

Do Recasts Promote Noticing the Gap in L2 Learning?

Hideki Sakai

Shinshu University, Japan

caff@nifty.com

Bio Data:

Hideki Sakai is an associate professor in the Faculty of Education at Shinshu University, where he teaches undergraduate and graduate courses in TEFL methodology and second language acquisition. His current research interests include second language interactional studies, classroom second language acquisition, and psycholinguistic aspects of listening and speaking.

Abstract

This paper reports the effects of implicit negative feedback in the form of recasts on noticing. Twenty Japanese-speaking learners of English were assigned to an experimental group (the recast group, $n = 10$) or a control group (the no-feedback group, $n = 10$). They engaged in an information gap task, during which the recast group received recasts to their erroneous utterances or non-corrective repetition of their targetlike utterances, whereas the no-feedback received no corrective feedback. The participants' verbal reports about noticing about their errors were elicited through stimulated recall. Both groups noticed errors or problems at the moment of production about 10% of the time. Further, the recast group noticed their errors through recasts 21.6% of the time. Results suggest that recasts alone did help L2 learners notice their errors.

Keywords: implicit negative feedback, recasts, noticing, stimulated recall, negative evidence

Introduction

Over the past few decades, conversational interaction and its effects on second language (L2) learning have drawn considerable attention in the field

of language acquisition (e.g., Mackey, 2007). According to Long's (1996) interaction hypothesis, "environmental contributions to acquisition are mediated by selective attention and the learner's developing L2 processing capacity" (p. 414). In other words, conversational interaction may facilitate L2 learning because it constitutes a context in which L2 learners' attention is directed toward some formal aspects of the language. In particular, Long argued that "Negative feedback obtained during negotiation work or elsewhere may be facilitative of L2 development, at least for vocabulary, morphology, and language-specific syntax, and essential for learning certain specifiable L1-L2 contrasts" (p. 414). Recasts, on which this study focuses, are one such type of negative feedback. By utilizing a stimulated recall technique to elicit introspective data, the present experimental study addresses the issue of how L2 learners interpret recasts for the learning of morphology, in particular, irregular past tense forms.

Recasts as Positive and Negative Evidence

According to Nicholas, Lightbown, and Spada (2001), recasts are "utterances that repeat a learner's incorrect utterance, making only the changes necessary to produce a correct utterance, without changing the meaning" (pp. 732-733). Recasts are considered to provide the linguistic data of both positive evidence (i.e., what is grammatical in the target language) and negative evidence (i.e., what is ungrammatical in the target language) at the same time. The following shows an example of a recast, taken from the data gathered for the current study.

(1) An Example of a Recast

Researcher: *What did he do at four in the evening?*

Koichi: *He shaked the bottle.* [an erroneous utterance]

Researcher: *He shook the bottle?* [a recast]

Koichi: *Yes.*

The learner, Koichi (a pseudonym), produced an overgeneralized past form of *shake*, which was followed by a recast in the next researcher's turn. The recast maintained the proposition of the original utterance but contained a reformulation of the erroneous element into the correct one, *shook*. In this

example, the recast was supposed to inform the learner of both positive evidence (i.e., the correct form was *shook*) and negative evidence (i.e., the form *shaked* was ungrammatical in English). The semantic and discoursal properties of recasts that repeat the information generated by learners and that are juxtaposed with the erroneous utterances make it easier for L2 learners to make cognitive comparisons between their interlanguage and the target language (Long, 1996, 2007; see also Saxton, 1997, 2000).

Thus far, a number of descriptive studies have shown that recasts are the most frequent negative feedback types (e.g., Doughty, 1994, Lyster, 1998a, 1998b; Lyster & Ranta, 1997; Nassaji, 2007; Oliver, 1995; Panova & Lyster, 2002; Roberts, 1995; Van den Branden, 1997). Further, recasts have been found to be effective for L2 learning (e.g., Ayoun, 2001; Carroll & Swain, 1993; Doughty & Varela, 1998; Ellis, 2007; Han, 2002; Ishida, 2004; Iwashita, 2003; Loewen & Nabei, 2007; Long, Inagaki, & Ortega, 1998; Mackey, 1995, 1999; Mackey & Philp, 1998; for a meta-analysis, see Mackey & Goo, 2007). Recently, L2 researchers have addressed the issue of how recasts benefit L2 learning (Gass & Mackey, 2007; Leeman, 2003; Long, 2007). One probable reason for much interest in recasts is that if the mechanisms underlying the functions of recasts are made clearer, recasts will be utilized more effectively in L2 classrooms. Recasts are considered to be implicit in that they do not provide an overt indication of the existence of errors to the learner (Ellis, 2007; Ellis, Loewen, & Erlam, 2006). Thus, they have been argued to provide negative evidence while avoiding interrupting the flow of communication (Long, 2007; Long, Inagaki, & Ortega, 1998).

However, at the same time, the very implicitness of recasts as negative feedback may bring about one problem for L2 learning. For example, Lyster (1998a) argued that recasts are ambiguous as negative feedback because recasts and non-corrective repetition, that is to say, repetition of learners' correct utterances, are difficult to distinguish from each other. On the other hand, Long (1996) argued that L2 learners do notice the corrective nature of recasts. As indirect evidence, research has shown that L2 learners respond to recasts and non-corrective repetitions differently (e.g., Doughty, 1994; for a review of L1 studies, see Long, 1996). For example, Doughty (1994) reported that L2 learners of French imitated recasts more frequently than the non-corrective repetition (21.5% vs 2.3%). However, it has been pointed out that differing responses to recasts and non-corrective repetitions do not

always imply that L2 learners perceive recasts as negative feedback (e.g., Mackey & Philp, 1998). Recently, a more direct introspective technique (i.e., stimulated recall) has been employed in order to elicit noticing data from the learners (Egi, 2007a, 2007b; Mackey, Gass, & McDonough, 2000; Nabei & Swain, 2002; Roberts, 1995).

Recasts and Noticing for L2 Learning

Stimulated recall is “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” (Gass & Mackey, 2000, p. 1). In stimulated recalls, learners are asked to report what they were thinking at the time of the target event. In other words, stimulated recalls are retrospective, not concurrent, introspective methods in that there is some interval between the time of the event and the time of verbal reports. The data elicited through stimulated recall concern what Schmidt (1990) called *noticing*, that is to say, only a subset of detected information in Tomlin and Villa’s (1994) sense. Tomlin and Villa (1994) defined *detection* as “the cognitive registration of sensory stimuli” (p. 192) and argued that detection does not entail awareness (p. 199). On the other hand, Schmidt (1990) defined *noticing* as “the availability for verbal report” (p. 132). Because stimulated recall is an elicitation of verbal reports, not every item of detected information will be elicited through stimulated recall. (For a discussion of theoretical issues of noticing and awareness, see Robinson, 1995; Simard & Wong, 2001.)

Previous studies using stimulated recall have showed that L2 learners had difficulty in perceiving the corrective intention in recasts precisely, in particular, when the recasts concerned grammatical-syntactic errors (Gass & Lewis, 2007; Mackey, Gass, & McDonough, 2000; Roberts, 1995). Recently, Egi (2007a) made an attempt to explore how L2 learners interpret recasts. She investigated the relationship between L2 learners’ interpretation of recasts and factors such as linguistic targets, recast length, and number of changes. Her participants were 49 learners of Japanese. They performed two communicative tasks with native speakers in dyads. Of the 49, 31 performed the tasks under immediate recall conditions; the other 18 attended a stimulated recall session after the task. Their recall protocols were categorized as (a) responses to content, (b) negative evidence, (c) positive evidence, and (d) negative and positive evidence. Responses to content were

defined as “comments indicating that the learner interpreted the recast as the interlocutor’s replies to the conversational content” (p. 522). Negative evidence was operationalized as “comments indicating that the learner was aware of error production or correction without a clear indication that the learner noticed the targetlike form provided in the recast” (p. 524). Positive evidence referred to “comments indicating that the learner attended to the targetlike model without reporting having produced errors or having been corrected” (p. 534). The last category, negative and positive evidence, was operationalized as “comments that indicated the learner’s (a) recognition of error production or correction and (b) attention to the targetlike form in the recast” (p. 524). Of the 476 morphosyntactic feedback episodes, her participants interpreted recasts as responses to content (21.22%), negative evidence (34.03%), positive evidence (18.70%), and negative and positive evidence (26.05%); of the 77 lexical feedback episodes, they interpreted recasts as responses to content (18.18%), negative evidence (33.77%), positive evidence (18.18%), and negative and positive evidence (29.87%). No statistical difference was found between the linguistic targets and the interpretations of recasts. On the other hand, the recast length and the number of changes were found to affect the participants’ interpretations of recasts. Shorter recasts helped L2 learners interpret the recasts as negative and positive evidence more frequently than longer recasts; recasts with fewer changes were more effective for learners to interpret them as negative and positive evidence than those with more changes. What is important in Egi’s study is that she developed a categorization system for coding verbal protocols on the basis of types of linguistic data (negative evidence, positive evidence, or both). However, her system still seems to be insufficient to address the issue of whether or how recasts facilitate learners’ noticing.

In the literature, noticing has been categorized mainly into (a) noticing a form in the input, (b) noticing one’s interlanguage deficiencies (or a hole), and (c) noticing the gap between the interlanguage and the target language (Doughty & Williams, 1998, p. 228; Swain, 1998, p. 66). In her output hypothesis, Swain (1985, 1993, 1995) proposed that output in the target language may promote the second type of noticing: “the activity of producing the target language may prompt second language learners to consciously recognize some of their linguistic problems” (1995, p. 126). Thus, it is necessary to distinguish noticing one’s errors through the activity

of production from noticing errors through recasts. In Example (1) mentioned above, Koichi produced an overgeneralized form *shaked* for the past tense. In the case that, at the moment of production, he noticed his error (that *shaked* was wrong), the role of the recast may have been only to provide positive evidence about how to say the past form of the verb. By contrast, in the case that he did not recognize his error at the moment of production, the recast may have had the dual function of indicating that *shaked* was wrong and of showing the correct form. In Egi's (2007a) categorization system, no clear distinction was made between L2 learners' noticing of errors by themselves and L2 learners' noticing of errors with the help of recasts. For example, she provided an example of the verbal protocol coded as negative and positive evidence as follows: "I had no idea how to say this verb, so I just said *hayai* ["fast"], which turned out to be wrong, and then you told me how to say 'hurry,' *isoide*" (p. 524). The first part of the comment indicates that this learner had already recognized her problem at the moment of production. Thus, the recast may not have provided the learner with negative evidence; rather, it may have functioned only as positive evidence. In other words, Egi's categorization did not make it clear whether recasts facilitate noticing errors or not.

To sum, although several studies (Gass & Lewis, 2007; Mackey, Gass, & McDonough, 2000; Roberts, 1995) pointed out the difficulty of noticing the corrective intention in recasts, Egi's (2007a) detailed analysis of verbal protocols about L2 learners' interpretations of recasts suggests that recasts are indeed interpreted as negative evidence (accounting for 34.03% for morphosyntactic feedback episodes and 33.77% for lexical feedback episodes) or negative and positive evidence (accounting for 26.05% for morphosyntactic feedback episodes and 29.87% for lexical feedback episodes). However, because there have so far been a very limited number of studies on L2 learners' interpretations of recasts, further research is needed. Furthermore, as pointed out earlier, in order to address the function of recasts as negative feedback, it is necessary to isolate the effects of recasts from those of the activity of output itself. Thus, the present study was designed to separate the confounding effects of recasts and language production. This study involved using a no-feedback group as a comparison group. By comparing the recast group with the no-feedback group, this study attempted to examine the effects of recasts, excluding the noticing effects of

production. This study is significant in that by utilizing a stimulated recall technique, it attempted to explore the role of recasts in details.

The research question posited for this study was: Do recasts alone provide negative evidence about the grammaticality of L2? In other words, this study examined whether L2 learners, who did not notice their linguistic problems or errors at the moment of production, notice their errors through recasts. On the basis of Egi (2007a), it was hypothesized that the recast group will notice more errors in production than the no-feedback group.

Method

Participants

The participants were initially 26 Japanese university student volunteers from a required general English course at a national university in Japan. Since the results of an exit questionnaire indicated that six participants had experience abroad and/or had started to learn English prior to entrance to junior high school, those six participants were eliminated from the analysis of the current study in order to make the two groups comparable in terms of prior L2 learning experiences. Finally, the number of participants was 20 (10 male and 10 female). The twenty participants were randomly assigned to the recast group ($n = 10$) or the no-feedback group ($n = 10$) as a control group. Table 1 shows the pertinent information (sex, age, and years of learning English) about the participants. All of them started to learn English in their 1st year of junior high school, that is to say, in the 7th grade. Prior to the experiment, they had had received about 7 years of formal instruction in English at junior and senior high school as well as at university.

Table 1

Information about Participants (Sex, Age, and Years of Learning English)

Recast Group ($n = 10$)				No-Feedback Group ($n = 10$)			
Learner	Sex	Age	English study	Learner	Sex	Age	English study
Tomoko	F	20;01	7;11	Maki	F	20;05	7;11
Yuki	F	20;00	7;11	Ayumi	F	20;05	7;11
Rinko	F	21;07	8;11	Chitose	F	18;04	6;03
Yuko	F	18;04	6;04	Kaori	F	19;06	7;03

Momo	F	19;02	6;04	Shuko	F	19;03	6;04
Toshiya	M	20;07	7;11	Taka	M	19;01	6;11
Yusuke	M	20;03	7;11	Ichiro	M	18;06	6;04
Taro	M	20;09	7;11	Seiji	M	18;11	6;04
Akira	M	20;03	7;11	Hideki	M	18;09	6;04
Koichi	M	19;10	7;11	Kotaro	M	18;09	6;04
<i>M</i>		20;01	7;08	<i>M</i>		19;02	7;00

Note. All names are pseudonyms.

Target Grammatical Structure

The target structure chosen for this study was irregular past tense forms for mainly two reasons. First, several researchers had dealt with this structure in studies on the effects of interaction (e.g., Ellis, 1987; Nobuyoshi & Ellis, 1993; Takashima, 1995). Second, it is well pointed out that L2 learners may produce overgeneralized forms, that is, *-ed* forms, instead of the appropriate irregular past forms (e.g., Dulay, Burt, & Krashen, 1982, p. 158; Lightbown & Spada, 1993, p. 56). In order to investigate whether recasts provide negative evidence for the learner, it is necessary to create a condition in which the learner will produce interlanguage forms different from the targetlike forms. For example, an L2 learner produces *drinked* as the past form of the verb *drink* and is provided with a recast including the targetlike form *drank*. This case may enable researchers to examine whether the learner will not only notice the correct form *drank*, but also recognize that the overgeneralized form *drinked* is not acceptable.

Prior to this study, another eleven university students who had similar backgrounds to the participants in this study were asked to write the past forms of verbs. This pilot study yielded ten verbs for which most learners produced generalized forms: *feed*, *beat*, *cast*, *strike*, *draw*, *swing*, *bite*, *shake*, and *fight*. These ten words were then used as the target verbs for the tasks used in the main study.

Procedure

The procedure of the study is summarized in Table 2. The participants in both groups took part in the experiment individually. The experiment consisted of three tasks, in which they communicated with the researcher, and a stimulated recall session. First, the participants performed Task 1,

Yuki’s Story, which served as a pretest. Then, they performed Task 2, Ken’s Story, the treatment task. In Task 2, the participants in the recast group received recasts of their erroneous utterances or non-corrective repetition of their targetlike utterances from the researcher. On the other hand, the participants in the no-feedback group received no corrective feedback or positive evidence. Then, both groups carried out Task 3, Yuki’s Story,¹ followed by the stimulated recall session, in which participants were asked about what they had thought during Task 2. The interval between Task 2 (Ken’s story) and the stimulated recall was about 10 minutes.

Table 2: Procedures

	Time	Recast group	No-Feedback group
Explanation	5 min	- Consent form	- Consent form
Task 1 (pretest)	10 min	- Yuki’s Story	- Yuki’s Story
Task 2 (treatment task)	10 min	- Ken’s Story received recasts or non-corrective repetition	- Ken’s Story received no feedback but received some positive encouragement like “Yes” and “OK.”
Task 3 ^a	10 min	- Yuki’s Story	- Yuki’s Story
Verbal reports	15 min	- Stimulated recall	- Stimulated recall
Questionnaire	10 min	- Exit questionnaire	- Exit Questionnaire

a. The results of Task 3 are not reported in the current study (see Note 1).

Thus, in this study, the independent variable was the provision of recasts in Task 2; the dependent variable was obtained from the verbal protocols in the stimulated recall.

Tasks

To serve as a pretest, a story-making task was provided before the treatment

session. In this task, the participants were asked to make a story about what Yuki did. First, they were given a chart that contained a time line from 8:00 to 22:00 and ten small picture cards that depicted the actions of the target verbs, which had been selected through a pilot test. They were told to shuffle the picture cards and to place each card into a time slot on the chart with the card faces down. Then, they were allowed to turn the cards over. They were told to narrate in English what Yuki (a female character) did yesterday. The planning time of 1 minute was given, during which they were not allowed to make written notes. Then, they were told to start with the sentence, “Yuki got up at seven yesterday,” so that the past time reference was obligatory. Through this task, the participants’ production data was obtained.

The treatment task was an information gap task, Ken’s Story. The participants were given a timeline and ten small picture cards (see Appendix). The pictures depicted the actions of ten target verbs, although the actions were different from those of Yuki’s Story. For example, a picture for the verb *shake* showed that Yuki was shaking hands in Yuki’s Story and that Ken (a male character) was shaking a bottle in Ken’s story. First, they were told to shuffle the picture cards and place them on the chart. Then, they were instructed to turn the cards over. The researcher held the same set of ten picture cards and the time chart and asked the participants, “What did Ken do at ten in the morning?” so that the cards could be placed in the same order as on the participants’ chart. For the recast group, responses to the participants’ elicited utterances in the form of recasts were provided when the participants made an error or in the form of non-corrective repetitions when the participants’ utterances were targetlike.² The reason for the provision of non-corrective repetition was that, without repetition of correct utterances, the participants would notice the corrective nature of recasts easily. Participants in the no-feedback group were provided with responses like “Yes.” “I see.” and “OK.” to every utterance. Thus, no information about the grammaticality of the target structures was given to the no-feedback group. The following examples illustrate the two conditions.

(2) The Recast Group

Yusuke: *He ... he swunged the bat.* [an erroneous utterance]
 Researcher: *Ah, he swung a bat?* [a recast]
 Yusuke : *Yes.*

(3) The No-Feedback Group

Kaori: *He shaked a bottle.* [an erroneous utterance]

Researcher: *Um. OK. I got it.*

Thus, the participants were provided with almost the same number of opportunities to produce the target structures. However, there was no provision to control for opportunities for the participants to react to the feedback.

Stimulated Recall

The stimulated recall technique was employed as a method of eliciting noticing data.³ The instruction of the stimulated recall was created on the basis of Egi (2004) and Gass and Mackey (2000). The participants were asked to report what they had been thinking about for every utterance. As prompts to stimulate their recall of noticing during the interactions, they listened to audiotaped interactions of the treatment session and looked at the pictures and timeline. The recall session was conducted in Japanese, the L1 for the participants and the researcher.

Analysis

The researcher transcribed the participants' performance in all the tasks and their verbal protocols in the stimulated recall. A graduate assistant checked all the transcriptions. The following analyses were based on the transcriptions.

The production of the past tense forms in Task 1 was analyzed as a measure of knowledge of the target structures. More specifically, the accuracy rate of the past tense forms was calculated using two methods, following Pica (1983; see also Takashima, 1995, pp. 72-75): supplied in obligatory context (SOC) analysis and target-like use (TLU) analysis. The formula for SOC analysis is: [(the number of correct supplings in obligatory contexts x 2) + (the number of misformations in obligatory contexts x 1)] / (total obligatory contexts x 2). The formula for TLU analysis is: (the number of correct supplings in obligatory contexts) / [(the number of obligatory contexts) + (the number of supplings in non-obligatory contexts)]. This study reports type counts, which are not influenced by repetitive use of the same words.

To obtain the indices of the SOC and TLU analyses, the researcher coded the participants' utterances into (a) correct use of the past tense form (Correct), (b) use of wrong past forms (Misformation), (c) non-use of past tense forms in the obligatory contexts (Incorrect), and (d) wrong use of past tense forms in contexts that do not require past tense forms (Overuse). A graduate student assistant coded 25.0% of the data, and the intercoder agreement reached 96.4% (269 out of the 279 instances).

To code the verbal reports in the stimulated recall session, an individual analysis proceeded in two steps: (a) analyzing the participants' production in the treatment task (Task 2, Ken's Story) and (b) coding their verbal reports. First, the production in the treatment task was coded into five categories: correct, misformation, incorrect, overuse, and other. The category *Other* refers to cases in which the participants were not provided with recasts containing their verbs, as in the following example.

(4) The Category *Other*

Researcher: *What did he do at three in the afternoon?*

Toshiya : *Horse he take a carrot for horse.* [Other]

Researcher : *Ah, he fed a carrot to the horse?* [a recast]

Toshiya : *Yes.*

Recasts like this example contained corrective information about the past tense reference and the choice of verb. In this study, such cases were excluded from the analysis. This analyzing procedure of Task 2 was similar to that of Task 1; the former was intended to be used for a further analysis of verbal protocols, whereas the latter was utilized for the calculation of accuracy rate of the past tense forms. Another assistant coded all the productions in Task 2 as well. The intercoder agreement was 95.0% (266 out of the 280 instances). The disagreements were resolved after discussion.

Second, the verbal reports regarding the participants' non-targetlike use of past tense forms (that is, Misformation and Incorrect) were analyzed and coded according to four categories as shown in Table 3 (see also Sakai, 2004). These categories were based on the relevant literature (Doughty & Williams, 1998, p. 228; Swain, 1998, p. 66). First, their reports about noticing were divided into (a) noticing a hole (noticing what they cannot express in their interlanguage system) and (b) noticing the gap (noticing that

what they say is different from the target language). The latter type of noticing was further divided into noticing a gap (a) at the moment of production and (b) through a recast.

Table 3
Coding Categories for Verbal Protocols

Category	Noticing Type	Definition
Problem	Noticing a hole	- make comments about what they cannot express or what they only partially know
Production	Noticing a gap at the moment of production	- notice errors in their utterance, or that their utterance is wrong, at the moment of speaking
Recast	Noticing a gap through a recast	- notice errors, or that their utterance is wrong, by receiving a recast
Unnoticing	Noticing at the interview / No report	- notice errors, or that their utterance is wrong, at the moment of the verbal report, or fail to make a report

The following are examples for each category. The original comments of the stimulated recall were in Japanese.

(5) Noticing a hole (coded as *Problem*)

Task 2: Ken's Story

Momo: *He dro.. drinken coffee. He have a coffee.* [an erroneous utterance]

Researcher: *He drank coffee?* [a recast]

Momo: *Yes.*

Researcher: *OK.*

Stimulated Recall:

Momo: *Well, about this, I stopped because I was thinking*

of what the past form of the word drink was. Actually, I wanted to say a cup, full of coffee, but I did not know the expression. That was before speaking. Yes. I did not know the past form of the word drink before starting to speak.

(6) Noticing a gap at the moment of production (coded as *Production*)

Task 2: Ken's Story

Researcher: *What did he do at six in the evening?*

Taro: *Ah, he taken care horse.* [an erroneous utterance]

Researcher: *He took care of a horse?* [a recast]

Stimulated Recall:

Taro: *Well, about this, I wanted to express "feeding" when I saw the picture. But I didn't know how to say it. Ah, so while I was thinking, I came up with the expression "take care of," so I thought I would say it instead. I don't know why, but I said taken for took. This may be because I was upset under pressure to say something. In my brain, I wanted to say "took care" correctly. But I said "taken care," I don't know why.*

Researcher: *When did you notice about taken?*

Taro: *After saying it, I thought I would correct it. But at that time you responded. So I thought it was OK. That's all.*

(7) Noticing a gap through a recast (coded as *Recast*)

Task 2: Ken's Story

Researcher: *Next what did he do at eight in the evening?*

Yuki: *He shaked umm he shaked a ...* [an erroneous utterance]

Researcher: *Ah. He shook a bottle?* [a recast]

Yuki: *Yes.*

Researcher: *OK. I got it.*

Stimulated Recall:

Yuki: *Well, let's see, I thought the past form of shake was shook, and, and, I did not come up with the word bottle. I thought that I knew it, but I did not remember it. That's all.*

Researcher: *Yes. When did you notice about shook?*

Yuki: *Well, when I was told by you, I thought that was it.*

(8) Noticing at the interview or no report (coded as *Unnoticing*)

Task 2: Ken's Story

Yusuke: *And ... he, he swunged the bat.* [an erroneous utterance]

Researcher: *Ah, he swung a bat?* [a recast]

Yusuke: *Yes.*

Researcher: *OK. I got it.*

Stimulated Recall:

Yusuke: *Well, before starting to speak, I thought of a word to mean "swinging a bat." That's all.*

The researcher coded all the verbal reports. The assistant who coded the participants' production in Task 2 and another assistant coded the verbal reports obtained from the recast group and the no-feedback group, respectively. The intercoder agreements were 96.1% (49 out of the 51 instances) and 98.7% (76 out of the 77 instances). Disagreements were resolved after discussion.

Statistical Analysis

First, non-parametric Mann-Whitney tests and chi-square test were performed to check the comparability of the two groups in terms of the accuracy of the past tense forms and the distribution of past tense forms.

To examine whether the two groups differed for each category of noticing, three non-parametric Mann-Whitney tests were performed with Holm's sequential Bonferroni method⁵ to control for Type 1 error.

Results

Comparability of the Two Groups

The participants were randomly assigned to one of the two conditions; so to check the comparability of the two groups, the accuracy of the past tense forms in the pretest (Task 1) was examined. The results of SOC were $Mdn = .65$, $IQR = .46$ for the recast group and $Mdn = .64$, $IQR = .40$ for the no-feedback group; the results of TLU were $Mdn = .54$, $IQR = .55$ for the recast group and $Mdn = .53$, $IQR = .45$ for the no-feedback group.⁴ Non-parametric Mann-Whitney tests did not detect statistically significant differences in the accuracy of the past tense forms in Task 1 between the two groups ($p = .869$ for the SOC analysis; $p = .782$ for the TLU analysis).

Table 4 shows the distribution of each category of past tense forms for the two groups. The chi-square test revealed no statistically significant difference between the two groups ($\chi^2(2) = 0.658$, $p = .720$). In other words, the recast group and the no-feedback group did not differ in the production of past tense forms on Task 2. Therefore, these results suggest that the participants in the two groups had relatively similar proficiency levels in terms of past tense forms.

Table 4

Distribution of Each Category for the Past Tense Forms in Task 2

	Recast Group		No-Feedback Group	
Correct	64	(48.5%)	71	(48.0%)
Incorrect	50	(37.9%)	61	(41.2%)
Misformation	18	(13.6%)	16	(10.8%)
Total	132	(100%)	148	(100%)

Note. For the recast group, the number in the category *Other* was 20 out of 132 (3 for Correct, 14 for Incorrect, and 3 for Misformation).

Verbal Reports

Table 5 shows the frequencies of each noticing category for the two groups (see also Figure 1). For the recast group, the category *Unnoticing* accounted for 70.5% (36 out of the 51 instances); for the no-feedback group, this category was observed more frequently (89.6%, 69 out of the 77 instances). Table 6 indicates that the medians of frequency rates of *Unnoticing* for the recast and the no-feedback groups were .63 (i.e., 63%, $IQR = .36$) and 1.00 (i.e., 100%, $IQR = .18$) respectively. Tables 5 and 6 show that for both the recast and the no-feedback groups, the categories *Problem* and *Production*

were observed in similar percentages (7.9%, 4 instances for the recast group; 10.4%, 8 instances for the no-feedback group) and that the medians of frequency rates of *Problem* and *Production* were both .00 (*IQRs* = .14 and .18 respectively). For the recast group, noticing a gap though a recast accounted for 21.6% (11 out of the 51 instances, *Mdn* = .27, *IQR* = .50). Thus, the results suggest (a) that since both groups reported *Problem* and *Production* to a similar degree, the participants noticed errors or what they could not express in English irrespective of the provision of recasts, and (b) that since the recast group noticed errors at a higher rate than the no-feedback group, the provision of recasts facilitated the participants' noticing greatly. In other words, the main differences between the two groups came about by means of noticing through recasts.

Table 5
Distribution of Noticing Types by Group

Verbal Report	Recast Group				No-Feedback Group			
	Incorrect	Mis	Total		Incorrect	Mis	Total	
Problem	0	1	1	(2.0%)	2	0	2	(2.6%)
Production	2	1	3	(5.9%)	5	1	6	(7.8%)
Recast	5	6	11	(21.6%)	0	0	0	(0.0%)
Unnoticing	29	7	36	(70.5%)	54	15	69	(89.6%)
Total	36	15	51	(100.0%)	61	16	77	(100.0%)

Note. Mis = Misformation.

Table 6
Medians and Interquartile Ranges of Frequency Rates of Noticing Types by Group

Verbal Report	Recast Group		No-Feedback Group	
	<i>Mdn</i>	<i>IQR</i>	<i>Mdn</i>	<i>IQR</i>
Problem/Production ^a	.00	.14	.00	.18
Recast	.27	.50	.00	.00
Unnoticing	.63	.36	1.00	.18

Note. Frequency rates were obtained by dividing the frequency for each noticing type by the total number of noticing. They were calculated for each participant. Then, medians and interquartile ranges of frequency rates were calculated for each group.

a. The frequency rates of *Problem* and *Production* were combined for a further statistical analysis that tests the effects of output.

These findings were supported by statistical analyses. First, statistically significant differences were obtained between the two groups for the *Recast* category and the *Unnoticing* category ($p = .008$ for *Recast* and $p = .022$ for *Unnoticing*). Thus, the hypothesis that the recast group will notice more errors in production than the no-feedback group was confirmed. The comparison of the two groups for *Problem* and *Production* was not significant ($p = .720$). Therefore, the participants in the two groups noticed their errors or what they could not express in English to a similar degree.

Moreover, Table 5 provides breakdowns by error types (Incorrect and Misformation). Regarding the results for Misformation, the recast group noticed 8 of the 15 instances, whereas the no-feedback group noticed only 1 of the 16 instances. The results suggest that it may be difficult for the participants to notice errors coded as Misformation, such as *shaked* and *fighted*, by themselves and that recasts may facilitate such noticing of overgeneralized forms.

It is important to ascertain whether only a few individuals contributed to the high rate of noticing of errors through recasts. A post hoc analysis of the data showed that such individual biases were not observed. Of the ten participants in the recast group, three (Rinko, Momo, and Yusuke) did not report noticing of errors through recasts; four (Yuko, Toshiya, Akira, and Koichi) reported their noticing of errors through recasts for misformation (1 instance) and incorrect (1 instance); two (Tomoko and Yuki) reported their noticing of errors through recasts for misformation (1 instance); and one (Taro) reported his noticing of errors through recasts for incorrect (1 instance). Although the number of cases for each individual was quite small, the positive effect of recasts on noticing was observed for most participants (7 out of 10).

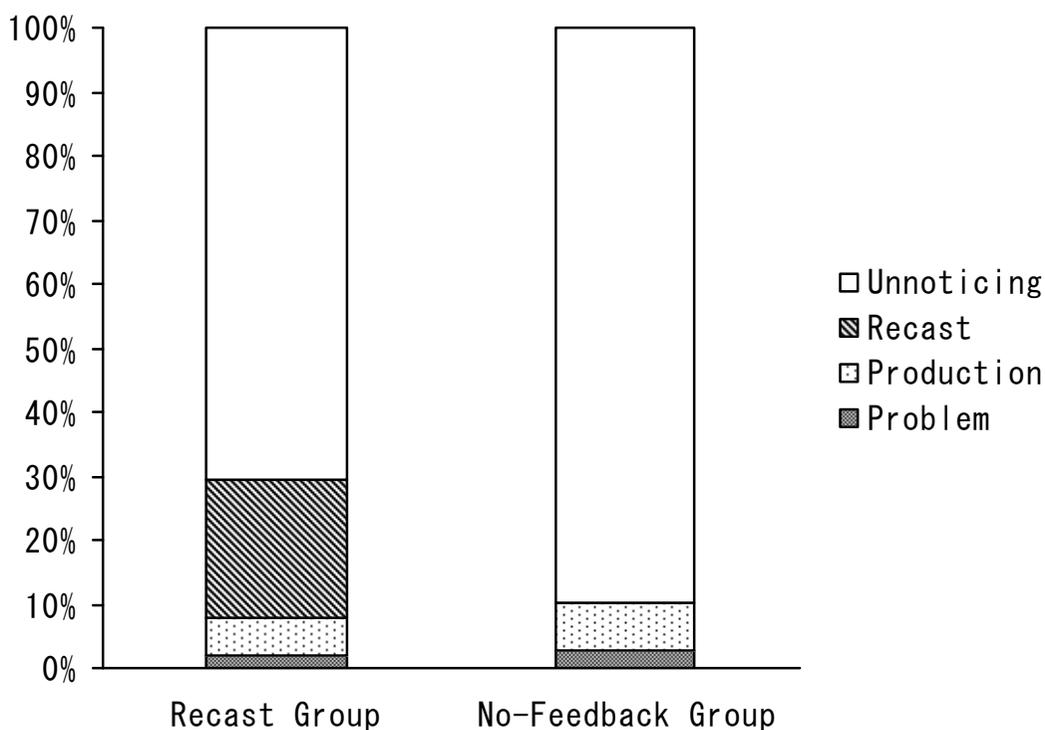


Figure 1. Distributions of noticing types by groups.

Discussion

Analyses of verbal reports showed that the provision of recasts was shown to be effective in promoting L2 learners’ noticing, in particular, of overgeneralization errors. Also, they showed that the activity of production led to a certain degree of noticing of errors or linguistic problems irrespective of the provision of recasts. This study attempted to isolate the effects of recasts from those of the activity of production on L2 learners’ noticing. By doing so, this study provides separate empirical support for the claims of the interaction hypothesis (Long, 1996, 2007) and the output hypothesis (Swain, 1985, 1993, 1995) with regard to noticing.

In his interaction hypothesis, Long (1996, 2007) claims that negative feedback facilitates L2 acquisition by drawing L2 learners’ attention to nonsalient and problematic forms. He pointed out that “Heightened attention makes detection both of new forms and of mismatches between input and output more likely, and such mismatches may also provide at least some of

the information a learner needs about what is *not* permissible in a language” (p. 453). The findings of this study lend support to his claim. First, the results showed that the recast group noticed significantly more errors than the no-feedback group and suggested that recasts are effective in promoting L2 learners’ noticing. Furthermore, as shown in the Results section, the recast group noticed overgeneralized forms (i.e., Misformation) as errors far more frequently than the no-feedback group (8 out of the 15 instances for the recast group; 1 out of the 16 instances for the no-feedback group). As Long argued, the results suggest that recasts are effective in drawing L2 learners’ attention to otherwise nonsalient forms and in informing L2 learners that overgeneralized forms such as *shaked* and *fighted* are ungrammatical in the target language.

The findings of this study also support Swain’s (1985, 1993, 1995) output hypothesis, which claims that by producing the target language, L2 learners may notice what they cannot express in the target language. According to Swain, this type of noticing, that is to say, noticing a hole, may make L2 learners more attentive to further input, leading to more opportunities for L2 acquisition. The results showed that both the recast and the no-feedback groups noticed their linguistic problems or errors through the activity of production (7.9% for the recast group; 10.4% for the no-feedback group). However, it is important to note that the low rates suggest that the noticing function of output may be limited, in particular, for overgeneralized morphological errors like the irregular past tense forms. In other words, L2 learners who produce overgeneralized past forms of irregular verbs by adding the *-ed* suffix to a verb stem may have difficulty in noticing such errors by themselves (i.e., without any feedback) at the moment of production. Because the rule of the addition of the *-ed* suffix is grammatical for some set of verbs and productive in English, L2 learners may need negative evidence that applying that rule to some set of verbs (i.e., irregular verbs) is not grammatical. The activity of production itself may not provide such negative evidence.

Finally, coding systems for the verbal protocols are discussed. Previous studies used systems developed by the researchers for the purposes of their study. For example, Mackey, Gass, and McDonough (2000) and Gass and Lewis (2007) attempted to investigate whether L2 learners will notice the linguistic targets (lexical, morpho-syntactic, or phonological aspects) of

negative feedback correctly. Egi (2004) aimed at examining the reactivity and veridicality of the stimulated recall method and therefore included such categories as *no memory* and *no thoughts* in her system. Importantly, Egi (2007a, 2007b) categorized the verbal protocols on the basis of the types of linguistic data (negative evidence, positive evidence, and negative and positive evidence). Thus, her categorization system made it possible to examine what type of data L2 learners may receive from recasts. Finally, this study divided the verbal protocols according to the noticing types proposed by Swain (1998) and Doughty and Williams (1998); in particular, this study attempted to isolate the effects of production alone from the effects of recasts. In the categorization system used in this study, the category *noticing a gap through a recast* covers both (a) negative evidence and (b) negative and positive evidence in Egi's categorization system. Therefore, the categorization system of this study examined whether L2 learners received negative evidence, that is, whether they noticed their errors or not; nevertheless, it did not make clear whether they received positive evidence, that is, whether they noticed the target forms. The categories for coding must be developed with great care in accordance with the purpose of the study. One implication for future study is that it might be advantageous to combine the coding systems of Egi (2007a, 2007b) and the present study. More specifically, verbal protocols can be categorized on the basis of noticing types as in the current study so that the effects of output alone will be eliminated; and they can be further categorized according to whether L2 learners receive positive evidence, negative evidence, or both. A post hoc analysis following this procedure revealed that, out of the 11 verbal reports coded as *noticing a gap through a recast*, five were coded as negative evidence; six were coded as negative and positive evidence. Although the numbers were small, this post hoc analysis suggests that recasts provide both negative and positive evidence for L2 learners.

Conclusion

The research question posited for this study was: Do recasts alone provide negative evidence about the grammaticality of L2? By analyzing the verbal protocols of L2 learners' noticing elicited through a stimulated recall technique, this study provided a confirmative answer to this question: In other words, the results suggest that recasts do facilitate L2 learners' noticing

of errors even when they did not notice their linguistic problems or errors at the moment of production. Some researchers have argued that recasts are ambiguous as negative feedback (Lyster, 1998a) and that it is difficult for L2 learners to notice the corrective intentions of recasts accurately (Gass & Lewis, 2007; Mackey, Gass, & McDonough, 2000; Roberts, 1995). On the other hand, by analyzing verbal protocols of L2 learners' noticing in detail, Egi (2007a) found that L2 learners do receive negative evidence from recasts. Along this line, this study provided additional evidence for the function of recasts to provide negative evidence.

Based on the findings of this study, two pedagogical implications can be made. First, although production itself may lead to some degree of noticing, providing L2 learners with opportunities for interactions in which they can receive negative feedback such as recasts from a more competent interlocutor or teacher may be more effective in order to promote noticing of errors. In other words, when learners produce an overgeneralized and interlanguage form such as *shaked* or *drinked*, the findings of this study suggest that they may have difficulty in noticing their errors by themselves (i.e., through the activity of production). That is the point in which recasts may be effective. Thus, as a second pedagogical implication, it is suggested that teachers can provide recasts, depending on error types. Recasts may be effective for overgeneralized past forms of irregular verbs.

In interpreting and generalizing the findings, however, it is important to take into account a few limitations of the study. First, this study was experimental in nature. The participants in the recast group received the feedback intensively; the context was quite different from natural conversational interactions. The findings that recasts promote L2 learners' noticing need to be corroborated by more studies with different contexts including natural settings. Second, verbal reports about noticing cannot avoid suffering from memory loss (Egi, 2004; Gass & Mackey, 2000; Robinson, 1995). For example, Robinson (1995) stated that "the experience of noticing may be fleeting and thus difficult to recall" (p. 299). To overcome this problem, this study used audiotaped recordings and task materials as prompts to stimulate the recall of noticing. However, future studies may require different methods to measure noticing, such as think-aloud protocols and immediate reports, within their research design.

Notes

1. The phase of Task 3 was intended to examine the learning effects, that is, the improvement from Task 1 (pretest) to Task 3 (posttest). However, the results of Task 3 are not reported in the current study because they are beyond the scope of this study that focuses on the analysis of verbal reports on noticing.
2. The recast group was provided with either recasts or non-corrective repetition. Because all the erroneous utterances were followed by recasts, the recast group did not receive any repetition of their erroneous utterances from the researcher.
3. Two issues, reactivity and veridicality, were taken into consideration (Egi, 2004; Gass & Mackey, 2000). Reactivity concerns the influence of stimulated recall on the target performance. Adams (2003) showed that stimulated recall can provide additional learning opportunities if it is positioned before the posttest (see also Nabei & Swain, 2002). In this study, stimulated recall was carried out after all the tasks were completed in order to avoid influencing the performance on Task 3. Veridicality concerns the accuracy of verbal reports in stimulated recall. Egi (2004) not only pointed out the memory decay problem associated with stimulated recall, but also the misinformation effect (p. 256). In the current study, the misinformation effect might be observed because Task 3, positioned between the treatment task (Task 2) and the stimulated recall, might have influenced the verbal reports; in other words, the L2 learners might have confused their thought processes in the treatment task with those in Task 3. To reduce the misinformation effect, two different tasks with different pictures were prepared for Task 2 and Task 3 even though both tasks were intended to elicit the same verbs as mentioned in the Tasks section.
4. Because the sample size was small and the data were not normally distributed, non-parametric procedures, Mann-Whitney tests, were selected (Hatch & Lazaraton, 1991, p. 270). Accordingly, instead of means and standard deviations, medians (*Mdns*) and interquartile ranges (*IQRs*) were used as a measure of central tendency and dispersion respectively. The interquartile ranges indicate the ranges of the middle 50% of the data.
5. Holm's sequential Bonferroni method is one of the methods to avoid making a Type I error, referred to as "the probability of rejecting a null hypothesis when it is true" (Green & Salkind, 2005, p. 417), in multiple comparisons. This method was used for this study according to Green and Salkind (2005, pp. 418-419). The family wise alpha for this study was set at the .05 level. Because three comparisons between the two groups were tested in this study, the three-step procedure was as follows: First, the smallest *p* value among the three comparisons was evaluated against the .016 level, obtained by dividing .05 by the number of comparisons; second, the second smallest *p* value was evaluated against the .025 level, obtained by dividing .05 by 2 (the number of comparisons – 1); and finally, the largest *p* value was evaluated against the .05 level. If the null hypothesis is not rejected at a particular step in the procedure, the analysis is stopped at that point. In that case, the remaining comparisons are assessed as nonsignificant.

References

- Adams, R. (2003). L2 output, reformulation and noticing: Implications for IL development. *Language Teaching Research*, 7, 347-376.
- Ayoun, D. (2001). The role of negative and positive feedback in the second language acquisition of the passé composé and imparfait. *The Modern*

- Language Journal*, 85, 226-243.
- Carroll, S., & Swain, M. (1993). Explicit and implicit negative feedback: An empirical study of the learning of linguistic generalizations. *Studies in Second Language Acquisition*, 15, 357-386.
- 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 D.C.: 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.
- Dulay, H., Burt, M., & Krashen, S. (1982). *Language two*. NY: Oxford University Press.
- Egi, T. (2004). Verbal reports, noticing, and SLA research. *Language Awareness*, 13, 243-264.
- Egi, T. (2007a). Interpreting recasts as linguistic evidence: The roles of linguistic target, length, and degree of change. *Studies in Second Language Acquisition*, 29, 511-537.
- Egi, T. (2007b). Recasts, learners' interpretations, and L2 development. In A. Mackey (Ed.), *Conversational interaction in second language acquisition* (pp. 249-267). Oxford: Oxford University Press.
- Ellis, R. (1987). Interlanguage variability in narrative discourse: Style shifting in the use of the past tense. *Studies in Second Language Acquisition*, 9, 1-20.
- Ellis, R. (2007). The differential effects of corrective feedback on two grammatical structures. In A. Mackey (Ed.), *Conversational interaction in second language acquisition* (pp. 339-360). Oxford University Press.
- Ellis, R., Loewen, S., & Erlam, R. (2006). Implicit and explicit corrective feedback and the acquisition of L2 grammar. *Studies in Second Language Acquisition*, 28, 339-368.
- Gass, S. M., & Lewis, K. (2007). Perceptions about interactional feedback:

- Differences between heritage language learners and non-heritage language learners. In Mackey A. (Ed.), *Conversational interaction in second language acquisition* (pp. 79-99). Oxford University Press.
- Gass, S. M., & Mackey, A. (2000). *Stimulated recall methodology in second language research*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Gass, S. M., & Mackey, A. (2007). Input, interaction, and output in second language acquisition. In B. VanPatten & J. Williams (Eds.), *Theories in second language acquisition* (pp. 175-199). Mahwah, NJ: Lawrence Erlbaum.
- Green, S. B., & Salkind, N. J. (2005). *Using SPSS for Windows and Macintosh: Analyzing and understanding data* (4th ed.). Upper Saddle River, NJ: Pearson Education Inc.
- 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*. New York: Newbury House.
- Ishida, M. (2004). Effects of recasts on the acquisition of the aspectual form –te i- (ru) by learners of Japanese as a foreign language. *Language Learning*, 54, 311-394.
- Iwashita, N. (2003). Negative feedback and positive evidence in task-based interaction: Differential effects on L2 development. *Studies in Second Language Acquisition*, 25, 1-36.
- Leeman, J. (2003). Recasts and second language development: Beyond negative evidence. *Studies in Second Language Acquisition*, 25, 37-63.
- Lightbown, P. M., & Spada, N. (1993). *How languages are learned*. Oxford: Oxford University Press.
- Loewen, S., & Nabei, T. (2007). Measuring the effects of oral corrective feedback on L2 knowledge. In A. Mackey (Ed.), *Conversational interaction in second language acquisition* (pp. 361-377). Oxford University Press.
- Long, M. H. (1996). The role of the linguistic environment in second language acquisition. In W. C. Ritchie, & T. K. Bhatia (Eds.), *Handbook of second language acquisition* (pp. 413-468). San Diego, CA: Academic Press.
- Long, M. H. (2007). *Problems in SLA*. Mahwah, NJ: Lawrence Erlbaum.
- Long, M. H., Inagaki, S., & Ortega, L. (1998). The role of implicit negative

- feedback in SLA: Models and recasts in Japanese and Spanish. *The Modern Language Journal*, 82, 357-371.
- 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 classrooms. *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. (1995). *Stepping up the pace: Input, interaction and interlanguage development: An empirical study of questions in ESL*. Unpublished doctoral dissertation, University of Sydney, Australia.
- 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. (Ed.) (2007). *Conversational interaction in second language acquisition*. Oxford University Press.
- Mackey, A., Gass, S., & McDonough, K. (2000). How do learners perceive interactional feedback? *Studies in Second Language Acquisition*, 22, 471-497.
- Mackey, A., & Goo, J. (2007). Interaction research in SLA: A meta-analysis and research synthesis. In A. Mackey (Ed.), *Conversational interaction in second language acquisition* (pp. 407-462). Oxford University Press.
- Mackey, A., & Philp, J. (1998). Conversational interaction and second language development: Recasts, responses, and red herrings? *The Modern Language Journal*, 82, 338-356.
- Nabei, T., & Swain, M. (2002). Learner awareness of recasts in classroom interaction: A case study of an adult EFL student's second language learning. *Language Awareness*, 11, 43-63.
- Nassaji, H. (2007). Elicitation and reformulation and their relationship with learner repair in dyadic interaction. *Language Learning*, 57, 511-548.
- Nicholas, H., Lightbown, P. M., & Spada, N. (2001). Recasts as feedback to language learners. *Language Learning*, 51, 719-758.
- Nobuyoshi, J., & Ellis, R. (1993). Focused communication tasks and second language acquisition. *ELT Journal*, 47, 203-210.

- Oliver, R. (1995). Negative feedback in child NS/NNS conversation. *Studies in Second Language Acquisition*, 17, 459-481.
- Panova, I., & Lyster, R. (2002). Patterns of corrective feedback and uptake in an adult ESL classroom. *TESOL Quarterly*, 36, 573-595.
- Pica, T. (1983). Methods of morpheme quantification: Their effect on the interpretation of second language data. *Studies in Second Language Acquisition*, 6, 69-78.
- 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, HI: University of Hawaii at Manoa.
- Robinson, P. (1995). Attention, memory, and the “noticing” hypothesis. *Language Learning*, 45, 283-331.
- Sakai, H. (2004). Roles of output and feedback for L2 learners’ noticing. *JALT Journal*, 26, 25-54.
- Saxton, M. (1997). The contrast theory of negative input. *Journal of Child Language*, 24, 139-161.
- Saxton, M. (2000). Negative evidence and negative feedback: Immediate effects on the grammaticality of child speech. *First Language*, 20, 221-252.
- Schmidt, R. (1990). The role of consciousness in second language learning. *Applied Linguistics*, 11, 129-158.
- Simard, D., & Wong, W. (2001). Alertness, orientation, and detection: The conceptualization of attentional functions in SLA. *Studies in Second Language Acquisition*, 23, 103-124.
- Swain, M. (1985). Communicative competence: Some roles of comprehensible input and comprehensible output in its development. In S. M. 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. *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 University Press.
- Swain, M. (1998). Focus on form through conscious reflection. In C. Doughty & J. Williams (Eds.), *Focus on form in classroom second*

language acquisition (pp. 64-81). New York: Cambridge University Press.

Takashima, H. (1995). *A study of focused feedback, or output enhancement, in promoting accuracy in communicative activities*. Unpublished doctoral dissertation presented to Temple University Japan.

Tomlin, R., & Villa, H. (1994). Attention in cognitive science and second language acquisition. *Studies in Second Language Acquisition*, 16, 183-203.

Van den Branden, K. (1997). Effects of negotiation on language learners' output. *Language Learning*, 47, 589-636.

Appendix Sample Pictures for Task 2 (Ken's Story)

