The Long-term Effects of Reflective Activities on Oral Output

YANAI Tomohiko*

Abstract

This study examines the effects on speaking ability resulting from learners' reflection on their own oral output. The participants in the study were two groups of Japanese university students. Both groups were trained so that they could give longer and more informative answers to questions. The students in one group were required to do a variety of reflective activities as homework assignments, while those in the other group were required only to listen to their own recorded output in class without doing reflective activities. An analysis of the students' linguistic performance in the final examination showed positive long-term effects for reflection tasks on syntactic complexity and lexical variety but no effects on fluency or accuracy. The results suggest that reflective post-task activities may be effective for specific domains of interlanguage development.

【Key words】 Oral output, learners' reflection, interlanguage development

Introduction

The focus of this study is on the role of learners' reflective activities for the development of speaking ability. Two groups of students participated in the study. One of the groups was given a homework for reflection on their oral production, while the other was not. It was assumed that the former group would have the opportunity to 'notice', by reflection, what they did not know or knew only partially. The reflection was further expected to help them acquire new knowledge or modify their existing knowledge.

With respect to the role of output and noticing for second language learning, Swain and Lapkin(1995) stated:

> in producing the L2, a learner will on occasion become aware of (i.e. notice) a linguistic problem (brought to his/her attention either by external feedback (e.g. clarification requests) or internal feedback). Noticing a problem 'pushes' the learner to modify his/her output.

Here, 'noticing' is meant to be a cognitive process learners engage in when they find their own output

Received October 27, 2004.

*柳井 智彦 (やない・ともひこ), Department of English, Faculty of Education and Welfare Science, Oita University, 700 Dannnoharu, Oita 870-1192, Japan

problematic, in contrast to other researchers' definition of the term that 'noticing' is what occurs when a learner pays conscious attention to linguistic input (see Ellis, 1997, p.141). The present study examines the significance of reflective activities assumed to promote noticing in Swain and Lapkin's sense.

Several classroom activities for promoting noticing have been proposed. For example, Swain and Lapkin(1998) used a jigsaw task where a pair of students noticed their linguistic problems while cooperating to make up a story line for a series of pictures. Lynch(2001) let learners transcribe their own output to improve their language. Similarly, Mennim(2003) required the students to transcribe a rehearsal of their oral presentation for correction of forms and improvement of pronunciation.

The practice reported in this paper is different from the above studies in three respects. Firstly, reflective noticing activities were not restricted to any specific type; learners were free to choose from a variety of optional activities according to their interest and ability. These ranged from an activity of finding their linguistic problems by listening to their own output, to an activity of collecting useful vocabulary related to the topics of the oral practice. Secondly, the students in this study continued the reflective activities for as long as 15 weeks. The participants in the above studies, on the other hand, received a short-term treatment, i.e. five weeks in Swain and Lapkin(1998) and probably several weeks in Lynch(2001) and Mennim(2003). Thirdly, this study will show that reflective activities do make a difference to learners' long-term performance, a point which Lynch(2001) needed to explore.

The oral ability which it was intended to develop and examine in this study centers on the production of more elaborate responses to questions. Language teachers have been dissatisfied that learners' answers to teacher questions are extremely short and simple. Wu(1993), for example, demonstrated that the overwhelming majority of responses to teacher questions collected in the study were 'restricted' (in contrast to 'elaborate') answers, as the following extracts represent:

T: Have you been to an airport before?
S: No.

T: Where do you see it?
S: On television.

(Wu, 1993, p.57)

Teaching procedure and the independent variable

The two groups of participants in this study each attended an 'Oral English' class lasting for one semester in different years. Each ninety-minute class consisted of two parts: communication activities based on the textbook (60 minutes), and 'response exercises' (30 minutes). The 'response exercises', specially developed for this study, were given independently of the contents of the textbook. The exercises were made up of three sections. Section 3 differed between the two groups, being the independent variable of the study. The content of each section was as follows:

Section 1: Presentation of a target strategy with examples.

Here I demonstrated a speaking strategy which would help learners produce longer and more

elaborate answers. An example is: 'Include words of *frequency* in your answer when appropriate'. The following are the demonstration sentences:

Do you use a computer?

--Yes, I *often* use one.

--Well, I use one *several times a month.*

Other speaking strategies covered the notions of 'degree', 'time', 'place', and so on. One strategy was presented in one lesson.[1)]

Section 2: Applying the strategy in the response exercises.

Students listened to the questions asked by the teacher through the headset in the language laboratory and answered them, recording both the teacher's questions and their answers on the tapes. They were given 15 seconds for a response and directed to say as many things as possible.

Section 3: Reflection on the oral performance: the independent variable.

Students in one group replayed the tapes immediately after they had finished answering the questions and listened to them to see how well they had performed. This immediate reflection was the only reflection required of this group. Students in the other group did not do the immediate reflection. Instead, they were required to hand in, the next week, their reflective studies written at home on a small worksheet. Because the teacher predicted that students would find it difficult to know how to do the reflective studies on their own, he delivered a guide for help. The guide presented a list of activities, as shown below, from which students chose according to their interest and ability. The guide offered two types of activities, one named 'Direct reflective activity', and the other, 'Indirect reflective activity'. The former required students to listen to their recordings on the tape to hear exactly what they had spoken in the class, while the latter did not require students to listen.

(Direct reflective activities)

--Writing a new version of answers reformulated from the recorded answers on the tape.

--Making retrospective comments on the English recorded on the tape.

(Indirect reflective tasks)

--Collecting useful words and phrases for the topics dealt with in the response exercises in class by consulting dictionaries and reference books.

--Copying useful words and sentences given in class and memorizing them.

-- Noting weak points in grammar.

--Consolidating the day's points.

--Writing answers to the questions in the textbook.

--Doing creative composition by using important expressions presented in the class.

-- Making questions to put to the teacher.

-- Tasks created by the student.

Method

Subjects

Two groups of students at a state-run university were the subjects in this study. Both groups

consisted of 41 second-year students. One group attended 'Oral English' taught by this author in the year of 2002. They were required to carry out the reflective activities as homework assignments. This group will be called the RE group (REflection group) hereafter. The other group attended the class of the same title taught by the same instructor in the year of 2001. As this group was not required to do the reflective activities, they will be called the NR group (Non-Reflection group). The students of each group were majoring in the teaching of school subjects in the faculty of education. Their levels of proficiency in English were assumed to be similar, because the distribution of the holders of STEP (Society of Testing English Proficiency) certificates was almost the same: the second grade certificate had been obtained by 4.6% of the RE group and 2.3% of the NR group, and the pre-second grade certificate by 30.2% and 32.6%, respectively.

Oral response test

An oral response test was administered in the language laboratory as part of the final examination of the course. The students were required to answer a series of questions which the teacher asked live and to tape their answers on the audiotapes set in the tape recorders in the individual booths. 15 seconds were allowed for answering each question, the end of the time for responding being signaled by the beep from a timer. At the beginning of the test, the teacher told the students that they should try to say as many things as possible in their answers. The students answered ten questions, all of which had the form of Yes/No questions such as "Do you usually get up early?", "Can you drive a car?", and "Are you going to cook dinner tonight?" The responses to the question, "Does your mother drink beer?" were excluded from the analysis because the question was educationally inappropriate. Eventually the responses to nine questions were subjected to a data analysis. All the questions in the test were similar to those which had been used in the practice materials in class in terms of their contents and grammatical structures, but no questions in the test were identical with those in the practices.

Measures

All the students' responses were transcribed from the audiotapes and analysed in the light of the four aspects of linguistic performance: fluency, accuracy, complexity, and lexical variety. A great many studies have proposed different methods for measuring these aspects and examined the validity of them (see for example, Foster et al., 2000). Because the fuller study of the proposed methods lies outside the scope of a brief paper, the measuring methods adopted in this study will be described.

Fluency in the students' utterances was measured by the number of words in a response. Note that two scores for fluency were obtained. For one score, repeated or reformulated words were counted, and for the other score, they were not counted. If such words are counted, a score of 'unpruned tokens' is obtained; if not, one of 'pruned tokens' is calculated. The results by the two counting methods will be presented later. Accuracy was measured by counting the number of error-free clauses. Syntactic complexity of the students' utterances was measured by the number of words per T-unit (see Hunt(1966) for the definition of T-unit and the determination of its length). A T-unit consisting of a main clause and a subordinate clause is usually longer and looks more complex

than a simple sentence with no subordinate clause. In this study, only error-free T-units were subjected to word count for complexity. Furthermore, short formulaic expressions such as 'Yes, I do' or 'No, I wasn't' signaling whether the response was affirmative or negative were also excluded from the data. This was done because these expressions, if regarded as a T-unit, would make the mean length of T-unit for a speaker unreasonably short. Finally, lexical variety was measured by counting the number of 'types', or valid different words, in a response. This is equal to the number of different words remaining after eliminating firstly, words which were used in the corresponding question, and, secondly, words which are repeated within an answer.

In the following example illustrating the assessment of lexical variety, the underlined words are not counted as types because they have occurred in the question, and the italicized words are not counted because they are repeated. The number of types in this example is therefore nine.

(Example)

Q: Did you eat bread this morning?

A: Yes, I did. *I* usually eat bread or rice in the morning. *I* am ate *bread in* this *morning*. (number of types=9)

The counting method for types, by which even the words used in the corresponding question are eliminated, is peculiar to this study. The method was created to obtain the number of different words which a learner used with no borrowing from the interlocuter's expressions.

Results

Fluency

As shown in Table 1, the RE group achieved a higher average number of words per response than the NR group when repetitions and reformulations were included in the analysis (i.e.score of unpruned tokens). However, there was no significant difference between the mean scores of the two groups when repetitions and reformulations were excluded (i.e.score of pruned tokens).

Accuracy

As Table 1 shows, the NR group produced a total of 667 clauses for their responses, of which 419 clauses (62.8%) were grammatically correct, while the RE group produced 572 clauses, of which 339(59.3%) were correct ones. There was no statistical difference between these percentages.

Complexity

As Table 1 shows, the mean T-unit length was 5.12 words for the NR group and 5.65 words for the RE group. It was revealed that RE students' T-units were significantly longer than NR students', mainly because they used modifiers (adverbials and adjectival phrases) and infinitives more frequently. T-units containing subordinate clauses, which are generally a major cause of syntactic complexity, were small in number for both groups (2.7% for the NR group and 4.6% for the RE group). Thus, RE students appear to have used more words, but usually in simple sentences.

When RE students listened to their recorded output and attempted to make a reformulated

version from it as a reflective activity, they often used more words to express their ideas more fully, as the example below demonstrates.

(Example)

Question: Have you ever enjoyed a barbecue?

(Audio-tape) Yes, I have. I have enjoyed a barbecue.

Table 1 Mean scores for the four aspects of linguistic performance

	NR group	RE group		significance
Fluency				
unpruned tokens per response	12.21	13.14	$t(736)=2.79$	$p<0.01$
pruned tokens per response	11.33	11.61	$t(736)=0.94$	ns
Accuracy				
total number of clauses	667	572		
number of erro-free clauses	419(62.8%)	339(59.3%)	$\chi^2(1)=1.64$	ns
Complexity				
length of error-free T-units	5.12	5.65	$t(601)=3.47$	$p<0.01$
Lexical variety				
types per response	6.36	7.67	$t(736)=6.42$	$p<0.01$

(Reformulation) Yes, I have. I have enjoyed a barbecue near the sea with friends last summer.

RE students seem to have gradually acquired skill in expressing their ideas more precisely, gaining a higher score for complexity in the final test.

Lexical variety

It was found that the scores for the number of types (i.e. valid different words) varied significantly between the two groups. The RE students used a greater variety of words in response to a question than the NR students. It should be noted, however, that the present study did not use the more common 'type/token ratio' method of measuring lexical variety. Instead, as mentioned earlier, a method for counting 'types' peculiar to this study was employed, which took into account the interlocutor's use of words. Valid types in the learner's utterance were calculated with reference to words which had been used in the question, not solely with reference to the learner's utterance.

Discussion

The results of this study suggest that learners' reflective activities about their oral performance

have significant effects on their interlanguage development, particularly on linguistic complexity and lexical variety. Giving learners ample opportunities for oral production is, as the Output Hypothesis has advocated, necessary for second language learning, but 'just speaking and writing are not enough.'(Swain, 1993). The students in the NR group in the present study were given the same amount of practice in oral output as the RE group, but their performance in the final test was inferior to that of the students in the RE group, who had been required to carry out the post-task activities for reflection.

The reflective activities, however, were effective only for specific domains of interlanguage development. No significant differences were obtained for accuracy or fluency (when measured by pruned tokens) between the two groups in the study, whereas significant differences in their complexity scores and lexical variety scores were found. One reason for the lack of significant difference in the accuracy scores may be that RE students, when carrying out their reflection, were more eager to find appropriate words and expressions than to pay attention to linguistic forms. Scrutinizing words and expressions may have become a habit for RE students through their weekly reflection, resulting in better scores for lexical variety and complexity in the final test, but not for accuracy of forms. This interpretation is in accordance with the 'trade-off' hypothesis between accuracy and complexity proposed by Foster and Skehan(1996). As to the lack of significant difference in fluency measured by the average number of pruned tokens per response, this seems to have been caused by the same habit as mentioned above. Because of their repetitive, reflective activities at home, leading them to observe and reformulate their own utterances, RE students seem to have become more cautious of choosing words and expressions, resulting in more frequent use of such dysfluency markers as repetitions or reformulations. The number of unpruned tokens therefore increased but that of pruned tokens remained almost the same as for NR students.

Conclusion

This study has offered some evidence for the positive effect of learners' reflective activities in improving their long-term performance. The results show that the reflective activities, when carried out for as long as 15 weeks, did yield superiority in one type of linguistic performance, namely responding to questions. The reflective studies carried out by the learners were not restricted to any specific type. The learners were free to choose among optional reflective methods suggested by the teacher, in contrast to previous studies where only a required type of reflection was attempted for experimental reasons.

It remains necessary to explore what types of reflective activity or what combination of activities would yield better results, since the learners in the study tried many types and combinations of reflective methods and preference for them differed with individual learners and on different days.

Note

[1] It should be mentioned that both groups of students were given 'parroting practices' before they began exercises based on the speaking strategies. In a parroting practice, students were expected to repeat in their

responses the exact words used in a question, paying attention to grammar and expressions, as the following examples show:

Q: Were you a high school student three years ago?
A: Yes. I was a high school student three years ago.
Q: Did you play tennis in the park last Sunday?
A: Yes. I played tennis in the park last Sunday.

One group practised parroting during one class time and the other group during four class times. It is not certain how the different frequency of the parroting exercises affected the results of this study.

References

Ellis, R. 1997. *Second Language Acquisition.* Oxford: Oxford University Press.
Foster, P. and P. Skehan.1996. 'The influence of planning and task type on second language performance'. *Studies in Second Language Acquisition* 18/3: 299-323.
Foster, P., A. Tonkyn, and G. Wigglesworth. 2000. 'Measuring spoken language: a unit for all reasons'. *Applied Linguistics* 21/3: 354-375.
Hunt, K. 1966. 'Recent measures in syntactic development.' *Elementary English*43: 732-739, reprinted in M. Lester (ed.). *Readings in Applied Transformational Grammar*. New York: Holt, Rinehart and Winston, Inc.
Lynch, T. 2001. 'Seeing what they meant: transcribing as a route to noticing'. *ELT Journal* 55/2: 124-132
Mennim, P. 2003. 'Rehearsed oral L2 output and reactive focus on form'. *ELT Journal* 57/2: 130-138.
Swain, M. 1993. 'The Output Hypothesis: just speaking and writing aren't enough'. *The Canadian Modern Language Review* 50/1: 158-164.
Swain, M. and S. Lapkin, 1995. 'Problems in output and the cognitive processes they generate'. *Applied Linguistics* 16/3: 371-391.
Swain, M. and S. Lapkin, 1998. 'Interaction and second language learning: Two adolescent French immersion students working together'. *Modern Language Journal* 82/3: 320-337.
Wu, K. 1993. 'Classroom interaction and teacher questions revised'. *RELC Journal* 24/2: 49-68.

音声アウトプットに及ぼす「振り返り」の効果

柳 井 智 彦

【要旨】 本研究は自身の音声アウトプット（英語の質問への応答）を振り返ることがスピーキングにいかに効果を及ぼすかを検証することを目的とする。2群を構成し，一方には家庭学習として多様な振り返り活動を行わせた。他方の群には振り返り活動は行わせず，単に授業直後に自分の発話の録音を聞く時間のみを与えた。3ヶ月後のスピーキングテストにおいて，統語的複雑さと語彙の多様さにおいて有意な差が生じたが，流暢性や正確さには効果は認められず，この活動は中間言語の発達に限定的に作用したと考えられる。

【キーワード】 音声アウトプット，学習者の振り返り活動，中間言語の発達