Brief Note

A Comparison of Presentation-Practice-Production and Task-based Language Teaching Approaches for L2 Vocabulary Learning of Sign Language

Rika ENOKIDO*, **Atsuko SATO**** and Hideki SAKAI*****

This study experimentally investigated the acquisition of sign language vocabulary by 45 students who are native Japanese speakers and sign language beginners. The aim was to clarify how the presentation-practice-production (PPP) and task-based language teaching (TBLT) approaches affect the acquisition of sign language vocabulary. Following Shintani (2013), the participants were divided into two experimental (PPP and TBLT) groups and a control group, and the experimental groups took three 30-minutes lessons respectively. Task-based and individual word production tests were conducted in pre-test, post-test, and delayed test to measure the effects of the teaching methods. The results showed that both approaches were effective in sign language vocabulary learning; furthermore, it was found that PPP was more effective for nouns and verbs. The results suggest that learners may tend to understand signs roughly due to the iconicity of sign language in the comprehension-type tasks used for the TBLT approach and that some time may be needed for learners to notice the details of the sign language; on the other hand, the explicit instruction used for the PPP approach may help adult learners become aware of the details of the signs and learn them accurately.

Key words: Sign language, second language acquisition, vocabulary, presentation-practice-production, task-based language teaching

I. INTRODUCTION

In second language acquisition (hereinafter referred to as SLA) research, three cognitive processes have been differentiated as a necessary condition for SLA: (a) focus on forms (hereinafter referred to as FonFs), a learner's explicit awareness of the formal aspects of language extracted or isolated from communicative contexts; (b) focus on meaning (hereinafter referred to as FonM), a learner's exclusive focus on the semantic aspects of language; and (c) focus on

form (hereinafter referred to as FonF), a learner's attention to the formal aspects, when needed, during meaningful communicative activities. FonFs leads to synthetic instructional approaches in which at first lexical and grammatical language items are taught explicitly and then learners are provided with activities so that they can apply the explicit knowledge for communicative use. On the other hand, FonM supports communicative approaches in which learners are engaged in communication in the target language and can learn L2 incidentally or implicitly through communication. However, some researchers (e.g., Swain, 1985) pointed out that meaningful input of the target language and opportunities to use

＊ Graduate School of Comprehensive Human Sciences, University of Tsukuba
＊＊ Institute of Human Sciences, University of Tsukuba
＊＊＊ Faculty of Education, Shinshu University

the language for communicative purposes are not sufficient for SLA, and it was proposed that learners' occasional shift of attention to linguistic forms in the course of communication (in other words, FonF) is necessary for SLA (Long & Robinson, 1998). FonF takes place during a negotiation for meaning in interaction between learners and others in the forms of corrective feedback, pushed output, and interactionally modified input (Long & Robinson, 1998). At present, approaches based on FonFs (e.g., presentation-practice-production) and FonF (e.g., task-based language learning) are prevalent in language teaching.

As its name shows, lessons of presentation-practice-production (hereinafter referred to as PPP) consist of three phases. In presentation, new grammar and word usage are explicitly explained; in practice, learners are engaged in activities to improve the skills of using the grammar and words; and, finally, in production, learners utilize the target items in communicative activities (Matsumura, 2017).

Task-based language teaching (hereinafter referred to as TBLT) is an instructional approach in which learners are asked to carry out tasks (communicative activities with clear purposes to be achieved). The target structures or words are not set, but learners can utilize whatever they know. Thus, it is expected that, while learners focus on meaning, they can sometimes direct their attention to linguistic form during task-based activities. The term *task* is used in a narrower sense than in daily life (Matsumura, 2017). Ellis and Shintani (2014) proposed the four conditions of tasks: First, a task focuses primarily on meaning; second, some information gaps exist among its performers; third, performers can use their own resources; and fourth, outcomes of task achievement are clearly set. For example, chatting about a certain topic cannot be considered a task because it does not yield any outcomes indicative of task completion although the first three conditions are met. An activity in which two students have different pictures respectively and are asked to spot the differences in the pictures can be considered a task because all four conditions are met: learners focus on primarily meaning; they have different information (i.e., different pictures); they can use any structures and words that they know; and the outcome is clear (i.e., differences they have found).

The two approaches have been compared to find any relative efficacy. For example, Shintani (2013) investigated whether the PPP and TBLT lessons affect the acquisition of nouns and adjectives for young Japanese children learning English as a second language. The participants were 45 Japanese six-year-olds with no prior experience in L2 learning. The target words were 24 nouns and 12 adjectives. Thirty-minute lessons were given twice a week for five weeks. The PPP lessons involved five activities in which participants were shown the target words, repeated the words, practiced saying the words depicted by picture cards, and were engaged in production activities such as bingo games. On the other hand, the TBLT lessons involved three communicative activities in which the learners moved and picked pictures in response to the teacher's requests and performed bingo games. Learning of the target words was measured in two different ways: a discrete-item word production test and a Same or Different task test. A pre-test was given one week before the lesson, a post-test one week later, and a delayed test four weeks later. The results showed that there was no significant difference between PPP and TBLT instruction for nouns, but that the TBLT group outperformed the PPP group for adjectives. Shintani explained that this was because, in TBLT instruction, the learners needed to clarify adjectives to identify the cards so that more negotiation for meaning and student-initiated production naturally took place than in PPP instruction.

In teaching sign language, SLA theories can be applied. Nakano (2021) conducted a literature review on sign language teaching methods as a

second language. Based on the classification of sign language teaching methods in various countries, the curriculum of Auslan (Australian sign language) and CNSE (lengua de signos española: LSE) employ CBI (content-based instruction) and TBLT. Except for Auslan and CNSE, there are few opportunities for learners to produce sentences in sign language, or, even if there are opportunities, the learners just arrange a part of a conversation sample or are required to use certain expressions. Regarding sign language acquisition in Japan, Nakano (2021) pointed out that the training curriculum for sign language volunteers by the Ministry of Health, Labor and Welfare, one of the curricula for teaching a sign language as a second language, is not necessarily a systematic curriculum, as it takes a teacher-centered approach and has little content on the structure of the sign language, so it is essential to consider how to teach sign languages from the perspective of SLA.

As Nakano reviewed, several sign languages are taught using approaches following SLA theories. It is important to examine which approach (e.g., FonFs or FonF), or both, may be effective for sign language learning, but as far as the authors know, there have been no empirical studies comparing PPP and TBLT approaches in sign language. Therefore, this study, which focuses on vocabulary learning, aims to investigate whether PPP and TBLT approaches affect sign language learning. The four following research questions were set for this study:

RQ1: To what extent are the characteristics of PPP and TBLT reflected in the process features?

RQ2: Does PPP affect the learning of nouns, verbs, and adjectives in sign language for university and graduate students who are beginning to learn sign language as a second language?

RQ3: Does TBLT affect the learning of nouns, verbs, and adjectives in sign language for university and graduate students who are beginning to learn sign language as a second language?

RQ4: Are there any differences between PPP and TBLT in their effects on the production of nouns, verbs, and adjectives in sign language for university and graduate students who are beginning to learn sign language as a second language?

II. METHOD

1. Participants

Forty-five Japanese university or graduate students with no experience or beginners of sign language participated in this study. The 45 participants consisted of 3 freshmen, 4 sophomores, 5 juniors, 23 seniors, and 10 graduate students. The researcher divided the participants into two experimental (TBLT and PPP) groups, and one control group, with 15 participants each.

2. Target Words

The researcher chose 9 nouns, 12 verbs, and 15 adjectives (6 descriptive adjectives and 9 color adjectives) in sign language as target words. They were selected based on the following criteria: the words are familiar to people who speak Japanese; in the case of verbs, hand movement differs from gesture; and, in the case of adjectives, using common descriptors for the target nouns. Most nouns, verbs, and adjectives appear in the textbook *Watashitachi no shuwa gakushu jiten I* (compiled by the "Our Sign Language" reorganization committee, 2021) as fourth- or fifth-grade words in the national sign language test in Japan. The PPP group was exposed to different sets of 12 words each lesson (Table 1). In contrast, the TBLT group was exposed to the 36 target words shown in Table 1 at random and repeatedly over the three lessons.

3. Procedure

(1) Method: All lessons and tests were conducted via Zoom. Two different lessons and materials were designed for the PPP and the TBLT groups. Each group of 15 people was further divided into two classes of 5 to 10 participants. However, when several participants were absent, extra classes were held individually. The two experimental groups had one

Table 1   Three Sets of Target Words for the PPP Group

	Set A	Set B	Set C
Nouns	dog, cat, rabbit	frog, bear, bird	fish, book, pencil
Verbs	stand up, sit down, walk, jump	ski, play soccer, play table tennis, swim	dance, sing, paint, sleep
Adjectives	red, purple, white, big, small	black, gray, blue, heavy, light	green, yellow, brown, long, short

Table 2   Flow of Experiment

Group	PPP	TBLT	Control
Target words	12 words × 3 sets	36 words at random	
Pre-test 1 week before	1. Discrete-item word production test 2. Same or Different task test		
	▼		
Lessons	Lesson 1 Lesson 2 Lesson 3	Lesson 1 Lesson 2 Lesson 3	
	▼		
Post-test 1 week later	Same as pre-test		
	▼		
Delayed test 4 weeks later	Same as pre-test		

lesson once a week for three weeks, and the control group did not receive any instruction but took three tests. Each lesson took about 30 minutes. All lessons and tests were conducted by the lead author who had approximately one year of sign language learning experience. Two students who use sign language as a second language cooperated in selecting the target words and evaluating the tests. The flow of the experiment is shown in Table 2.

(2) Teaching: Based on Shintani (2016), instructional activities were created. Three activities were designed for the PPP group in this study. Activity 1 was "watch and repeat" including presentation and practice. The teacher introduced 12 target words and showed the sign language to the students. The teacher taught them how to move their hands and explained the origin of the sign in Japanese at the same time. The students then imitated the teacher's signs. Activities 2 and 3 concerned the production part. Activity 2 was a "card passing game." Each student spread picture cards of 12 target words on the table by themselves. The teacher asked one student to choose one of the 12 target words by showing the sign. The other students then signed the target word and took a card of it as quickly as possible. Activity 3 was a "bingo game." The students chose nine cards from the 12 picture cards and arranged them freely in a 3×3 format. One student was asked to sign each of the target words in any order the student wanted. The other students expressed the sign of the word and turned the card

over if they had it in their selection. The student who got three words in a line before the others was the winner. At the beginning of each lesson, the goal and the task procedures were explained to the participants in Japanese. However, instruction was mostly conducted in sign language.

For the TBLT group, four input-based tasks were conducted in sign language. Following Shintani (2013, 2016), comprehension tasks were selected because the participants were absolute beginners of sign language. The tasks were designed in such a way that the outcome could only be achieved if students were successful in comprehending the input. Each task involved the students watching the teacher's commands (e.g., "Please show a rabbit" in sign language) and responding to them (e.g., by choosing the correct card and showing it to the teacher). At the beginning of each lesson, the goal and the task procedures were explained to the participants in Japanese. However, instruction was mostly conducted in sign language. Activity 1 was a "gesture game." The teacher asked the students to gesture (e.g., "Stand up, please," or "Let's play soccer" in sign language), and the students then made that gesture according to the teacher's direction. Activity 2 was a "card passing game." The teacher produced a sentence that included at least one noun and one verb, or two nouns. Students showed two cards of the words that were included in the teacher's direction as quickly as possible. Activity 3 was a "bring me a pen game." The teacher asked the students to bring something colored, a big or small book, a heavy or light book, or a long or short pencil. (For example, if the teacher ordered "Please bring a long pencil" in sign language, the students then showed a long pencil.) The winner was the student who showed these items the fastest. Activity 4 was a "bingo game." The students chose 9 cards from 10 or 11 cards and arranged them freely in a 3×3 format. The teacher showed the signs one by one. The students presented the card if they had it in their

bingo set. The first student to reach bingo was the winner.

(3) Testing: The study included the following two production tests to measure aspects of productive knowledge of the target signs. The content and methods of the tests were based on Shintani (2013). In the discrete-item word production test, the students were asked to look at pictures representing the target words on 36 flashcards and produce one word at a time in sign language. The teacher elicited the student's production by asking questions such as "What is this?" in sign language. If the students did not understand the question, the teacher used Japanese. For the cards for dimensional adjectives, the teacher asked the student to complete a sentence, for example, "This book is . . ." in Japanese. In Same or Different task test, a student communicated with the teacher in sign language to judge whether his or her pictures were the same or different. For this task, the student and the teacher respectively had a sheet of 27 pictures. Each picture consisted of a noun-adjective or a noun-verb combination of the target words. In the test, the teacher asked questions (e.g., "What animal is it?" and "What color is it?") to elicit the students' production of the target adjectives, nouns, and verbs in sign language. For both tests, the teacher worked in a one-on-one manner with the students. The maximum score of the two tests was 36 each; one point was given if the student could correctly sign the target word. The interaction was video recorded. The recording was used to verify that the students were producing the correct target words.

4. Analysis Method

The three lessons for the PPP and TBLT groups were video recorded and transcribed. The video data were used to identify individual learners' sign language expressions. The method of analysis followed Shintani (2013). Categorizing of the teacher's and students' production was also the same as in her study. Firstly, the transcribed data were analyzed in terms of the number of target words in

the input and output in the two experimental groups. A student's word was counted as one when more than one student chorally expressed the target signs for output. For this study, the input to a student came from the teacher or from another student. The sources of input were not differentiated in this study. Thus, when a student expressed a sign as an input to other students, it was counted as the input of a target word. Secondly, the transcripts were analyzed to determine how many of the input target words were isolated or embedded. If a word appeared by itself, it was counted as isolated; if it appeared with other words, it was counted as embedded. Thirdly, the transcripts were analyzed to examine how many of the student's target words were requested or optional; an optional one was furthermore counted as self-initiated if the student spontaneously expressed it and as borrowed if they imitated the teacher or other students. The test scores were analyzed for descriptive statistics and a series of parametric tests were applied for nouns, verbs, and adjectives, and for within-group and between-group comparison using js-STAR. Chi-square tests were used for the analysis of categorical data.

5. Research Ethics

For the recruitment, the researcher explained the purpose of the study to the participants. The control group consisted of the participants who were not able to attend the three learning sessions but agreed to take tests. The researcher obtained consent from the participants regarding the following ethical considerations on Google Forms: (1) participation is not compulsory; (2) participation can be cancelled after consent; (3) there will be no disadvantages due to nonparticipation or withdrawal from participation in the middle of the program; (4) privacy is protected and data strictly managed by the principal investigator; and (5) data are not used for purposes other than data.

III. RESULTS

1. Process Features

(1) Frequency of input and output: To examine whether there is a difference in the amount of input and output between PPP and TBLT teaching methods, the total amount of input and output was counted in the PPP and TBLT groups (Table 3). Input was the teacher's representation of the target word and output was the learner's representation of the target word. For input, the PPP group represented nouns 134 times, verbs 159 times, and adjectives 223 times, while the TBLT group represented nouns 397 times, verbs 285 times, and adjectives 343 times. For output, the PPP group had 85 nouns, 109 verbs, and 154 adjectives, while the TBLT group had 9 nouns, 5 verbs, and 4 adjectives. Chi-square tests showed that there was a significant difference in both input (nouns: $\chi^2(1) = 130.26$, $p < .01$; verbs: $\chi^2(1) = 35.76$, $p < .01$; adjectives: $\chi^2(1) = 25.44$, $p < .01$) and output (nouns: $\chi^2(1) = 61.45$, $p < .01$; verbs: $\chi^2(1) = 94.88$, $p < .01$; adjectives: $\chi^2(1) = 142.41$, $p < .01$) for nouns, verbs, and adjectives between the two groups. That is, the TBLT group was exposed to more frequent target words in input, but the PPP group had a higher amount of target words in output.

(2) Types of input: The frequency of target words by the teacher and students was classified as embedded or isolated to investigate whether there are differences in the way teachers represent target words in the two teaching methods. If the teacher signed a word with other words, it was embedded, and if the teacher produced a word by itself, it was isolated. Table 4 shows embedded and isolated word production by the teacher. In the PPP group, no parts of speech were embedded and were independent 134 times for nouns, 159 times for verbs, and 223 times for adjectives; in the TBLT group, they were embedded 355 times for nouns, 176 times for verbs, and 303 times for adjectives and were independent 42 times for nouns, 109 times for verbs, and 40 times for adjectives. Chi-square tests showed that there

Table 3　Frequency of Input and Output

	Input			Output		
	PPP	TBLT	Difference	PPP	TBLT	Difference
Nouns	134	397	sig.*	85	9	sig.*
Verbs	159	285	sig.*	109	5	sig.*
Adjectives	223	343	sig.*	154	4	sig.*

Note. sig. = significant *significant at p = .01

Table 4　Teacher's Embedded and Isolated Word Production

	PPP			TBLT		
	Nouns	Verbs	Adjectives	Nouns	Verbs	Adjectives
Embedded	0	0	0	355	176	303
Isolated	134	159	223	42	109	40

Table 5　Students' Requested/Optional and Self-Initiated/Borrowed Word Production

		PPP			TBLT		
		Nouns	Verbs	Adjectives	Nouns	Verbs	Adjectives
Requested		82	103	147	0	0	0
Optional	Self-initiated	0	5	6	0	0	0
	Borrowed	3	1	1	9	5	4

was a significant difference in embedded (nouns: χ^2 (1) = 355.00, p < .01; verbs: χ^2 (1) = 176.00, p < .01; adjectives: χ^2(1) = 303.00, p < .01) and isolated target words (nouns: χ^2 (1) = 48.09, p < .01; verbs: χ^2 (1) = 9.33, p < .01; adjectives: χ^2 (1) = 127.34, p < .01) for nouns, verbs, and adjectives between the two groups. In other words, the TBLT group had embedded input; in contrast, the PPP group had more isolated input for target words.

(3) Types of output: To consider whether there are differences in the way learners express target words in the two teaching methods, learners' production was classified as requested or optional; in the case of optional production, it was further classified into self-initiated or borrowed. Table 5 shows the number of requested or optional (self-initiated or borrowed) production by the students. The PPP group had 82 nouns, 103 verbs, and 147 adjectives that were requested; 5 verbs, 6 adjectives, but no nouns that were self-initiated; and 3 nouns, 1 verb, and 1

adjective that were borrowed. The TBLT group had 9 nouns, 5 verbs, and 4 adjectives that were borrowed, but no requested or self-initiated expressions. Chi-square tests showed that there was a significant difference in the students' requested production between the PPP and the TBLT groups for nouns, verbs, and adjectives (nouns: χ^2 (1) = 82.00, p < .01; verbs: χ^2 (1) = 103.00, p < .01; adjectives: χ^2 (1) = 147.00, p < .01). Chi-square tests showed that there was not a significant difference in students' optional self-initiated production for nouns, but significant differences for verbs and adjectives (nouns: χ^2 (1) = 2.67, p = .10; verbs: χ^2 (1) = 5.00, p < .05; adjectives: χ^2 (1) = 6.00, p < .05) Students' optional borrowed production was a significant tendency for nouns, but a nonsignificant difference for verbs and adjectives (nouns: χ^2 (1) = 3.00, p = .08; verbs: χ^2 (1) = 2.67, p = .10; adjectives: χ^2 (1) = 1.80, p = .18).

2. Productive Vocabulary Acquisition

A discrete-item word production test and the Same

or Different task test was conducted to measure productive knowledge of the target vocabulary. A pre-test, post-test, and delayed test were conducted to examine whether the PPP, the TBLT, and the control groups differed in their acquisition of the target words for three parts of speech. The test scores were analyzed with two-way ANOVA. The groups were at three levels (the PPP, the TBLT, and the control groups), and the tests were at three levels (pre-test, post-test, and delayed test).

(1) Nouns: The descriptive statistics of the test scores for the nouns are shown in Table 6. The results of the two tests are shown in Figures 1 and 2. The maximum score for each test was 9. The results of the individual item tests showed that the mean of the pre-test was 0.67 ($SD = 0.60$) for the PPP group, 0.67 ($SD = 0.60$) for the TBLT group, and 0.73 ($SD = 0.68$) for the control group; the mean of the post-test was 7.13 ($SD = 1.59$) for the PPP group, 5.53 ($SD = 1.89$) for the TBLT group, and 1. 20 ($SD = 0.91$) for the control group; and the mean of the delayed test was 7.53 ($SD = 1.02$) for the PPP group, 6.13 ($SD = 1.41$) for the TBLT group, and 2.07 ($SD = 1.48$) for the control group.

For the discrete-item word production test, two-way ANOVA (comparing within-group variance with between-group variance) was performed. The result showed the main effect of groups ($F_{(2,42)} = 71.30$, $p < .01$), the main effect of tests ($F_{(2,84)} = 234. 61$, $p < .01$), and the interaction effect of group and test difference ($F_{(4,84)} = 38.42$, $p < .01$). Multiple comparisons with Holms correction were administered. For the between-group comparisons, in the post-test and the delayed test, the mean values were significantly higher in the PPP and the TBLT groups than in the control group, and significantly

Table 6　Descriptive Statistics for Nouns

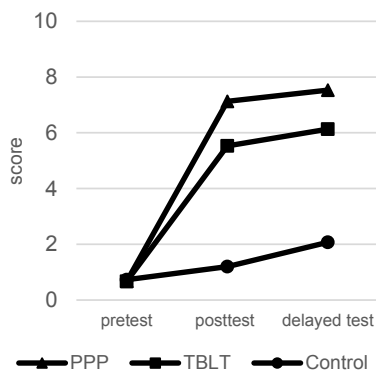| Test	Group	Pre-test		Post-test		Delayed test	
		M	*SD*	*M*	*SD*	*M*	*SD*
Discrete-item	PPP (n = 15)	0.67	0.60	7.13	1.59	7.53	1.02
word production	TBLT (n = 15)	0.67	0.60	5.53	1.89	6.13	1.41
	Control (n = 15)	0.73	0.68	1.20	0.91	2.07	1.48
Same or	PPP (n = 15)	0.60	0.71	7.13	1.59	7.73	0.85
Different task	TBLT (n = 15)	0.60	0.61	5.67	2.09	6.00	1.71
	Control (n = 15)	0.67	0.70	1.27	0.85	2.13	1.82

Fig. 1　Results of the Discrete-item Word Production Test of the Three Groups
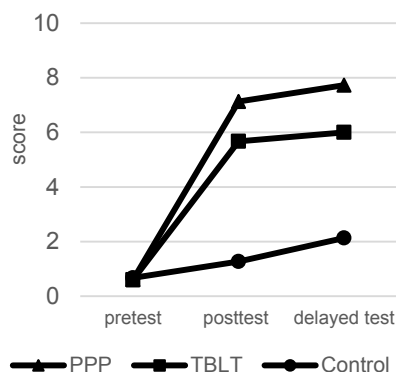

Fig. 2　Results of the Same or Different Task Test of the Three Groups

higher in the PPP group than in the TBLT group. For the within-group comparisons, in the PPP and the TBLT groups, the mean values of the post-test and the delayed test were significantly higher than the pre-test, and there was no significant difference between the post-test and the delayed test, so participants maintained the vocabulary. In the control group, the score was much lower than that of the experimental groups, but it improved significantly from the pre-test to the delayed test.

For the Same or Different task test, the mean of the pre-test was 0.60 ($SD = 0.71$) for the PPP group, 0.60 ($SD = 0.61$) for the TBLT group, and 0.67 ($SD = 0.70$) for the control group; the mean of the post-test was 7.13 ($SD = 1.59$) for the PPP group, 5.67 ($SD = 2.09$) for the TBLT group, and 1.27 ($SD = 0.85$) for the control group; and the mean of the delayed test was 7.73 ($SD = 0.85$) for the PPP group, 6.00 ($SD = 1.71$) for the TBLT group, and 2.13 ($SD = 1.82$) for the control group. Two-way ANOVA was performed. The result showed the main effect of groups (F (2,42) = 60.91, $p < .01$), the main effect of tests (F (2,84) = 203.05, $p < .01$), and the interaction effect of group and test difference (F (4,84) = 31.59, $p < .01$). Multiple comparisons with Holms correction were administered. For the between-group comparisons, in the post-test and the delayed test, the PPP and the TBLT groups were significantly higher than the control group, and the PPP group was significantly higher than the TBLT group. For the within-group comparisons, in the PPP and the TBLT groups, the mean values of the post-test and the delayed test were significantly higher than the pre-test, and there was no significant difference between the post-test and the delayed test; therefore, the participants maintained the vocabulary. In the control group, the score was much lower than that of the experimental groups, but it improved significantly from the pre-test to the delayed test.

(2) Verbs: The descriptive statistics for the test scores of each group at each level are shown in Table 7. The results of the two tests are shown in Figures 3 and 4. The maximum score for each test was 12. For the discrete-item word test, the mean of the pre-test was 0.27 ($SD = 0.44$) for the PPP group, 0.60 ($SD = 1.08$) for the TBLT group, and 0.47 ($SD = 0.88$) for the control group; the mean of the post-test was 9.13 ($SD = 1.82$) for the PPP group, 5.93 ($SD = 2.11$) for the TBLT group, and 1.13 ($SD = 1.02$) for the control group; and the mean of the delayed test was 9.60 ($SD = 1.40$) for the PPP group, 6.47 ($SD = 2.19$) for the TBLT group, and 1.87 ($SD = 1.36$) for the control group. Two-way ANOVA was performed. The result showed the main effect of groups (F (2,42) = 73.25, $p < .01$), the main effect of tests (F (2,84) = 296.81, $p < .01$), and the interaction effect of group and test difference (F (4,84) = 58.38, $p < .01$). Multiple comparisons with Holms correction were administered. For the between-group comparisons, in the post-test and the delayed test, the mean values were significantly higher in the PPP and the TBLT groups than in the control group, and were significantly higher in the PPP group than in the TBLT group. For the within-group comparisons, in the PPP and the TBLT groups, the mean values of the post-test and delayed test were significantly higher than the pre-test, and there was no significant difference between the post-test and delayed test, so participants maintained the vocabulary. In the control group, the score was much lower than that of the experimental groups, but it improved significantly from the pre-test to the delayed test. For the Same or Different task test, the mean of the pre-test was 0.67 ($SD = 0.47$) for the PPP group, 0.53 ($SD = 0.62$) for the TBLT group, and 0.73 ($SD = 0.93$) for the control group; the mean of the post-test was 9.47 ($SD = 1.86$) for the PPP group, 6.33 ($SD = 1.58$) for the TBLT group, and 0.93 ($SD = 1. 00$) for the control group; and the mean of the delayed test was 9.93 ($SD = 1.44$) for the PPP group, 6.53 ($SD = 1.86$) for the TBLT group, and 1.93 ($SD = 1.48$) for the control group. Two-way ANOVA was performed. The result showed

Table 7 Descriptive Statistics for Verbs

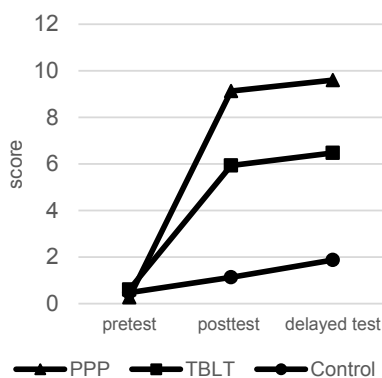| Test	Group	Pre-test		Post-test		Delayed test	
		M	*SD*	*M*	*SD*	*M*	*SD*
Discrete-item	PPP (n = 15)	0.27	0.44	9.13	1.82	9.60	1.40
word production	TBLT (n = 15)	0.60	1.08	5.93	2.11	6.47	2.19
	Control (n = 15)	0.47	0.88	1.13	1.02	1.87	1.36
Same or	PPP (n = 15)	0.67	0.47	9.47	1.86	9.93	1.44
Different task	TBLT (n = 15)	0.53	0.62	6.33	1.58	6.53	1.86
	Control (n = 15)	0.73	0.93	0.93	1.00	1.93	1.48

Fig. 3 Results of the Discrete-item Word Production Test of the Three Groups
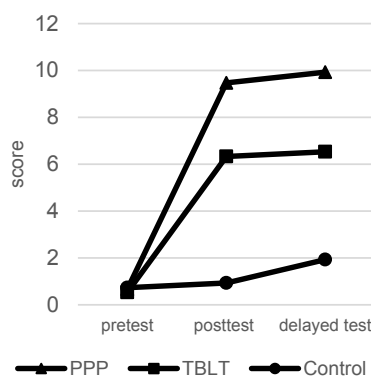

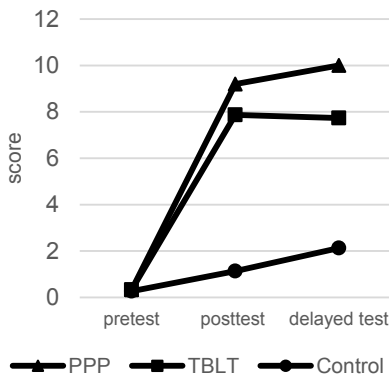Fig. 4 Results of the Same or Different Task Test of the Three Groups

the main effect of groups (F (2,42) = 95.60, $p < .01$), the main effect of tests (F (2,84) = 379.67, $p < .01$), and the interaction effect of group and test difference (F (4,84) = 82.45, $p < .01$). Multiple comparisons with Holms correction were administered. For the between-group comparisons, in the post-test and the delayed test, the mean values were significantly higher in the PPP and the TBLT groups than in the control group, and was significantly higher in the PPP group than in the TBLT group. For the within-group comparisons, in the PPP and the TBLT groups, the mean values of the post-test and delayed test were significantly higher than the pre-test, and there was no significant difference between the post-test and the delayed test, so the participants maintained the vocabulary. In the control group, the score was much lower than that of the experimental groups, but it improved significantly from the pre-test and the
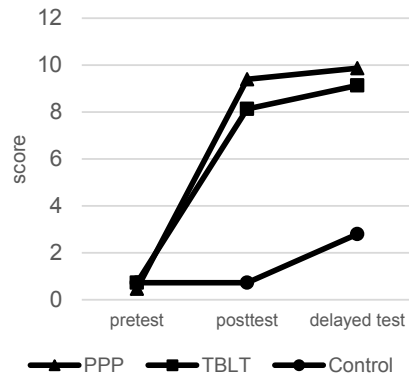
post-test to the delayed test.

(3) Adjectives: The descriptive statistics for test scores of each group at each level are shown in Table 8. The results of the tests are shown in Figures 5 and 6. The maximum test score for each test was 15. For the discrete-item test, the mean of the pre-test was 0.33 ($SD = 0.60$) for both the PPP and the TBLT groups and was 0.27 ($SD = 0.44$) for the control group; the mean of the post-test was 9.20 ($SD = 3.17$) for the PPP group, 7.87 ($SD = 2.70$) for the TBLT group, and 1.13 ($SD = 1.02$) for the control group; and the mean of the delayed test was 10.00 ($SD = 2.88$) for the PPP group, 7.73 ($SD = 3.34$) for the TBLT group, and 2.13 ($SD = 1.50$) for the control group. Two-way ANOVA was performed. The result showed the main effect of groups (F (2,42) = 47.71, $p < .01$), the main effect of tests (F (2,84) = 151.83, $p < .01$), and the interaction effect of group and test difference

Table 8 Descriptive Statistics for Adjectives

Test	Group	Pre-test		Post-test		Delayed test	
		M	SD	M	SD	M	SD
Discrete-item	PPP (n = 15)	0.33	0.60	9.20	3.17	10.00	2.88
word production	TBLT (n = 15)	0.33	0.60	7.87	2.70	7.73	3.34
	Control (n = 15)	0.27	0.44	1.13	1.02	2.13	1.50
Same or	PPP (n = 15)	0.47	0.72	9.40	3.79	9.87	2.96
Different task	TBLT (n = 15)	0.73	0.85	8.13	1.96	9.13	2.39
	Control (n = 15)	0.73	0.85	0.73	0.85	2.80	1.68

Fig. 5 Results of the Discrete-item Word Production Test of the Three Groups

Fig. 6 Results of the Same or Different Task Test of the Three Groups

(F (4,84) = 24.06, p < .01). Multiple comparisons with Holms correction were administered. For the between-group comparisons, in the post-test and the delayed test, the mean values were significantly higher in the PPP and the TBLT groups than in the control group. There was no significant difference between the PPP and the TBLT groups in the post-test, but, in the delayed test, the scores were significantly higher for the PPP group than for the TBLT group. For the within-group comparisons, in the PPP and the TBLT groups, the mean values of the post-test and the delayed test were significantly higher than the pre-test, and there was no significant difference between the post-test and the delayed test, so participants maintained the vocabulary. In the control group, the score was much lower than that of the experimental groups, but it improved significantly from the pre-test to the delayed test. For

the Same or Different task test, the mean of the pre-test was 0.47 (SD = 0.72) for the PPP group and 0.73 (SD = 0.85) for the TBLT and the control groups; the mean of the post-test was 9.40 (SD = 3.79) for the PPP group, 8.13 (SD = 1.96) for the TBLT group, and 0.73 (SD = 0.85) for the control group; and the mean of the delayed test was 9.87 (SD = 2.96) for the PPP group, 9.13 (SD = 2.39) for the TBLT group, and 2.80 (SD = 1.68) for the control group. Two-way ANOVA was performed. The result showed the main effect of groups (F (2,42) = 45.61, p < .01), the main effect of tests (F (2,84) = 194.05, p < .01), and the interaction effect of group and test difference (F (4,84) = 33.95, p < .01). Multiple comparisons with Holms correction were administered. For the between-group comparisons, in the post-test and the delayed test, the PPP and the TBLT groups were significantly higher than the control group. There was no significant

Rika ENOKIDO, Atsuko SATO and Hideki SAKAI

difference between the PPP and TBLT groups in either the post-test or the delayed test. For within-group comparisons, in the PPP and the TBLT groups, the mean values of the post-test and the delayed test were significantly higher than the pre-test, and there was no significant difference between the post-test and the delayed test, so participants maintained the vocabulary. In the control group, the score was much lower than that of experimental groups, but it improved from the pre-test and the post-test to the delayed test.

IV. DISCUSSION

1. Discussion of the Process Features

Research Question 1 concerned whether the characteristics of PPP and TBLT approaches are reflected in the process features. Following Shintani (2013, 2016), significant differences in process features are shown in Table 9. In the PPP approach, the target forms were instructed directly; therefore, intentional learning and production were promoted. The PPP group was exposed to isolated target words, and learners frequently produced the target items. Although most of the representations were requested, some learners spontaneously produced signs that were similar so that they could check the difference in form when they learned another word. In the TBLT group, learners were frequently exposed

to embedded target items. It could be that the amount of teacher's production increased because the target adjectives were frequently used to explain nouns. As a matter of technical method, it could also be that the timing of watching the teacher online by students was different, so the teacher expressed the target signs repeatedly. Optional borrowed production emerged, but optional self-initiated production did not appear. This result differs from Shintani (2013). The reason may be due to the iconicity (Takei, 2006) that is a characteristic of sign language. It is possible that the students understood what the signs represented roughly even if they did not observe how to move their hands in detail. Additionally, even if the learners did not understand the adjectives, they could perform tasks if they knew the nouns, so they did not need to ask them of the teacher. Thus, the students did not show optional self-initiated production for task achievement. For these reasons, interaction may be increased by choosing higher abstract sign language vocabulary as target items. Another possible reason is that students were not at a level where they could ask questions in sign language in three lessons. A student asked the teacher, "What is the difference between 'red' and 'purple'?" in Japanese in the second lesson. Also, a student who mistook "gray" for "bird" asked, "Which is gray, and which is a bird?" after the third class. Learners may become more aware of repeated

Table 9 Comparison of the Process Features between the PPP and TBLT Groups

Process Features	PPP	TBLT
Input		
Frequency of target words	Frequent input of the target words	Frequent input of the target words
Embedded vs. isolated	Target words occurred in isolation.	Target words were more often embedded than isolated.
Output		
Production frequency of target words	Target words were produced frequently.	Target words were produced less frequently.
Optional vs. requested	Target words were produced mostly when requested.	Target words were produced only optionally.
Self-initiated vs. borrowed	Verbs and adjectives were produced through self-initiation.	Target words occurred by borrowing the teacher's production.

adjectives. Assuming that we are moving from watching to interaction, learners may need a longer period for this transition.

Conversely, the PPP group showed self-initiated production for verbs and adjectives. Some target verbs and adjectives were similar (e.g., "red": close your mouth and trace your lips to the right with the tip of your right index finger and "white": open your mouth and move the tip of your right index finger pointing to your teeth to the left). Consequently, some students may output what they have already learned while learning other signs to absorb knowledge about similar shapes or movements of their hands.

2. Discussion of Productive Vocabulary Acquisition

Research Question 2 concerned whether PPP affects the production of nouns, verbs, and adjectives in sign language for university and graduate students who are beginning to learn sign language as a second language. The PPP approach appears effective for nouns, verbs, and adjectives because the score improved much more than in the control group, and the score was maintained from the post-test to the delayed test in both tests. This is because students comprehended the origin of the signs and details in the PPP approach. The PPP group had opportunities to understand the target signs explicitly; by contrast, the TBLT group noticed it through the teacher's production implicitly.

Research Question 3 concerned whether the TBLT approach affects the production of nouns, verbs, and adjectives in sign language for university and graduate students who are beginning to learn sign language as a second language. The TBLT approach is effective because the score improved much more than in the control group, the same as PPP approach, and the score was maintained from the post-test to the delayed test in both tests for nouns, verbs, and adjectives. However, the score of the TBLT group was lower than that of the PPP group for nouns and verbs; however, for adjectives, the TBLT group performed as well as the PPP group

in the post-test of the discrete-item test and the task-based test. The result was different from Shintani (2013) who concluded that the TBLT approach is more effective for adjectives. As shown in Shintani (2013), negotiation for meaning and learner-initiated productions occurred using the TBLT approach, but, in this intervention, only little negotiation for meaning and learner-initiated production took place. This difference in process may have led to the difference in results. As other reasons, for some target words, the relationship between expression and meaning is difficult to grasp implicitly, which may cause difficulty in noticing or establishing the connection even if it was observed in the lessons. In addition, the TBLT group played gesture games as a task whereas they were not requested to produce the target words. Thus, some students might easily be confused about signs and gestures. Moreover, there are more subtle differences in expression of sign language vocabulary words than the students had expected, and it might be difficult to notice the difference implicitly. Furthermore, although the same conditions were used in both teaching methods, the signs represented in three dimensions were performed on a flat online surface. In the PPP approach, sign words were explicitly taught along with verbal explanations, whereas in the TBLT approach the students saw the teacher's signs and understood them implicitly, which may have made it difficult to accurately convey the sign language representations.

Research Question 4 concerned whether there are any differences between the PPP and the TBLT approaches in their effect on the production of nouns, verbs, and adjectives in sign language for university and graduate students who are beginning to learn sign language as a second language. Summarizing the results of RQ2 and 3, both the TBLT and PPP approaches are effective for sign language acquisition because both experimental groups tremendously improved in both tests for nouns, verbs, and

adjectives; further, both experimental groups kept the score from the post-test to the delayed test in most tests. The PPP and TBLT approaches showed a similar effect for adjectives. However, the PPP approach was more effective for nouns and verbs in this study because the PPP group demonstrated significantly higher scores than the TBLT group for nouns and verbs. This suggests that the explicit approach and production activities are more efficient than the implicit approach in communicative activities for the acquisition of sign language vocabulary for adults under this condition. However, as mentioned in the discussion about RQ1, sign languages have iconicity as one of their characteristics. Highly iconic signs were chosen as the target items. Therefore, the PPP group could comprehend the details of the signs; however, the TBLT group understood the meaning roughly and did not interact with details of the sign word expressions. These processes may be reflected in the results of acquisition.

Although there was no research question connected to the control group, this study found that most control group tests showed improvement between the pre-tests and the delayed tests. It may be easier to remember in a visual language if participants learn a few words during interaction in a task-based test. Some signs may be easy to understand by watching due to their iconicity.

V. CONCLUSION

This study compared the effect of PPP and TBLT approaches on the acquisition of sign language vocabulary for nouns, verbs, and adjectives by investigating the process features and outcomes. The results showed that both the PPP and the TBLT approaches were effective for sign language vocabulary acquisition; however, the PPP approach was more effective for nouns and verbs, but that both approaches produced the same effect for adjectives under this condition. The results of this study were different from Shintani (2013) in that the

PPP approach was more effective than TBLT in the acquisition of nouns and verbs, while both PPP and TBLT approaches were effective in the acquisition of adjectives.

As pedagogical implications of this study, to learn sign language, which has iconicity, comprehension tasks used in this study may lead to a rough understanding of signs, and no interaction occurred in short lessons. It may need time for the students to become aware of the details of the sign language. On the other hand, explicit instruction may help adult learners to become aware of the details and learn the signs accurately.

There are three limitations of this study. First, the target words were highly iconic, and the learners understood roughly, so no interaction on the details of sign language representation occurred in the TBLT group. Further study needs to be conducted on sign language vocabulary with more abstract items for negotiation for meaning to occur in interaction. Second, only three classes were conducted for learning sign language vocabulary. In further studies, prolonging the intervention period may lead to interaction between the teacher and students and correction of the details of the target signs. Third, it is difficult to refer to sign language grammar and sign language communication because the present study targeted only nouns, adjectives, and verbs, and there are aspects of syntactic and pragmatic aspects that have not yet been examined.

Lastly, this study cannot be generalized because the study was limited in that it involved teaching sign language to normal-hearing people, people who had already acquired Japanese, and university students. However, it is hoped that this study will be of help in studies of sign language acquisition as a second language.

References

Ellis, R., & Shintani, N. (2014) *Exploring language pedagogy through second language. acquisition*

research. Routledge, London.

Long, M. H., & Robinson, P. (1998) Focus on form Theory, research, practice. In C. Doughty, & J. Williams (Eds.), *Focus on form in classroom second language acquisition.* Cambridge University Press, New York, 15-41.

Matsumura, M. (2012) *Task-Based English Class Design.* Taishukan shoten. (in Japanese)

Matsumura, M. (2017) Task-based concepts and language teaching methodology. In Matsumura, M. (Ed.), *Task-Based Instruction of English as a Second Language: Principles of TBLT.* Taishukan shoten, 5-36. (in Japanese)

Nakano, S. (2021) A Review of Studies Related to the Sign Language Pedagogy as a Second Language. *Research in Educational Practice and Development, Gunma University*, 38, 255-265. (in Japanese)

"Our Sign Language" reorganization committee (Eds.) Japan National Center of Sign Language Education (Sign Language Supervision) (2021) *Our Sign Language Dictionary I*, Revised 6th Edition, Japanese Federation of the Deaf. (in Japanese)

Shintani, N. (2013) The effect of focus on form and focus on forms approach on the acquisition of productive knowledge of L2 vocabulary by young beginning-level learners. *TESOL Quarterly, 47*, 36-62.

Shintani, N. (2016) *Input-based tasks in foreign language approach for young learners.* John Benjamins, Amsterdam.

Takei, W. (2006) Does Iconicity of Language Facilitate Language Acquisition? : Perspectives from Research on Sign Language Acquisition. *The Japanese Journal of Communication Disorders*, 23, 143-151. (in Japanese)

手話のL2語彙習得におけるPresentation-Practice-Production と
Task-based Language Teaching の影響に関する実験的検討

榎戸　里佳*・左藤　敦子**・酒井　英樹***

　　本研究では、日本語母語話者で手話初心者の大学生45名を対象に手話の語彙習得について実験的に検討した。第二言語習得で有効とされるpresentation-practice-production（PPP）と task-based language teaching（TBLT）の教授法が手話単語の習得に与える影響を明らかにすることを目的とした。Shintani（2013）を基にして、参加者を2つの実験（PPP、TBLT）群と対照群に分け、実験群には30分の授業を3回行った。指導法の効果を測定するため、タスクベースと個別項目のテストを事前テスト、事後テスト、遅延テストで実施した。その結果、PPPとTBLTの教授法はどちらも手話単語の習得に効果的であり、PPPは名詞と動詞に対してより効果的であることを示した。この結果は、TBLTでの理解型タスクでは、学習者は手話の写像性により手話を大まかに理解する傾向があり、学習者が手話の細部に気づくのに時間がかかる可能性があることを示唆している。一方、PPPに使用された明示的な指導は、成人学習者が手話の詳細な表現に気づき、正確に学習するのに役立つ可能性がある。

キー・ワード：手話　第二言語習得　語彙　presentation-practice-production　task-based language teaching

　　* 筑波大学大学院人間総合科学学術院
　 ** 筑波大学人間系
*** 信州大学学術研究院教育学系