learning Effects According to Noticing Types**

Learning effects	Noticing the hole (problem)			Noticing the gap (recast/model)		
	Recast Group	Model Group	Total	Recast Group	Model Group	Total
Repair	4 (21.1%)	2 (7.7%)	6 (13.3%)	19 (63.3%)	9 (75.0%)	28 (66.7%)
Other	15 (78.9%)	24 (92.3%)	39 (86.7%)	11 (36.7%)	3 (25.0%)	14 (33.3%)
Total	19	26	45	30	12	42

## Discussion

The results of this study are summarized as follows.

1. Production led the learners to notice their linguistic problems in the target language. Approximately 10% of the errors (8.5% for the model group; 7.5% for the recast group) were reported as linguistic problems at the moment of speaking (see Table 4).
2. Although both of the feedback types (models and recasts) helped the learners notice the gap between their interlanguage and the target language, recasts, which were provided immediately after their utterances, were more effective in leading to noticing than were models, which were provided a few minutes later, after their utterances (see Table 4).
3. There were no significant differences in learning effects by noticing types between the two groups (see Tables 5 & 6). Models and recasts may have worked in the same way in terms of learning effects.
4. Noticing the gap through feedback (models and recasts) showed a relatively higher effect in leading to repairs than noticing a hole did (see Table 7).

### *Roles for Output*

The findings suggest that although L2 learners notice their linguistic problems by producing the target language, they might have much difficulty in attending to the subsequent input to take the necessary information in. According to the results (Table 4), the learners in this study reported that they had noticed their linguistic problems more frequently than the other noticing types, particularly in the model group. However, in terms of learning effects (Tables 5 & 6), noticing the holes led to no more repairs at the later production than a lack of noticing did. As in the previous studies (Izumi et al., 1999; Izumi & Bigelow, 2000), this study did not lend support to Swain's claim (1993, 1995, 1998) that through the activity of output, learners who notice their linguistic problems in their own interlanguage may have the need to pay closer attention to the subsequent input, resulting in learning. In this study, the input provided to the learners was exact reformulation of the learners' utterances, which

contained the structural information necessary for the learners. Therefore, even if provided with the necessary input, most of the learners did not utilize and incorporate the information in the model input provided in the form of models or recasts.

One might suppose that the learners did not retain what they noticed until the model input was given. Nevertheless, in fact, the time differential between noticing the hole and the subsequent input seems to have made no significant difference. In the recast condition, the input in the form of recast was provided immediately after the learners' utterances; in the model condition, the input in the form of a model was provided a few minutes later. Learning effects between the model condition and the recast condition were almost the same.

One possible explanation for the failure to make use of the subsequent input in later production may have been the developmental readiness of the learners (Mackey & Philp, 1998; Pienemann, 1998; Schmidt, 1990).<sup>13</sup> In other words, the linguistic problems noticed by the learners may have been far beyond their current L2 competence. In Example (10), Saori stated that she had wanted to say, "The girl doesn't know whose watch that is." She came up with the question word *whose*, but did not know how to use the word in the sentence. That is, she faced a problem with subordination. As Pienemann (1998) pointed out, subordination is a construction in a relatively higher level of L2 development, which requires learners to process linguistic information beyond the clauses. It may be that because Saori's English abilities were not yet at the stage of subordination, she had difficulty in paying attention to and utilizing the subsequent input, even though she was aware of her linguistic problems.

(10) Noticing linguistic problems

Tape:           *(She ... she ... um ... she don't ... don't ... she doesn't know ... um ... um ... she doesn't know watch ... whose watch...)*

Saori:           *Well, I didn't come up with the expression of the girl who didn't know whose watch that was. Then, yes, "whose" occurred in my mind. I didn't know how to use "whose." A little bit irresponsibly, I did. Yes.*

Schmidt (1990) also points out that “availability for noticing and stages of L2 development are closely related” (pp. 142-143). The relationship between L2 learners’ readiness and ability to notice is beyond the scope of this study, but is worthy of further investigation.

### *Roles for Feedback*

As to the relative effect of recasts and models on noticing, the results showed that noticing errors through recasts occurred more frequently than through models (Table 4). The difference between the two feedback types was the interval between the utterance and the feedback. As reported previously (Doughty & Varela, 1998; Long, Inagaki, & Ortega, 1998; Mackey & Philp, 1998), recasts make it easier for learners to compare their own utterances with the target structures. In other words, recasts are somewhat more effective in making cognitive comparisons than models are. Although models brought about noticing, the learners had more difficulty in making cognitive comparisons as time passed. This finding may support Long’s (1996) argument:

Negative feedback of this type (i.e., in the form of implicit correction immediately following an ungrammatical learner utterance) is potentially of special utility because it occurs at a moment in conversation when the NNS is likely to be attending to see if a message got across, and to assess its effect on the interlocutor. (p. 429)

As with learning effects, the model group and the recast group showed similar patterns in terms of repairing (see Tables 5 & 6). This may indicate that models functioned in the same way as recasts in terms of learning effects. The two feedback types are common in that they provide L2 learners with negative evidence and positive evidence at the same time. Probably this dual function of models and recasts results in similar learning effects. Also, it is possible that once learners notice errors through recasts or models, the provided feedback is likely to yield similar learning effects. Thus, Schmidt suggests that noticing plays an important role in mediating between the external information and the interlanguage system (Schmidt, 1990, 1993, 1995).

In summary, both recasts and models may similarly lead to repair in later production because of their common dual function of providing negative evidence and positive evidence simultaneously. Recasts, however, seem to enhance L2 learners’ noticing more effectively than

models do. In this study, feedback was provided artificially in a dense way, that is, every time the learners produced an utterance. This study suggests that feedback has the potential to make L2 learners notice the gap in interactions; nevertheless, the findings are not intended to be generalized to more natural situations.

### Conclusions

In conclusion, Japanese EFL learners' verbal reports obtained through retrospective interviews were used to investigate whether they noticed their errors or linguistic problems during communicative tasks and, if so, when they noticed them. This study found that (a) although the learners noticed linguistic problems through the activity of speaking, they had difficulty in attending to and incorporating the subsequent input for the later production; (b) recasts were more effective in prompting noticing of errors than models were; and (c) noticing the gap through the feedback showed a relatively higher effect in leading to repairs than noticing the hole by output did.

Some limitations, however, need to be addressed. First, the proficiency levels of the participants were not available in this study. It cannot be assumed that the participants represented L2 learners of various proficiency levels. Therefore, I do not intend to generalize the findings of this study to all L2 learners, but rather I believe that this study provided some evidence on noticing by means of feedback or the activity of production. Secondly, the participants had been studying in my class at the time of the study. Our teacher-students relationship may have influenced their noticing. It is not clear, however, whether or not the relationship promoted the learners' noticing. Nevertheless, it must be pointed out that because in Japan most learners of English receive oral target language input primarily in the course of interactions in the classrooms, this experimental situation is not so unlike the learning environment of the participants.

Noticing can never be fully captured because it is the learner's internal process. However, if we hypothesize that noticing plays a mediating role between the input and the environment, then it becomes necessary to carry out research on the noticing issue and accumulate data in order to make a generalization. The findings of this study suggest that noticing the gap and noticing the holes may have differential effects on L2 learning.

### Acknowledgments

*An earlier version of this paper was presented at the 1st annual convention of the Japan Society of English Language Education (JASELE), Hiroshima, August 8, 2001. I would like to thank Professor Kinue Hirano of Joetsu University of Education, Ms. Carolyn L. Kaltenbach, as well as the editor and two anonymous reviewers of JALT Journal for their valuable suggestions and comments on an earlier draft of this paper. Any errors or omissions are mine.*

Hideki Sakai is an associate professor at Shinshu University, Faculty of Education. His current research interests include noticing, feedback, and classroom second language acquisition.

### Notes

1. Lyster (1998a) defined recasts as “the teacher’s implicit provision of a correct reformulation of all or part of a student’s ill-formed utterance” (p. 58). Thus, he seems to suppose that the role of interlocutor is limited to a teacher because his study focused on the classroom interactions. In contrast, Long (1996) stated that “*negotiation for meaning*, and especially negotiation work that triggers *interactional* adjustments by the NS [native speaker] or more competent interlocutor, facilitates acquisition.... Negotiation for meaning by definition involves denser than usual frequencies of semantically contingent speech of various kinds (i.e., utterances by a competent speaker, such as repetitions, extensions, reformulations, rephrasings, expansions and recasts), which immediately follow learner utterances and maintain reference to their meaning....” (pp. 451-452). Thus, he seems to consider that recasts are not necessarily provided by a native speaker. Although an anonymous reviewer pointed out the possibility that L2 learners provide recasts to each other, it is not clear whether L2 learners of the same proficiency level can perform recasts of each other’s utterances. This would require empirical study in the future.
2. It should be noted that the definition of recasts varies in studies on the effects of recasts: As Nicholas, Lightbown, and Spada (2001) stated, “one important difference in operational definitions in L2 studies is that recasts have sometimes included more than just a reformulation of a learner’s incorrect utterance” (p. 749). For example,

in Doughty and Varela's (1998) study, recasts were accompanied by repetition of the learner's ill-formed utterance. Doughty and Varela (1998) termed this combination of feedback as "corrective recasting" (p. 123). Mackey and Philp (1998) compared the two conditions between interaction with intensive recasts and interaction without intensive recasts. The former condition was described as "the artificial 'flooding' of interaction with recasts" (p. 353), resulting in 350% more than the latter condition. Long, Inagaki, and Ortega (1998) seemed to follow the definition of recasts as those without any emphasis, but their results were ambiguous. They carried out two studies and found the superiority of recasts over the provision of samples of the target structures, which they called modeling, in only one of the two structures in one of the studies. At present, therefore, it is quite difficult to argue for the effects of recasts alone based on these studies. I am grateful to an anonymous reviewer for pointing this out to me.

3. The learner names in this paper are fictitious.
4. According to Lyster (1998a), confirmation checks are those parts of recasts, which reformulate "all or part of the utterance with rising intonation and no additional meaning" (p. 48). As stated in the Procedure section, recasts were provided with rising intonation in this study, that is, equivalent to confirmation checks.
5. Lyster and Ranta (1997) referred to the learner's response to the feedback provided by the teacher as "uptake" (p. 49) and found that the percentage of uptake of recasts was 31%, which was quite low compared to the other types of feedback: elicitation, clarification request, metalinguistic feedback, explicit correction, and repetition. More relevant to the argument in the present study is Lyster's (1998a) finding that 95% of the non-corrective repetition did not lead to uptake (p. 67). This provides another piece of empirical evidence suggesting that L2 learners respond to recasts and non-corrective repetition differently.
6. Gass and Mackey (2000) considered stimulated recall to be "one subset of a range of introspective methods that represent a means of eliciting data about thought processes involved in carrying out a task or activity" (p. 1). In other words, introspection ranges from concurrent to retrospective think-aloud protocols.

7. I am using the term models in a narrower sense than in Long, Inagaki, and Ortega's (1998) study that compared the effects of recasts and models. In their study, models were pre-determined recorded samples of the target structure and functioned as only positive evidence. As a result, the propositional content of the models was out of the learners' control. In the current study, models were created on the basis of the learners' utterances.
8. Prior to the start of data collection, pilot sessions were carried out with three adult learners in order to develop and refine the experiment procedure. These learners were not included in the analysis of the current study.
9. There have been some controversies as to whether noticing entails awareness or not (Robinson, 1995; Simard & Wong, 2001; Tomlin & Villa, 1994). Tomlin and Villa (1994) argued that attention consists of three separable functions: alertness, orientation, and detection (p. 198). Among the three functions, they considered detection, defined as "the cognitive registration of sensory stimuli" (p. 192), to be the most significant because detected information can be processed further for learning. Moreover, they claimed that "such detection does not require awareness" (p. 199). Robinson (1995), based on memory studies, stated that "noticing can be identified with what is both detected and then further activated following the allocation of attentional resources from a central executive" (p. 297). Schmidt (1993) himself considered noticing to be "related to rehearsal within working memory and the transfer of information to long-term memory, to intake, and to item learning" (p. 213). Although he admitted that detection does not require awareness, Robinson (1995) pointed out that "subliminal exposure effects are unlikely to have effects over intervals longer than a few hundred milliseconds, are rapidly lost from memory, and cannot in any useful sense be claimed to be evidence of learning" (p. 298). In addition, Simard and Wong (2001) reviewed the studies cited by Tomlin and Villa (1994) and called into question Tomlin and Villa's (1994) position that awareness is dissociated from detection. At the moment, it can be agreed that noticing, operationally defined as "availability for verbal report" (Schmidt, 1990, p. 132), is necessary for further processing of learning, that is, SLA.
10. I basically followed the stimulated recall methodology guidelines described in Gass and Mackey (2000). However, I refer to the

recalling session in this study as *the retrospective interview* because guided questions were provided. It should be noted that the guided questions may have brought about validity problems in the recalling procedure, that is, nonveridicality (Gass & Mackey, 2000, p. 107). These problems should be kept in mind while interpreting the results.

11. The participants spoke in Japanese at the interview sessions. In this paper I gave examples of their protocols translated into English. Analysis was carried out on the original protocols.
12. An anonymous reviewer pointed to the problems of identifying errors and stated that “We are told that ‘Baby ask ... what is that’ [as in Example (9)] contains three errors. But ‘what is that’ is perfectly correct if the learner is using direct speech and we have no way of knowing whether the learner intended direct or reported speech.” Indeed it is extremely difficult to identify errors. For this study, I used an external criterion; that is, changes made in the feedback were counted as errors. Thus, the identification of errors was performed objectively in this study. In the case of Example (9), the recast reformulated the learner’s utterance in three ways. Based on the changes made in the recast, the error points were identified. So I did not calculate the inter-rater reliability for the identification of errors. This method of identification of errors may be considered to be valid because the study attempted in part to examine how the learners utilize the information provided through feedback.
13. As an anonymous reviewer pointed out, another possible explanation is that L2 learners might have had difficulty in attending to form and meaning simultaneously (VanPatten, 1990). However, it should be noted that in the second picture description task, the students described the pictures that they had already worked on in the first picture description task. Thus, the second task was not new to L2 learners in a true sense. This probably lessened cognitive load in the second task.

## References

- de Bot, K. (1996). The psycholinguistics of the output hypothesis. *Language Learning*, 46, 529-555.
- Doughty, C. (1994). Fine-tuning of feedback by competent speakers to language learners. In J. Alatis (Ed.), *GURT 1993: Strategic interaction and language acquisition* (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). Cambridge: 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). Cambridge: Cambridge University Press.
- Ellis, R. (1994). A theory of instructed second language acquisition. In N. Ellis (Ed.), *Implicit and explicit learning of languages* (pp. 79-114). San Diego, CA: Academic Press.
- Gass, S. M. (1988). Integrating research areas: A framework for second language studies. *Applied Linguistics*, 9, 198-217.
- Gass, S. M. (1997). *Input, interaction, and the second language learner*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Gass, S. M., & Mackey, A. (2000). *Stimulated recall methodology in second language research*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Hatch, E., & Lazaraton, A. (1991). *The research manual: Design and statistics for applied linguistics*. New York: Newbury House.
- Izumi, S., & Bigelow, M. (2000). Does output promote noticing and second language acquisition? *TESOL Quarterly*, 34, 239-278.
- Izumi, S., & Bigelow, M. (2001). Methodological and theoretical issues in testing the effects of focus on form. *TESOL Quarterly*, 35, 181-189.
- Izumi, S., Bigelow, M., Fujiwara, M., & Fearnow, S. (1999). Testing the output hypothesis: Effects of output on noticing and second language acquisition. *Studies in Second Language Acquisition*, 21, 421-452.
- Leow, R. P. (1997). Attention, awareness, and foreign language behavior. *Language Learning*, 47, 467-505.
- Long, M. H. (1991). Focus on form: A design feature in language teaching methodology. In K. de Bot, R. B. Ginsberg, & C. Kramsch (Eds.), *Foreign language research in cross-cultural perspective* (pp. 39-52). Amsterdam: John Benjamins.
- Long, M. H. (1996). The role of the linguistic environment in second language acquisition. In W. C. Richie & T. K. Bhatia (Eds.), *Handbook of second language acquisition* (pp. 413-468). New York: 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.
- Long, M. H., & 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). Cambridge: Cambridge University Press.
- Lyster, R. (1998a). Recasts, repetition, and ambiguity in L2 classroom discourse. *Studies in Second Language Acquisition*, 20, 51-81.
- Lyster, R. (1998b). Negotiation of form, recasts, and explicit correction in relation to error types and learner repair in immersion programs. *Language Learning*, 48, 183-218.

- Lyster, R., & Ranta, L. (1997). Corrective feedback and learner uptake: Negotiation of form in communicative classrooms. *Studies in Second Language Acquisition*, 19, 37-66.
- Mackey, A., Gass, S., & McDonough, K. (2000). How do learners perceive interactional feedback? *Studies in Second Language Acquisition*, 22, 471-497.
- Mackey, A., & Hilp, J. (1998). Conversational interaction and second language development: Recasts, responses, and red herrings? *Modern Language Journal*, 82, 338-356.
- Nakamura, A. (1995). *CD tsuki jidoeiken kuria mondaishu*. Tokyo: Seibido.
- Nicholas, H., Lightbown, P. M., & Spada, N. (2001). Recasts as feedback to language learners. *Language Learning*, 51, 719-758.
- Oliver, R. (1995). Negative feedback in child NS/NNS conversations. *Studies in Second Language Acquisition*, 17, 459-481.
- Pienemann, M. (1998). *Language processing and second language development: Processability theory*. Amsterdam: John Benjamins.
- Roberts, M. A. (1995). Awareness and the efficacy of error correction. In R. Schmidt (Ed.), *Attention & awareness in foreign language learning* (pp. 163-182). Honolulu: University of Hawaii at Manoa.
- Robinson, P. (1995). Attention, memory, and the "noticing" hypothesis. *Language Learning*, 45, 283-331.
- Sakai, H. (2002). Error types and noticing: Roles for feedback and L2 learner repair. *Bulletin of Joetsu University of Education*, 21, 741-755.
- Schmidt, R. (1990). The role of consciousness in second language learning. *Applied Linguistics*, 11, 129-158.
- Schmidt, R. (1993). Awareness and second language acquisition. *Annual Review of Applied Linguistics*, 13, 206-226.
- 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* (pp. 1-63). Honolulu: University of Hawaii at Manoa.
- Schmidt, R., & Frota, S. (1986). Developing basic conversational ability in a second language: A case study of an adult learner of Portuguese. In R. Day (Ed.), *Talking to learn: Conversation in second language acquisition* (pp. 237-326). Rowley, MA: Newbury House.
- Simard, D., & Wong, W. (2001). Alertness, orientation, and detection: The conceptualization of attentional functions in SLA. *Studies in Second Language Acquisition*, 23, 103-124.
- Skehan, P. (1998). *A cognitive approach to language learning*. Oxford: Oxford University Press.
- Swain, M. (1985). Communicative competence: Some roles of comprehensible input and comprehensible output in its development. In S. M. Gass & C. G. 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. *The Canadian Modern Language Review*, 51, 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 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). Cambridge: Cambridge University Press.
- Swain, M., & Lapkin, S. (1998). Interaction and second language learning: Two adolescent French immersion students working together. *Modern Language Journal*, 82, 320-337.
- Tomlin, R., & Villa, H. (1994). Attention in cognitive science and second language acquisition. *Studies in Second Language Acquisition*, 16, 183-203.
- Truscott, J. (1998). Noticing in second language acquisition: A critical review. *Second Language Research*, 14, 103-135.
- Van den Branden, K. (1997). Effects of negotiation on language learners' output. *Language Learning*, 47, 589-636.
- VanPatten, B. (1990). Attending to form and content in the input: An experiment in consciousness. *Studies in Second Language Acquisition*, 12, 287-301.
- Whitlow, J. (2001). Comments on Shinichi Izumi and Martha Bigelow's "Does output promote noticing in second language acquisition?": Some methodological and theoretical considerations. *TESOL Quarterly*, 35, 177-181.

(Received July 3, 2002; revised May 31, 2003)

## Appendix

### Sample Picture Cards Used (Revised from Nakamura, 1995)

