Effective English Teaching in Japanese Secondary Schools:

Examining the Effect of Adopting Output Tasks

Mutsuko NAGASAKI

中学，高等学校における効果的な英語指導法

― アウトプット，協同ダイアログを用いた指導法の効果の検証 ―

長　崎　睦　子

本研究は第二言語学習におけるアウトプットの役割について考察したものである。Swain（1985, 1995）が提唱したアウトプットの機能のひとつである「気づき機能」に焦点をあて，個人でのアウトプットならびに協同で作り上げるアウトプットが学習者の英語仮定法過去完了への気づきやその学習を促進するかどうかを調査した。本研究は，日本人大学一年生79名を被験者とし，事前・事後テストを用いた教室第二言語習得研究である。結果を基に，中学，高等学校におけるアウトプット・タスクを用いた効果的な英語指導法について考察している。

Key words: アウトプット仮説，気づき機能，テキスト再生タスク，協同的対話

1. Introduction

1.1 Purpose of the Study

The present study aimed to investigate the role of production, namely, *output*, in second language acquisition (SLA). Particularly, it focused on examining the effect of two types of output, *individual output* and *collaborative output*. Individual output is a type of output produced by second language (L2) learners individually, while collaborative output is produced collaboratively by L2 learners in pairs or in group. This collaborative output is also called *collaborative work* or *collaborative dialogue* (Swain, 2000). It is very important to note at the beginning that the present study utilized the word collaborative output to indicate both collaborative work and collaborative dialogue. This was done to clearly identify any output which learners produced collaboratively in pairs, as opposed to output produced by an individual learner. Finally, based on the results, the study attempted to provide

執筆者の所属：基礎総合教育　論文受理日2008年10月10日

some implications for L2 classrooms at secondary schools in Japan.

1.2 Swain's Output Hypothesis

Although Krashen's Input Hypothesis[1] (1985) suggests that it is comprehensible input that is necessary and sufficient for second language acquisition, Swain (1985) claims that not only comprehensible input but also comprehensible output are necessary to improve both fluency and accuracy. As Krashen contends, the importance of input in second language acquisition cannot be denied. However, the question is whether it is sufficient for acquisition to take place.

Swain states that output, especially being *pushed* to produce, is crucial for second language acquisition because it helps learners to extend their linguistic repertoire as they attempt to produce output in a more precise and appropriate manner. Swain (1995) further discusses three functions of output: (1) *the hypothesis-testing function*[2], (2) *the noticing function,* and (3) *the metalinguistic function*[3], which are related to facilitating accuracy. The present study focused on the noticing function of output in particular and researched whether individual output and collaborative output promote *noticing* and acquisition of an English linguistic form, in this case, the past hypothetical conditional.

The noticing function is when output works to prompt learners to consciously recognize some of their linguistic problems. Through output, learners become aware of what they know or know only partially, which may lead them to obtain additional information in order to know about their L2. Swain and Lapkin (1995) indicate that when learners notice their linguistic problems through output, they make an analysis and work on modifying their output. Specifically, they may return to input with more attention to search for relevant input, or they may work out a solution which leads to new reprocessed output. They state, "what goes on between the first output and the second, we are suggesting, is part of the process of second language learning" (p.386). The present study noted these words and attempted to research whether learners notice their linguistic problems through output and search for the information they need to solve the problems if the same input is provided again.

Regarding the importance of noticing in second language acquisition, Schmidt (1990) argues that to acquire any particular aspect of L2, one first needs to notice it. He operationally defines noticing as availability for verbal report and states that noticing is a necessary and sufficient condition for converting input to intake, and subliminal language learning is impossible (p.129). For Schmidt, awareness at the level of noticing is necessary for learning to take place. The present study investigated whether this noticing occurs through output, and also whether it leads to second language learning.

1.3 Previous Studies

Based on Swain's Output Hypothesis, Izumi and his colleagues' studies (Izumi, Bigelow, Fujiwara, & Fearnow 1999; Izumi & Bigelow 2000) attempted to address whether output encourages ESL (English as a Second Language) learners to find their linguistic problems and to seek out the necessary information to solve them if appropriate input is subsequently given to them. Their argument was that if output promotes noticing, learners should process subsequently provided input with more focused attention. Therefore, the distinctive feature in the treatment of their studies was to provide input twice, prior to and after output, which enabled the researchers to investigate whether output promotes noticing. In other words, it enabled them to seek out evidence for the theory that learners' attention to form increases from the first input to the second input.

The results of their studies did not support their hypothesis that greater noticing of the target form, the past hypothetical conditional, would be seen in the experimental group, which had opportunities to produce output, compared to the control group, which had no output opportunities. Instead, they found that in both groups there was a significant increase in the participants' noticing of the target form from the first input to the second input. Regarding the other hypothesis, which stated that the experimental group's accuracy of the use of the target form would improve more than the control group's, Izumi et al. (1999) partially supported it. However, Izumi et al. (2000) did not support it because both groups improved their use of the target form significantly after completing the whole treatment (engaging in two types of output tasks), and there were no significant differences between the two groups on the posttests.

Interestingly, the results from the two studies indicated that only after the whole treatment was finished did accurate use of the past hypothetical conditional by the output participants increase significantly from the first output to the second output. Thus they pointed out that repeating the task seemed to have an effect on the participants' improvement. They suggest that providing learners with extended opportunities to produce output and receive relevant input is crucial in promoting their learning of grammatical structures.

Izumi (2002) further investigated the effectiveness of output on the noticing and acquisition of relativization by adult ESL learners. In this study, he found that learners who engaged in both input and output activities outperformed those who received the same input for the sole purpose of comprehension in learning gains. Izumi claims that the noticing function underscores the interconnectedness of input and output processes in second language acquisition, and concludes that the study shows that output can induce learners to process input effectively for greater *Interlangauge* (IL) development.

In sum, the results of the studies above are mixed and more research regarding this issue is definitely necessary. More studies employing different output conditions, target forms, participants, or settings collecting data are needed.

1.4 Collaborative Output and Second Language Learning

Izumi and Izumi (2004) indicate that meaningful output, which attracts learners' attention to crucial form-meaning relationships, has to be encouraged for output to have a significant impact on L2 development so that genuine syntactic processing is engaged. They suggest using a *dictogloss* task as one of the techniques to have learners engage in meaningful output because the beneficial effects of dictogloss have been demonstrated by some second language acquisition studies (e.g., Swain, 1998; Swain and Lapkin, 2001). Dictogloss (Wanjnryb, 1990) is a task in which students listen to a short text read by the teacher and while it is being read, they write down any words and phrases to reconstruct the text later in small groups. This suggestion is compatible with Swain's (1998) claim that IL restructuring through output successfully occurs in collaborative work such as in a dictogloss task.

Swain (1998) suggests that the circumstance of working on producing output collaboratively is where collaborative consciousness is raised and learners pool analytic capacities and co-construct linguistic knowledge. Furthermore, referring to an example of the interaction between two first language (L1) English students in the study of Kowal and Swain (1994), Skehan (1998) indicates that the scaffold provided by collaboration enables the limitation of (individual) short-term memory to be overcome, and co-construction allows the transformation of material to be achieved and potentially integrated into long-term memory. Thus, collaborative work to produce output, such as in a dictogloss task, seems to be important for output to successfully promote IL restructuring. In other words, collaborative output where consciousness raising and mutual help are likely to occur by collaboration may enable the limitations of individual short-term memory to be overcome and possibly lead to long-term memory.

If collaborative work for output creates an environment where learners discuss their linguistic problems and work together to solve the problems, presumably, these cognitive activities will deepen their attention to the formal features of their problematic forms more than when working individually. The author argues that working together may draw learners' attention to the forms which he/she cannot notice alone or to a depth which he/she cannot attain alone.

The present study was inspired by Swain's Output Hypothesis and Izumi and his colleagues' series of studies, and determined to further research the noticing function of output. The research design and methodology basically adopted that of Izumi et al. (1999) and Izumi and Bigelow (2000). There are

mainly three differences of the present study compared to Izumi et al.'s studies reviewed above: (1) the Collaborative Output Group was set up as an experimental group; (2) in order to investigate the role of collaborative output in L2 learning, the study adopted the methodology and analysis styles used by studies with a sociocultural framework as well as studies with a psycholinguistic framework[4]; and (3) the participants were Japanese students studying English at a Japanese university; thus, the data was collected in the English as a Foreign Language (EFL) setting.

2. Method

2.1 Research Questions

The present study attempted to investigate the effectiveness of output on second language acquisition, and two major research questions were raised.

1) Does output promote the noticing of the past hypothetical conditional and the acquisition of the form more than input only (non-output)?
2) Does collaborative work for output promote the noticing of the past hypothetical conditional and the acquisition of the form more than individual work for output?

Based on the review of the SLA theories and previous studies, the following hypotheses were formed.

Hypothesis 1: The participants who are required to produce output (the Output Group = O) will notice the past hypothetical conditional in input more and attain higher scores on the posttests, in terms of the form, than the participants who have no opportunity to produce output (the Non-Output Group = NO).

Hypothesis 2: The participants who are required to work for output collaboratively (the Collaborative Output Group = CO) will notice the past hypothetical conditional in input more and attain higher scores on the posttests, in terms of the form, than the participants who are required to work for output individually (O).

2.2 Target Form

The grammatical target form for the present study was the past hypothetical conditional in English, such as *If I had had the money, I would have taken the vacation* (from Celce-Murcia and Larsen-Freeman, 1999). Celce-Murcia and Larsen-Freeman pointed out that conditional sentences are very difficult for both English as a Second Language (ESL) and English as a Foreign Language (EFL) students to learn due to the following general reasons: (1) a conditional sentence consist of two

clauses, a subordinate clause and a main clause, which makes it more syntactically complex than many other structures; (2) the meanings that the various types of conditional clauses convey are subtle, so that they are difficult to understand even for native speakers of English; and (3) in order to cope with the past hypothetical conditional fully, the students need to clearly understand English grammar structures such as the tense-aspect system, modal auxiliaries, and negation in advance.

One of the purposes of the present study was to provide some useful implications for L2 learning at secondary schools in Japan. Therefore, when and how the past hypothetical conditional is introduced in Japanese secondary schools is briefly discussed. Some of the English textbooks authorized by *the Ministry of Education, Culture, Sports, Science and Technology (MEXT)* of Japan are examined.

According to *the Course of the Study* for Foreign Languages (English)[5] by MEXT, the past hypothetical conditional is first introduced in upper secondary school. Careful analysis of the authorized textbooks revealed that all the English subjects (Oral Communication I, II, English, I, II[6], Reading, and Writing) except Oral Communication deal with the past hypothetical conditional. According to the guidebook for *the Foreign Language (English) Course of the Study* for upper secondary school (1999), of six subjects, only one subject, either Oral Communication I or English I, is required. In order to take Communication II, Communication I is required, while English I is the pre-requisite for English II. In addition, to take Reading or Writing, either Communication I or English I needs to be completed. Therefore, which courses, as well as how many courses, the upper secondary school students actually take depends on the school at which they study. In other words, it is possible that some students take most or all of the courses, while others take only one, either Oral Communication I or English I.

In the case that the students take only Oral Communication I, one cannot deny the possibility that they won't receive any input or produce output regarding the past hypothetical conditional, although practically, taking only Oral Communication I might be rare. However, at least, it can be concluded that the past hypothetical conditional is dealt with as much as the other grammatical forms in reading, but it is not the case in oral communication. Furthermore, since writing is an elective course, some might not practice producing the target form through writing. Thus, it is probable that the participants for the present study, who had all studied at a secondary school in Japan, had received input of the past hypothetical conditional through reading, but they did not have any opportunities to listen to it or orally produce it in the oral communication context. Some may have hardly written it either.

2.3 Participants

The participants for this study were 79 first-year students from four English language classes at a

private university in Japan. All of the classes were from one of the required English courses for freshmen called “Cinema English I” which used British or American movies as materials to learn English, as well as to study about their cultures. A total of four groups were established, namely, the Non-Output Group (NO, n = 18), the Output Group (O, n = 20), the Collaborative Output Group (CO, n = 18), and the Control Group (C, n = 23). These four groups were established by randomly assigning four intact classes to either NO, O, CO, or C. The number of participants shown for each group is the number of students who completed all the pretests, posttests and the two treatment phases. The pretest administered before any treatment was given to the three experimental groups shows that there was no significant difference between NO and O or between O and CO in terms of the attainment of the past hypothetical conditional (Tables 4 and 5 in Section 3). The background questionnaire revealed the participants’ recognition of the past hypothetical conditional (Table 1).

Table 1 *Participants’ Recognition of the Target Form*

	NO (N = 18)				O (N = 20)				CO (N = 18)				C (N = 23)			
	×	△	○	◎	×	△	○	◎	×	△	○	◎	×	△	○	◎
Future	2	7	8	1	1	5	14	0	1	6	12	1	2	7	12	2
Present	2	7	8	1	1	7	12	0	1	7	11	1	2	7	13	1
Past	2	7	8	1	1	8	11	0	1	7	11	1	2	7	13	1

Note. Future = Future conditional, Present = Present conditional, Past = Past conditional
× = had never heard of and did not know the form, △ = had heard of but did not know anything about the form, ○ = had heard of and knew the form to some extent, ◎ = had heard of and knew the form well

2.4 Research Design

The present study was a classroom quasi-experimental research with a pretest-posttest design and involved three experimental groups and a control group. Three experimental groups, NO, O, and CO were established. The independent variable between NO and O was whether the participants were required to produce output, and this variable was set in order to investigate the effectiveness of output on the noticing and the learning of the target form. The independent variable between O and CO was whether they were required to produce output individually or collaboratively (Table 2). Therefore, it was possible to examine which type of output promotes noticing and learning more. The Control Group did not experience any treatment but only took the pretest and the posttests.

Table 2 *The Difference Between NO and O, O and CO*

NO	O
± Output	

O	CO
± Collaboration	

2.5 Schedule of Data Collection and Treatment

The treatment and the tests were conducted over six weeks according to the schedule shown in Figure 1. The treatment consisted of two target sessions. In each target session, one task was administered; thus the participants in the three experimental groups engaged in two tasks in total. Two sessions were held in order to provide them with the opportunity to repeat the same type of task and investigate whether task repetition is effective in promoting L2 learning. As Figure 2 shows, a target session consisted of (1) a combination of two reading phases and two writing phases for O, (2) a combination of two reading phases and two collaborative writing phases for CO, and (3) a combination of two reading phases and two phases of answering questions for NO.

Week 1: Background Information Questionnaire Pretest
Week 2: Target Session 1, "Forrest Gump's Story 1" Uptake-claim survey
Week 3: Posttest 1
Week 4: Target Session 2, "Forrest Gump's Story 2" Uptake-claim survey
Week 5: Posttest 2
Week 6: Individual Performance of Task (Forrest Gump's Story 1, 2) by CO[7]

Figure 1 Schedule of Data Collection

Two reading passages, Text 1 and Text 2[8], were prepared for the target sessions. Both texts were originally made for the present study based on the story from the American movie "Forrest Gump" (1994), which depicts the life of a man who was born with physical and mental handicaps. Each text contained four past hypothetical conditional sentences, which was about 44 % of the total text. Along with the text, pictures depicting the story were also provided. The three experimental groups were asked to read the text to comprehend it. While reading it, the O and CO participants were required to write down any words or information needed to reconstruct the text, whereas the NO participants were told to write down words and information to answer the comprehension and related questions. The students were given four minutes to complete this task.

After the reading passage was collected, paper was provided and the O participants worked individually to reconstruct the text as accurately as possible in terms of content and grammar (15 minutes). They were allowed to look at the notes they had taken during the reading phase while reconstructing the text. The CO participants were assigned to pairs by the teacher[9]. They were required to work together to reconstruct the story (15 minutes). A piece of paper was given to each pair, and they were encouraged to write a text which was as similar to the original as possible in terms of grammar and content. They were told to reproduce the text by referring to each other's notes and

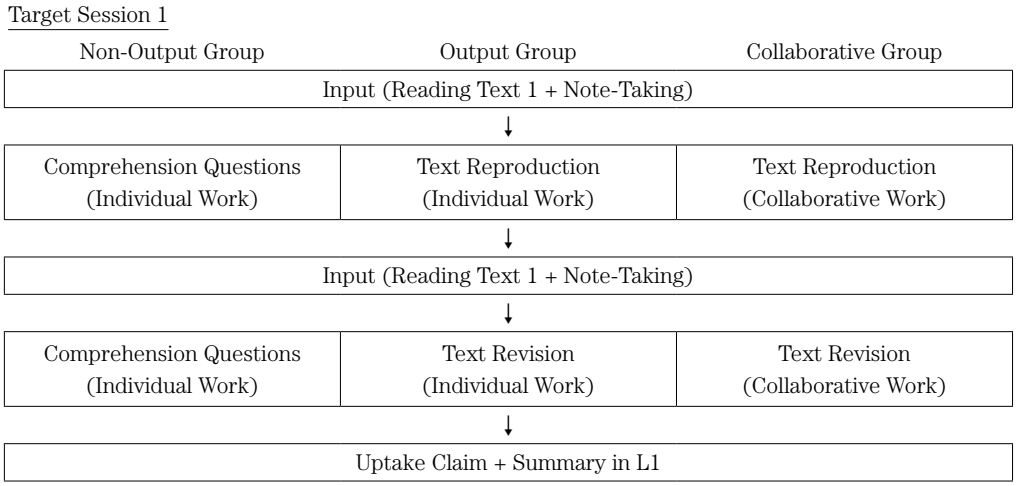

Figure 2 Treatment Sequence for Each Group
Note. The same sequence was followed in Target Session 2, in which Text 2 was used.

sharing resources. The conversations of each pair were audiotaped for the analysis of collaborative dialogue. The type of task used for O and CO is considered a text-reconstruction task, used by Izumi, Bigelow, Fujiwara, and Fearnow (1999) and Izumi and Bigelow (2000)[10]. The NO participants answered a series of yes/no comprehension questions as well as questions related to the topic of the text. These questions were written in Japanese, their first language, to avoid providing them with further input of the target form.

After collecting the reconstruction sheets as well as the comprehension and related question sheets, the students were asked to read the same text again for 3 minutes and take notes to reproduce (revise) it (O and CO)[11] for 10 minutes or to answer the comprehension and related questions (NO) for 4 minutes. Finally, the participants from all of the groups were asked to work individually and write a summary of the story in their L1, Japanese. Immediately after writing the summary, an uptake-claim survey[12], as used by Slimani (1992), was administered to the participants for further analysis of the students' noticing of the target form. The second target session was held two weeks later and used the same procedures described above for each group respectively. Text 2 was used as the reading text for the second session.

2.6 Noticing Measures

There were two dependent variables in the study. One variable was the noticing measures. These were adopted to examine the degree of noticing by NO, O, and CO. There were three types of noticing

measures: (1) note scores, which were calculated based on the note-taking done by all the participants while reading the text, (2) immediate uptake scores; namely, text-reconstruction scores, which were calculated based on the production by the O participants and the CO participants while reproducing the text, and (3) an uptake-claim survey, which was answered by all the participants after each target session and was used to examine whether and to what degree the participants paid attention to the target grammatical form during each target session.

2.7 Acquisition Measures

The other variable was the acquisition measures. Two types of tests, a multiple-choice recognition test and a picture-cued production test were used to investigate the degree of learning achieved by the three experimental groups respectively. These two tests were made for the present study based on the tests by Izumi & Bigelow (2000).

There were twenty questions in the multiple-choice recognition test in total. Ten included the target form, the past hypothetical conditional, while the other ten were distracters. Of the ten questions containing the target form, five began with an *if*-clause, and five began with a main clause. As a past hypothetical conditional sentence consists of two clauses, an *if*-clause and a main clause, each target question included two blanks, one for the first clause and the other for the second clause. Therefore, there were ten blanks for the *if*-clauses and ten blanks for the main-clauses. Since there were a total of twenty blanks, the maximum score each participant could obtain was twenty.

The picture-cued production test was also made following the test format utilized by Izumi and Bigelow. It consisted of three sections, each of which contained four questions. Two of the three sections dealt with the target form; hence, there were eight questions where the participants were supposed to make a past hypothetical conditional sentence. All the items contained pictures which represented the meaning, and the students were required to make a sentence which conveyed that meaning. Each target item contained two pictures, one of which represented the *if*-clause, and the other, the main clause. There was an arrow between the two pictures to make it clear that one picture indicated the cause, and the other indicated the result[13].

For the picture-cued production test, the study used TLU analysis and Interlanguage (IL) analysis[14] adopted from Izumi and Bigelow (2000). In the TLU analysis, the scores for the *if*-clauses and those for the main clauses were calculated separately. There were eight questions for each type of clause, and one point was given for each targetlike use of the target form. Thus, the maximum score for the *if*-clauses and the main clauses were eight respectively, and the total maximum score was 16.

3. Results and Discussion

3.1 Noticing – Note-Taking Scores

In order to examine the noticing issues of the first and second research questions, the results of (1) note-taking scores, and (2) text-reconstruction scores are presented. To analyze the participants' noticing of the past hypothetical conditional, the note-taking scores from NO, O and CO were calculated based on the notes they took while reading the text during the target sessions. In each target session, the participants were exposed to the input twice, before and after the first output trial. While receiving input through reading, they were asked to write down any words or phrases which they thought were necessary either to answer the comprehension questions (NO) or to reconstruct the text (O and CO). Each participant's note-taking scores were calculated in percentages. For example, the mean of the O participants' note-taking scores in the first input in Target Session 1 was 11% (Table 3). This means that on average, 11% of the words they wrote down during the first note-taking in the first target session were conditional-related items such as modals (would), aspect auxiliaries (have, had), copula in the past participle form (been), complementizer (if), and the past participle ending (-ed and -en).

Table 3 shows the statistical analysis of the note-taking scores for NO, O, and CO, and presents their scores in percentages. The results from Target Session 1 show that there was a significant difference from the first input to the second input in O, $t(19) = 3.15$, $p<.01$, and in CO, $t(17) = 3.76$, $p<.01$, but there was no significant difference in NO. Thus, the note-taking scores of the two groups with an output opportunity greatly improved, whereas the scores of the group without such an output

Table 3 ***Mean Note-Taking Scores within Groups***

	1st Input		2nd Input		Significance (1st → 2nd Input)
	M	*SD*	*M*	*SD*	*t*
Target Session 1					
NO (n=18)	7	11	10	13	1.39
O (n=20)	11	12	22	9	3.15**
CO (n=18)	6	8	26	19	3.76**
Target Session 2					
NO (n=18)	10	16	19	16	1.96
O (n=20)	22	11	15	12	1.71
CO (n=18)	19	16	23	18	0.77

Note. The scores are presented in percentages. *M* = mean, *SD* = standard deviation.
* p < .05. ** p < .01.

opportunity did not. These results show that output increased the students' attention to the target form, the past hypothetical conditional.

In contrast to the results from Target Session 1, there was no significant difference from the first input to the second input in any groups in Target Session 2. One explanation for this may be that the attention to the target form by the two output groups, O and CO, had already been enhanced during the first input in Target Session 2, because of their experience of output in Target Session 1. In fact, as Table 3 shows, the mean scores of O (22%) and CO (19%) in the first input in Target Session 2 were much higher than those in the first input in Target Session 1 (O=11%, CO=6%). That is, in Target Session 2, from the beginning the participants' noticing of the target form was quite high. This may account for the lack of a significant difference between the first input and the second input in Target Session 2.

The note-taking scores were also compared between groups, NO vs. O as well as O vs. CO. There was a significant difference between NO and O in the second input in Target Session 1, $t(36) = 3.32$, $p<.01$, and in the first input in Target Session 2, $t(36) = 2.65$, $p<.05$. These results indicate that output contributed to an increase in the students' attention to the past hypothetical conditional. On the other hand, no significant difference was found between O and CO. Thus, it can be assumed that there was no difference between working individually for output and working collaboratively for output in terms of the effectiveness on the noticing of the target form.

3.2 Noticing – Text-Reconstruction Scores

The text-reconstruction scores were also utilized to examine the noticing of the past hypothetical conditional features in O and CO. Targetlike use (TLU) analysis was used to score the students' production during the text-reconstruction. For the O group, each student's score was calculated dividing the number of correctly formulated sentences by the number of sentences he attempted to make. The same method was used for calculating each pair's score in the CO group[15].

In Target Session 1, there were significant differences between the first reconstruction and the second reconstruction in terms of *if* and main clauses combined, $t(7) = -2.61$, $p<.05$, and main clauses, $t(8) = -2.55$, $p<.05$, in the O group. In the CO group, there was a significant difference in terms of main clauses, $t(15) = -2.34$, $p<.05$. In Target Session 2, there was a significant difference in *if* and main clauses combined, $t(17) = -2.11$, $p<.05$ in O, while in CO, all three items showed a significant difference between the first reconstruction and the second reconstruction: *if* & main clauses, $t(17) = -4.53$, $p<.01$; *if*-clauses, $t(17) = -2.67$, $p<.05$; and main clauses, $t(17) = -3.34$, $p<.01$, in CO.

These results indicate that an output opportunity enabled both types of output groups to increase their accuracy in making past hypothetical conditional sentences. It seems reasonable to suppose that, after engaging in the first text reconstruction, the participants paid more attention to the target form in the subsequent input than in the first input. In brief, producing output resulted in increasing the participants' noticing of the target form, which subsequently led to a significant improvement in making the target sentences accurately.

There was no significant difference between O and CO, except in the *if*-clauses in the second reconstruction in Target Session 1, $t(31) = 2.08$, $p<.05$. This indicates that the O participants could make *if*-clauses more accurately than the CO participants in Target Session 1, although there was no difference between the two groups in Target Session 2.

3.3 Acquisition – Multiple-Choice Recognition Test Scores

This section examines the issue of acquisition based on the results of the study. This is addressed by presenting, (1) the multiple-choice recognition test scores, and (2) the production test scores. Table 4 shows the results of the multiple-choice recognition test within groups. In both O and CO, there was a significant increase in the scores from Pretest to Posttest I, O, $t(19) = -2.44$, $p<.05$, and CO, $t(17) = -3.04$, $p<.01$, from Posttest I to Posttest II, O, $t(19) = -2.50$, $p<.05$, and CO, $t(17) = -3.52$, $p<.01$, and from Pretest to Posttest II, O, $t(19) = -4.51$, $p<.01$, and CO, $t(17) = -5.52$, $p<.01$. In NO, there was a significant increase from Pretest to Posttest II, $t(17) = -2.77$, $p<.05$. There was no significant difference between any tests in the control group (C). These results indicate that the output task greatly contributed to improving the recognition test scores, and the input only task was also effective when it was conducted repeatedly. These findings may support the effectiveness of task repetition on L2 learning (e.g., Bygate 2001).

In terms of the results of the multiple-choice recognition test scores between groups (C and NO, NO and O, and O and CO), although there was a significant difference between NO and O on Posttest II, $t(36) = -2.93$, $p<.01$, there was no difference on any tests between O and CO. Additionally, a significant difference was found between NO and C, on Posttest II, $t(39) = -2.67$, $p<.05$.

The results indicate that the task with output resulted in increasing the scores of the multiple-choice recognition test scores more than the task without output; thus, output was more effective for the acquisition of the past hypothetical conditional than input only. However, it appeared that there was no difference between producing output individually and collaboratively in terms of the contribution to the acquisition of the form. Furthermore, the NO participants, who engaged in the

Table 4 *Multiple-Choice Recognition Test Scores within Groups*

	Pretest		Posttest I		Posttest II		Significance		
	M	*SD*	*M*	*SD*	*M*	*SD*	(Pretest to Posttest I)	(Posttest I to Posttest II)	(Pretest to Posttest II)
Group							*t*	*t*	*t*
C	5.70	3.53	6.96	4.92	7.17	3.01	-1.50	-0.31	-2.01
NO	7.11	2.56	8.61	3.90	10.22	4.30	-1.57	-1.87	-2.77**
O	8.20	4.03	11.80	6.19	14.85	5.33	-2.44*	-2.50*	-4.51**
CO	6.89	3.05	11.56	5.51	14.83	4.91	-3.04**	-3.52**	-5.52**

Note. Maximum scores were 20. * p < .05. ** p < .01.

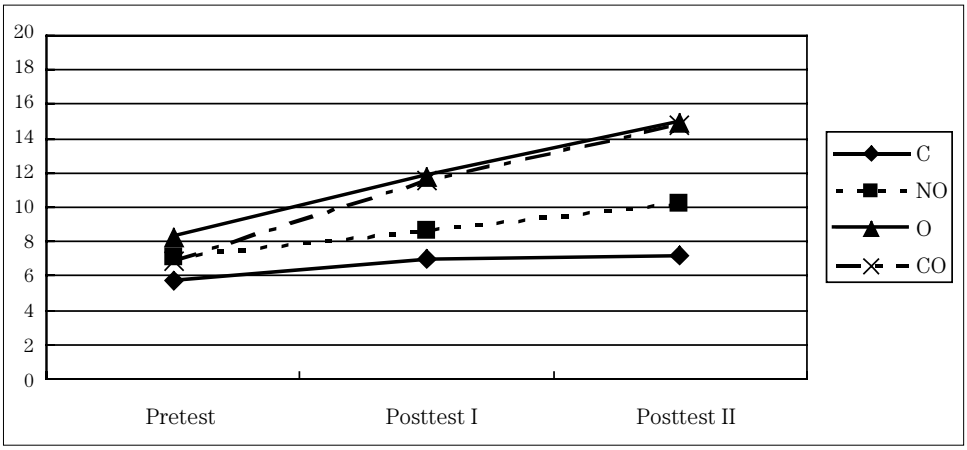

Figure 3 Multiple-Choice Recognition Test Mean Scores

input task twice, achieved significantly higher scores than those in the control group who only took the tests. Thus, it can be said that the accumulation of knowledge from the input tasks was effective in the learning of the past hypothetical conditional.

3.4 Acquisition – Production Test Scores

Table 5 shows the results of the production test based on the TLU analysis. Regarding the results from each group, O showed a significant increase concerning all three items, *if* and main clauses combined, *if*-clauses only, and main clauses only on all the tests: (1) Pretest-Posttest I, *if* & main, $t(19) = -3.61$, p<.01; *if*-clauses, $t(19) = -2.59$, p<.05; and main clauses, $t(19) = -3.55$, p<.01, (2) Posttest I-Posttest II, *if* & main, $t(19) = -4.57$, p<.01; *if*-clauses, $t(19) = -3.93$, p<.01; and main clauses, $t(19) = -3.73$, p<.01, (3) Pretest-Posttest II, *if* & main, $t(19) = -6.64$, p<.01; *if*-clauses, $t(19) = -5.89$, p<.01; and main clauses, $t(19) = -6.64$, p<.01. It appeared that output had a great impact on improving the students' production scores.

CO made a significant increase in terms of *if* and main clauses combined on all the tests: (1) Pretest-Posttest I, *if* & main, $t(17) = -4.38$, $p<.01$, (2) Posttest I-Posttest II, *if* & main, $t(17) = -2.91$, $p<.01$, (3) Pretest-Posttest II, *if* & main, $t(17) = -7.20$, $p<.01$. In addition, regarding the *if*-clauses only and the main clauses only, there was a significant difference between Pretest and Posttest I, *if*-clauses, $t(17) = -3.68$, $p<.01$; main clauses, $t(17) = -2.81$, $p<.05$, as well as Pretest to Posttest II, *if*-clauses, $t(17) = -6.54$, $p<.01$; and main clauses, $t(17) = -5.95$, $p<.01$. Hence, collaborative output work also seemed to greatly contribute to promoting the participants' accuracy of the past hypothetical conditional sentences.

In NO, there was a significant increase concerning all three items between Posttest I and Posttest II, *if* & main, $t(17) = -4.06$, $p<.01$; *if*-clauses, $t(17) = -2.87$, $p<.05$; and main clauses, $t(17) = -2.72$, $p<.05$, as well as Pretest and Posttest II, *if* & main, $t(17) = -4.34$, $p<.01$; *if*-clauses, $t(17) = -3.98$, $p<.01$; and main clauses, $t(17) = -2.40$, $p<.05$. Thus, being exposed to input only also apparently improved the participants' accuracy in making the past hypothetical conditional sentences. There were no differences between any tests in C.

Table 5 ***Production Test Scores within Groups, Targetlike Use Analysis***

	Pretest		Posttest I		Posttest II		Significance		
	M	*SD*	*M*	*SD*	*M*	*SD*	(Pretest to Posttest I) *t*	(Posttest I to Posttest II) *t*	(Pretest to Posttest II) *t*
C									
if & main	1.09	2.35	1.57	2.89	1.83	2.48	-1.27	-0.59	-2.01
if-clause	0.61	1.27	0.78	1.57	0.96	1.69	-0.81	-0.57	-1.70
main clause	0.48	1.47	0.78	1.98	0.87	1.58	-1.16	-0.24	-1.44
NO									
if & main	1.17	2.23	1.73	3.88	4.78	4.08	-0.90	-4.06**	-4.34**
if-clause	0.67	1.41	1.28	2.65	2.44	2.99	-1.68	-2.87*	-3.98**
main clause	0.50	1.29	0.44	1.29	2.33	2.93	0.17	-2.72*	-2.40*
O									
if & main	1.06	1.28	5.28	5.28	10.50	6.73	-3.61**	-4.57**	-6.64**
if-clause	0.70	1.08	2.30	2.72	4.95	3.61	-2.59*	-3.93**	-5.89**
main clause	0.35	0.67	2.80	3.12	5.30	3.39	-3.55**	-3.73**	-6.64**
CO									
if & main	1.28	2.02	6.28	4.70	10.39	5.40	-4.38**	-2.91**	-7.20**
if-clause	0.56	1.20	3.61	3.38	5.50	3.15	-3.68**	-2.05	-6.54**
main clause	0.72	1.49	2.67	3.36	4.44	3.17	-2.81*	-1.99	-5.95**

Note. Maximum scores were 16 for *if* & main clauses combined, 8 for *if*-clauses, and 8 for main clauses.
* $p < .05$. ** $p < .01$.

Again, the results obtained from within each group seem to suggest that repetition of a task which focuses on the same grammatical structure has an effect on the learning of the structure.

With regard to the results of the production test scores between NO and O, as well as O and CO, there was a significant difference between NO and O regarding all the items on Posttest I, *if* & main, $t(36) = 2.23$, $p<.05$; and main clauses, $t(26) = 3.09$, $p<.05$ (Welch's t test) as well as Posttest II, *if* & main, $t(36) = 2.99$, $p<.01$; *if*-clauses, $t(36) = 2.32$, $p<.05$; and main clauses, $t(36) = 2.87$, $p<.01$, except *if*-clauses on Posttest I. No difference was statistically found between O and CO. Between C and NO on Posttest II, there were significant differences regarding *if* and main clauses combined, $t(39) = -2.86$, $p<.01$; and main clauses only, $t(36) = -2.05$, $p<.05$.

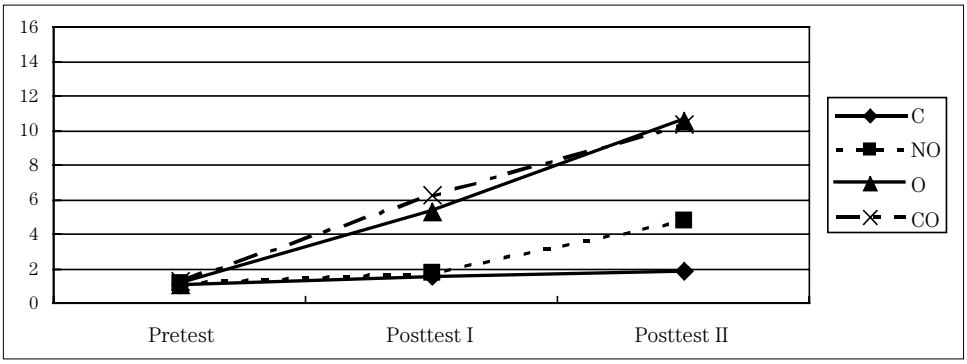

Figure 5 Production Test Mean Scores, Targetlike Use Analysis
Note. Maximum score was 16.

The data from the production test scores showed the same results as obtained from the multiple-choice recognition test data. Based on the TLU analysis of the production test scores, it appears that output contributed to improving the accuracy of the target form more than input only. However, again, working for output alone or together did not seem to make a difference in the learning of the form. Moreover, after engaging in the task twice, the learners who received input only improved their production test scores more than those in the control group who only took the tests. Therefore, this result also indicates that the accumulation of the task, in other words, being exposed to a great amount of input seemed to improve the accuracy of the past hypothetical conditional.

4. Conclusion and Implications for Japanese Secondary School

The first hypothesis, that the participants who are required to produce output (O) will notice the

past hypothetical conditional in input more and attain higher scores on the posttests, in terms of the form, than the participants who have no such opportunity (NO), was confirmed by the present study. Since output was the only activity that the learners engaged in between the two input trials, it is reasonable to say that output promoted the learners' noticing of the past hypothetical conditional. Output, rather than input only, resulted in greater learning of the target form.

Regarding the second hypothesis, which was the comparison of the O group and the Collaborative Output (CO) group, neither the noticing issue nor the acquisition issue, were supported. Contrary to the hypotheses, the present study found that individual output and collaborative output equally contributed to the noticing and the acquisition of the past hypothetical conditional[16].

The participants of the present study were first-year college students who had studied in Japanese secondary school. Their pretest results (Tables 4 and 5 in Section 3) revealed that before receiving any treatment given in the present study, they had not acquired the target form, the past hypothetical conditional. As the analysis of several authorized textbooks briefly reviewed in Section 2 indicates, it is probable that the participants for the present study had received input of the past hypothetical conditional through reading, but they did not have any opportunities to listen to it or orally produce it in the oral communication context. Furthermore, although the "Writing" course appears to take up the target form in detail, it is an elective course. Hence, it is possible that some learners did not have many opportunities to produce it. In other words, the study points out a lack of output compared to input in the L2 classroom at secondary school in Japan.

Based on the findings, the author argues that L2 classrooms should implement plenty of tasks which push or require learners to produce output. That is because the statistical analysis of the data obtained from L2 classrooms for this study supported that output promotes the noticing of linguistic forms, which Schmidt (1990) claims is crucial to L2 learning. In addition, both types of output, working for output alone and working for output together, enabled the learners to notice the gap between what they produced and what the appropriate linguistic forms should be. The subsequently provided input after output gave them an opportunity to seek the necessary information to fill the gap. Hence, it can be assumed that the follow-up input made it possible for them to compare their IL and the correct linguistic forms. Furthermore, an opportunity to revise the text, which was provided after the subsequent input, enabled the learners to modify their output by incorporating the information they found. Therefore, in addition to an output opportunity, it is significant to provide learners with opportunities to receive model input as well as to modify their own output in L2 classrooms.

Another implication is that *focused tasks*[17] such as the text-reconstruction task used in the present

study or a dictogloss task should be utilized to encourage learners to pay attention to the form on which the teacher expects them to focus. Focused tasks account for one of the reasons why output worked well to promote the noticing and the learning of the past hypothetical conditional in this study. Furthermore, the learners in each experimental group respectively worked on the same type of focused task twice: the NO group – the comprehension task, the O group – the text-reconstruction task, and the CO group – the text-reconstruction task in pairs. The test scores from not only the output groups (O and CO), but also the non-output group (NO) significantly increased from Pretest to Posttest II. Thus, the fact that the learners repeatedly engaged in a task which focused on the target form led to greater learning. In short, the repetition of the focused task was effective in drawing the learners' attention to the form, which consequently facilitated their learning.

One of the characteristics of focused tasks is to draw learners' attention to form while they are working on meaning. In brief, at first, their attention is on meaning. Within oneself or in groups, they may simply discuss the meaning of a word, sentence, or passage in the task. Whatever linguistic items they discuss, they are likely to focus on meaning first, as VanPatten (1990) suggests. Then, while reflecting on and discussing different points about the language, their attention might naturally be drawn to the form of the meaning.

This context in which learners' attention is drawn to the L2 form while focusing on meaning is called *focus on form* (Long 1991). The importance of instruction using *focus on form* tasks or techniques has been widely supported in second language research recently (e.g., Doughty and Williams, 1998). The present study revealed that focused tasks, for example structure-based output tasks such as the text-reconstruction task and the dictogloss task, can be considered especially effective in drawing learners' attention to the L2 form in the context of communicative activities. Implementation of these tasks should be encouraged in L2 classrooms.

According to *the Course of the Study*, the overall objective is to develop students' practical communication abilities. As long as this is the goal of the learning of English, it is important to provide plenty of opportunities for students to not only receive input but also produce output in the L2 classroom.

Focused tasks such as the text-reconstruction task used in the present study give learners opportunities to receive a large amount of input as well as to produce output. Additionally, it creates a meaningful context, where their attention is first drawn to meaning and then to form. More importantly, if a text-reconstruction task is engaged in in pairs, as was done with the CO group in this study, it offers learners many opportunities to engage in actual communication with others using the target language, negotiate with each other, and work together to solve any linguistic problems they

encounter. These activities in pairs greatly contribute to co-constructing linguistic knowledge, which leads to an individual learner's learning[18]. Although the study found that both types of output were equally effective for noticing and L2 learning, focused tasks through group work may provide even more meaningful context with learners.

Last, in addition to the importance of providing appropriate instruction and support before and during the task, it is crucial to give feedback to learners after completing the task. The present study found that the learners were able to use the linguistic items, which they discussed and correctly solved in pairs during the task, in the later individual performance of the same task. The study also found some instances in which incorrectly solved items from the pair work were used incorrectly by the learners later alone. Swain (1998) indicates that learners are likely to remember incorrect solutions which might be made through collaborative dialogue. Therefore, whether learners work alone or collaboratively, it is important for teachers to pay careful attention to each individual learner's performance during the task, as well as to their final product such as a written report.

This study emphasizes that output can be produced by engaging in various cognitive activities, and through output, learners reflect upon their own language use. They think, put their ideas together, teach, and scaffold each other in the case of CO. In the end, they build new knowledge. Output activates these cognitive activities; thus, it contributes to learning. Furthermore, what we produce through cognitive activities becomes a product at the end, which we can reflect upon again. Hence, output is both a cognitive activity and its product. Considering these roles of output, one cannot deny its great potential for L2 learning.

The present study supports the effective role of output; however, we need to further research what types of output and output tasks are most effective for L2 learning. The nature of interaction where output is produced collaboratively should be further examined as well. It is crucial to conduct more research on output and accumulate data to reveal the mechanisms by which output leads to second language acquisition and best contributes to the L2 classroom in terms of teaching and learning.

Notes

1 Krashen urges that language acquisition occurs when learners can access comprehensible input. Specifically, comprehensible input which contains "i + 1" is crucial to SLA. He defined "i" as a learner's current state of knowledge and "i + 1" indicates the next level. In other words, input containing structures which are slightly beyond the level of a learner's current grammatical knowledge is useful.

2 The hypothesis testing function is when learners use output to test how their hypothesis about the target language actually works.

3 The metalinguistic function of output is when learners produce output to consciously reflect on their

language.

4 Due to the page limit for this paper, the analysis through an approach based on the sociocultural theory is not presented here. For the results from the sociocultural perspective, refer to M. Nagasaki (2007).

5 The study referred to *the Course of the Study* implemented in 2002 for lower secondary school and the one implemented in 2003 for upper secondary schools. All of the textbooks for upper secondary school analyzed here were published in or after 2003. See references for the textbooks examined in the study.

6 English I & II focus on reading.

7 CO engaged in the same tasks they did collaboratively later alone in order to investigate whether what they discussed in pairs leads to individual learner's learning. This research question was investigated using the sociocultural approach. The results were not described here for the same reason as listed in note 4.

8 Text 1 was about the story of Forrest Gump's childhood, while Text 2 was about Forrest's youth.

9 For the later analysis of the audiotaped conversations from each pair, the teacher intentionally paired male students with female students as much as possible to make it easier to recognize each participant's voice in the conversation. Aside from this condition, the teacher paired up the CO participants randomly.

10 The text-reconstruction task which CO engaged in may be considered a dictogloss task, described in Section 1.4, as well. The differences between the two types of tasks are: (1) in the text-reconstruction task, the learners read the passage; whereas, in dictogloss they listen to the passage, (2) the text-reconstruction task asks learners to reconstruct the text as closely to the original as possible, while dictogloss encourages them to reconstruct the text in their own words, and (3) the text-reconstruction task requires learners to reconstruct the text individually, while dictogloss requires them to reconstruct it collaboratively in small groups.

11 The O and CO participants were asked to use a red pen to revise the text, so that the changes they made from the 1st reconstruction to the 2nd reconstruction would be clear.

12 The survey, made based on the one used in Slimani's study, asked the students to write down any words, spelling, grammar, language usage, and other things they could recall or learned from the task. However, the results of the survey are not reported here for the same reason as in note 4.

13 The contextualized pictures and the instructions provided created a situation in which the use of the past hypothetical conditional was required. In addition, a starter (e.g., *if John*) was provided for each target item in order to encourage the students to use the target form. Furthermore, the verbs were given for each target item to help the students comprehend the meaning of the two pictures.

14 The IL scoring system was utilized to detect the specific IL features which tended to be used in a non-targetlike way. The results based on the IL analysis are omitted here for the same reason as in note 4. For the analysis and results using the IL scoring system, refer to M. Nagasaki (2007).

15 Only the scores of the students who attempted to make at least two target sentences (at least two *if*-clauses as well as at least two main clauses) were calculated in order to avoid having unfair percentage scores coming from the calculation using small denominators.

16 The statistical analysis showed that there were no significant differences between individual output and collaborative output; however, further analysis of the collaborative output based on the sociocultural approach revealed its great potential for L2 learning. Refer to M. Nagasaki (2007).

17 The types of tasks used for each group in this study are all categorized as *focused tasks*. Ellis (2003) indicates that focused tasks are employed to elicit use of specific linguistic features, either by design or by the use of methodological procedures that focus attention on form in the implementation of the task. The present study argues that engaging in these focused tasks accounts for the greater learning of the three groups. This may explain why even engaging in the same type of task resulted in improved posttest scores.

18 See 16 above.

References

Celce-Murcia, M., & Larsen-Freeman, D. (1999). *The grammar book* : *An ESL/EFL teacher's course* (2nd ed). Boston: Heinle & Heinle.

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* (197-262). New York: Cambridge University Press.

Ellis, R. (2003). *Task-based language learning and teaching*. Oxford University Press.

Izumi, S. (2002). Effects of output and input enhancement. *Studies in Second Language Acquisition*, 24, 541-577.

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.

Izumi, S., & Bigelow, M. (2000). Does output promote noticing and second language acquisition? *TESOL Quarterly*, 34, 239-278.

Izumi, Y., & Izumi, S. (2004). Investigating the effects of oral output on the learning of relative clauses in English: Issues in the psycholinguistic requirements for effective output tasks. *The Canadian Modern Language Review, 60*, 587-609.

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

Krashen, S. (1985). *The Input Hypothesis: Issues and implications*. London: Longman.

Long, M. (1991). Focus on form: A design feature in language teaching methodology. In K. de Bot, R. Ginsberg, & C. Kramsch (Eds.), *Foreign language research in cross-cultural perspective* (pp. 39-52). Amsterdam: John Benjamins.

Nagasaki, M. (2007). The Role of Output and Collaborative Dialogue in Second Language Learning: The Effect of Learner Production on the Learning of English Past Hypothetical Conditional. Unpublished Dissertation.

Schmidt, R. (1990). The Role of Consciousness in Second Language Learning. *Applied Linguistics, 11*, 129-158.

Skehan, P. (1998). A cognitive approach to language learning. Oxford University Press.

Swain, M. (1985). Communicative competence: Some roles of comprehensible input and comprehensible output in its development. In S. Gass & C. Madden (Eds.), *Input in second language acquisition* (pp. 235-253). Rowley, MA: Newbury House.

Swain, M. (1995). Three functions of output in second language learning. In G. Cook & B. Seidlhofer (Eds.), *Principle and practice in applied linguistics* (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*. Cambridge University Press.

Swain, M. & Lapkin, S. (1995). Problems in output and the cognitive processes they generate: A step towards second language learning. *Applied Linguistics*, 16, 371-391.

The Ministry of Education, Culture, Sports, Science, and Technology. (2003, March). *The Course of the study for foreign languages*. (2003, March). Retrieved November 26, 2007, from http://www.mext.go.jp/english/shoutou/030301.htm

The Ministry of Education, Culture, Sports, Science, and Technology. (1999). *The guidebook for the course of the study for foreign languages*. Tokyo: Kairyudo.

VanPatten, B. (1990). Attending to content and form in the input: An experiment in consciousness. *Studies*

in Second language Acquisition, 12, 287-301.

Authorized English Textbooks

Birdland Oral Communication I. (2004), *Birdland Oral Communication II*. (2004), *Powwow English Writing*. (2004), *Open Door to Oral Communication II*. (authorized in 2003). Kyoto: Buneido / *Crown English Series I*. (2003), *Crown English Series II*. (2004), *Crown English Writing*. (2005). Tokyo: Sanseido / *Departure Oral Communication I*. (2004). Tokyo: Taishukan / *Empathy Oral Communication II*. (2004). Tokyo: Kyoiku Shuppan / *Polester Writing Course* (2004). Tokyo: Suuken / *Prominence English I*. (2006), *Prominence English II*. (2004), *Prominence English Reading*. (2005). Tokyo:Tokyo Shoseki / *Pro-Vision English Course I*. (2003), *Pro-Vision English Course II*. (2006), *Pro-Vision English Reading*. (2005). Tokyo: Kirihara / *Sailing Oral Communication I*. (2003). Tokyo: Keirinkan / *Screenplay Oral Communication II*. (2003). Nagoya, Aichi: Screenplay / *Voice Oral Communication II*. (2004), *Voyager English Course I*. (authorized in 2007), *Voyager English Course II*. (2007), *Voyager Reading Course*. (2006), *Voyager Writing Course*. (2005). Hiroshima: Daiichi Gakushusha