

**Short-Term Study Abroad:
Does It Improve University EFL Students' English?**

**Miki TOKUNAGA
Sachiyo HAYASHI**

東アジア英語教育研究会「研究論集」第8号 抜刷
2022年3月発行

Short-Term Study Abroad: Does It Improve University EFL Students' English?

Miki TOKUNAGA (National Institute of Technology Kurume College)

Sachiyo HAYASHI (Kumamoto Gakuen University)

Abstract

Short term study abroad programs of less than one month are popular among Japanese university students. While participants hope to come back with better English skills, can short-term programs make a difference? This research aimed to measure which aspects of learners' English proficiency improved when 19 Japanese university students with novice to intermediate proficiency studied abroad for three to four weeks. The test battery used included a grammaticality judgment test and a sentence translation test to measure the participants' grammatical knowledge and controlled production, and a picture description test and SST interview to measure their spontaneous production. The analysis showed no significant improvements in grammatical knowledge and accuracy in controlled production, while slight improvements were observed in spoken fluency.

1. Introduction

According to Japan Student Services Organization (2021), the number of Japanese university students studying abroad steadily increased until reaching 115,146 in 2018 but decreased to 107,346 in 2019 due to COVID-19. Of these participants, the majority (76,545 in 2018 and 71,267 in 2019 in the previous data) studied abroad for less than a month. Short-term study abroad programs are attractive to Japanese university students because they can participate during their long summer and spring breaks. Also, unless the students are receiving some kind of competitive financial support from their universities or other organizations, most short-term study abroad programs do not set minimum language proficiency requirements, so anyone who wants to and can afford to can participate.

There have been numerous studies on the effects of short-term study abroad, but many report on its effects on learners' confidence, motivation, and cultural awareness, based on self-reporting questionnaires and interviews, and not on the proficiency of the target language (e.g. Isa, 2016; Kimura & Shimizu, 2016; Matsuda, 2012; Nishina et al., 2017; Sasaki, 2017). Fujioka and Agawa (2007)'s self-reporting questionnaires of 61 Japanese university students after taking part in four-week study abroad programs showed increase in motivation towards studying English. Geis and Fukushima (1997) reported that teachers of 24 Japanese university students who studied abroad in the U.S. for six weeks felt that the most significant benefit was the increase in the students' level of motivation in the classroom. Upon returning from the program, those 24 students became role models in their English classes, as they were most at ease with talking in English. On the other hand, interviewing three Japanese university students after participating in a four-week study abroad program in the U.K,

Horness (2014) found that the participants' positive learning experience in the UK did not transfer back to studying in Japan. The participants were happy with their experiences, but their behaviors in English classes in Japan did not change after their return. Tateyama (2002) conducted surveys and interviews with 11 female junior college students who took part in a three-week English program in Australia. While the participants reported positive effects on cross-cultural attitudes and learning English, eight out of the 11 students disagreed or strongly disagreed that they could participate in discussions with native speakers of English. Five participants disagreed or strongly disagreed that they could express opinions in English.

Of course, the benefits of studying abroad are not just improving languages. Experiencing different cultures, meeting students from around the world at language schools, and living away from home are all invaluable experiences. However, many participants sign up hoping to improve their language proficiency. JASSO (2019) conducted a survey with 1,001 people in their 20s to 40s who had studied abroad within the last 15 years. When asked for the top 3 reasons for studying abroad, the top response with 58.5% was to learn the language. The number was even higher (61.7%) among the 384 respondents who studied abroad for less than 3 months. Also, Ryugaku Journal (2019), one of the most prominent private study abroad services organizations in Japan, reported that 87.2% of their 620 potential study abroad clients chose "acquiring the language" as one of the purposes of studying abroad, followed by "broadening my views" (56.1%), "learning communication skills" (40.7%), and "international exchange" (25.5%). (multiple answers were allowed).

Reviewing past studies on the effects of studying abroad on the target language proficiency, DeKeyser (2007) concluded that improvements in spoken fluency could be seen after studying abroad for a semester or longer, but not in accuracy. He also pointed out that in studies where student's progresses were measured by their self-reporting, students often overestimated the gains, not realizing how their perceived gains were limited to routine exchanges in the particular environments they were placed in. In Horness's (2018) study involving 88 Japanese university students, 76% of the participants in four-week study abroad programs believed, prior to going abroad, that the experience would increase their confidence in expressing their opinions in English, and 70% believed that it would increase their knowledge of English. However, upon returning, only 38% believed the experience increased their confidence in using English, and 47% believed that their English knowledge increased, showing that study abroad participants' expectation of improving their English was higher than the reality.

However, some studies carried out in Japan using objective measures, have reported improvements in students' English proficiency (e.g. Kimura 2009, 2011; Kobayashi 1999; Kuramashi 2019; Suzuki & Hayashi 2014). Otsuka and Negishi, (2009) analyzed students' spoken production in a five-minute group discussion activity before and after their two-week study abroad experience in New Zealand. The results showed statistically significant increase in their phonation time ratio (ratio of time spent speaking during the given time) and their articulation rate (number of syllables per minute, excluding pauses), suggesting that their spoken fluency has improved. Taura et al. (2009) found statistically significant improvements in TOEFL listening section scores among their 20 university students after

their three-week studying abroad experience in New Zealand. Nonaka and Seki (2016) compared pre and post TOEIC IP scores of their university students who had participated in three different types of study abroad programs in the United States between 2004 and 2014. During the 11 years, 87 students participated in their three-week program, 11 in the 3-month program, and 27 in the six-month program, and all the participants took TOEIC IP before and after their study abroad experiences. Comparing pre and post TOEIC IP scores, they found statistically significant increases in the total and listening section scores for all three groups, but only the six-month group showed statistically significant increase in the reading section scores. Focusing on the three-week group, they compared the results of the higher proficiency (pre TOEIC IP scores of above 450) group ($n = 20$) and the lower proficiency (pre TOEIC IP scores of below 320) group ($n = 20$), and found statistically significant improvements only in the lower proficiency group, with the average scores of the higher proficiency group actually dropping in the post test for both listening and the reading sections. The researchers explained that the lack of improvement in the higher-level group could be that, because their proficiency level was enough for the three-week study abroad experience, they did not have to push themselves as the lower proficiency group did, or that they did not do their best in the post test as their goal of studying abroad had been achieved and their motivation for the test had dropped.

2. Current Study

Even though combining results from the previous studies allows us to speculate that short-term study abroad experiences may increase participants listening skills and spoken fluency, none have conducted different types of tests to look at changes in different aspects of language proficiency within a group. This study conducted pre and post tests of four different kinds; a grammaticality judgment test, a Japanese to English translation test, a picture description test, and an interview test with a native English speaker, to investigate the following research questions.

1. Do short-term study-abroad experiences improve participants' English proficiency?
2. If so, what aspects of English proficiency are improved?
3. What factors seem to contribute to the improvement?

2.1 Participants

Participants of this study were 19 Japanese students of various majors who participated in short-term study-abroad programs organized by the private four-year-university they were attending. There were 15 females and four males, aged between 18 and 20, and all participants had Japanese as their L1. The researchers recruited the participants from their English classes and the university's English conversation lounge, and 20 students signed up for the research. One of the participants stayed overseas for 5 months, much longer than the original plan, so this participant's data was later removed. The 19 participants were asked to meet with the researchers to take a battery of tests before and after studying abroad. They were also asked to take a TOEIC IP, which was offered almost every month at the university, before they went abroad. The TOEIC IP scores were used to see the general English

proficiency of the participants. Participants were compensated with a 5,000-yen gift certificate and a refund on the TOEIC IP test fee. Participants' pre-departure TOEIC IP score ranged from 360 to 715, with an average score of 533. Five participants had never been abroad before, and four had previous study-abroad experiences of two weeks to 10 months. Ten of the participants reported they used no English outside of their English classes, while the others used English at the English conversation lounge, English conversation schools, or at part-time jobs.

The participants of this study went abroad on one of the following two programs organized by the university. The first program was one of the elective English classes for sophomores. For a total of four credit hours, instead of attending two 90-minute classes for two semesters, students who participated in this program attended one 90-minute class a week for two semesters, attended pre-study abroad workshops, and studied English abroad for three weeks either during the summer or winter breaks. At the time of this study, they were able to choose from seven destinations; Australia, Canada, Malta, New Zealand, Singapore, the U.K., and the U.S. The other program was a non-credit four-week program during the spring break at an affiliated university in Australia or in the U.S. The program's aim was to improve English proficiency and learn culture, history and nature of the country, and develop global awareness through interacting with the local residents and students.

Pre and post surveys were conducted to acquire information about the participants English study backgrounds and their environments (program attended, country, classes, accommodation, etc.) during their study-abroad experiences.

2.2 Test Batteries

All tests were conducted individually, in a student-teacher conference room at the university. Around three weeks prior to going abroad, the participants met with one of the researchers to take a survey, a grammaticality judgment test, a written Japanese to English sentence translation test, and a spoken picture description test. They also met with a native English speaker teacher, who is a certified SST interviewer/rater on a different day to take the SST interview. Two to three weeks upon returning, participants took the same round of tests, again on two separate days.

2.2.1 Grammaticality Judgment Test

In the grammaticality judgment test (GJT), participants were first asked to judge whether each sentence was grammatical, and then to correct errors for sentences they had judged as ungrammatical. GJTs have been widely used, with timed GJTs thought to measure constructs related to implicit knowledge and untimed GJTs to measure constructs of explicit or declarative knowledge (e.g., Ellis, 2009; Green & Hecht, 1992; Loewen, 2009; Roehr, 2008; Sakai, 2008; Shimada, 2010). Notwithstanding disputes about the distinctions between these constructs, many previous studies agree that timed and untimed tests measure different factors (e.g., Ellis, 2005; Godfroid, et al., 2015; Zhang, 2015). However, Shimada (2010) found that the difference in mean scores was larger between grammatical and ungrammatical items than between timed and untimed items among Japanese EFL

learners. Since the GJT was used to measure the participants' explicit or declarative knowledge in this study, only an untimed GJT was used.

The test consisted of 25 English sentences covering 17 grammar structures based on and revised from Ellis (2005), Shimada (2010) and Tokunaga (2017); 3rd person *-s*, plural *-s*, verb complement, *go + ing*, *It* subject, *since/for*, WH question, Yes/No question, past tense, relative clause, embedded question, present progressive, conjunction, preposition, negation, tag question, and article. Among the 25 sentences, 21 were ungrammatical, requiring correction. Each sentence was accompanied by a Japanese translation to minimize the effect of vocabulary knowledge and reading ability, and to focus on their understanding of target grammar structures. An example item (testing *go + ing*) is shown below:

Did you go to shopping in Tenjin last week?
あなたは先週天神に買い物に行きましたか。

For this item, correct responses would be choosing “ungrammatical” in the judging section, then deleting *to* in the correction section. There was no time limit for this test, and the same test was used for pre and post study abroad experiences. The test was scored for grammaticality judgment of 25 sentences and correction of 21 ungrammatical sentences.

2.2.2 Written Japanese to English Sentence Translation Test

The written Japanese to English sentence translation test was conducted to observe accuracy of participants' written production. The participants were asked to translate 13 short Japanese sentences into English with no time limit. The 13 sentences were checked for 40 target items covering the same 17 grammar structures as the GJT. Different sentences were used for before and after the study abroad experiences, but the sentences were very similar, covering the same number of target items. One set of pre and post test sentences including the Yes/No question and plural *-s* items is shown below with expected English sentences in brackets:

Pre test: あなたのお母さんは動物が好きですか。(Does your mother like animals?)

Post test: Matt は野菜が好きですか。(Does Matt like vegetables?)

The number of target grammar structures varied among sentences, and each target structure was independently scored dichotomously. For example, writing “Does your mother like animal?” for the first example above would be scored one for the Yes/No question, but zero for the plural *-s*. Thus, the test had 40 possible points.

2.2.3 Picture Description Tests

The picture description test was conducted to observe participants' fluency in spoken production. This test used a card (Fig. 1) showing six pictures connected with arrows in sequence, created for the purpose using free-to-use illustrations online. The storyline aimed to be simple but relatable to participants, and the same pictures were used for pre and post tests.

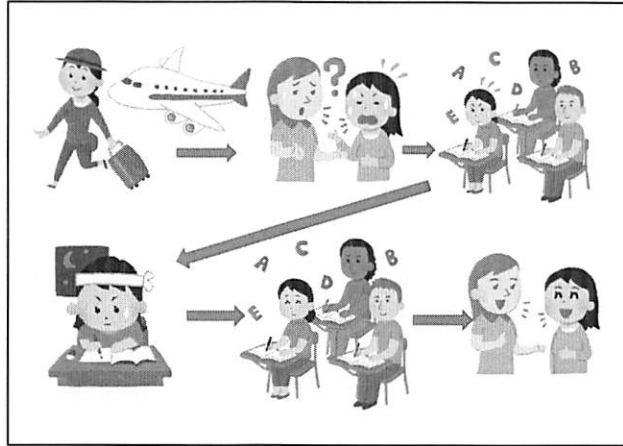


Fig.1 Picture card used for the spoken and written picture description tests

The participants were given one minute to look at the card and think about the story, then asked to describe what was happening in the pictures. Participants were asked to say “finish” when they were finished, but there was no time limit on how long they spoke. The test was audio recorded, with participants’ permission, then transcribed for analysis.

2.2.4 Standard Speaking Test (SST)

The Standard Speaking Test (SST) was conducted to measure the participants’ oral communication proficiency. SST is a face-to-face interview test developed from the ACTFL Oral Proficiency Interview (OPI) specifically for the Japanese market (ACTFL-ALC Press, 1996). Interviews comprise four rated stages: Warm-up, a short conversation to allow the interviewer and test-taker to develop familiarity; Picture description, where the test-taker is required to describe an illustration of an everyday scene such as a classroom or restaurant; Role-play, where the test-taker must complete a commercial or social transaction or negotiation, such as extending an invitation or purchasing a product; Story-board, where the test-taker must narrate a story describing the events shown in a four-panel or six-panel story-board. Once the test-taker has completed all stages, follow-up questions are asked concerning their personal experiences on topics related to the prompts. Interviewers must therefore tailor each interview to the test-taker's level and personalize it to focus on their personal interests and experiences.

While the original ACTFL is rated from 1 (novice) to 9 (advanced), SST does not include the highest rating of 9 because speakers at this level are exceedingly rare in Japan and SST is not intended to evaluate levels of proficiency beyond Intermediate-high (SST8). Each stage is rated 1 to 8 for text type, comprehension, communication, grammatical accuracy, pronunciation, and fluency. The text type rates the complexity of the utterances of the test-taker, whether they are single word responses, memorized chunks, simple or complex sentences. The comprehension rates how much of interviewer’s questions and task descriptions the test-taker seems comprehending, and the communication rates the test-taker’s

attempts to communicate with the interviewer, such as asking questions for clarification and trying to maintain the conversation.

Because no other trained ACTFL interviewers or raters were available, all interviews and ratings were conducted by the same certified SST interviewer/rater. Ratings were conducted at least one week after each set of interviews was finished, and each interview was rated multiple times, with a gap of at least one week between rating sessions.

3. Results

3.1 Grammaticality Judgment Test

In the GJT, participants were asked to judge the grammaticality of 25 sentences then correct the ones they judged ungrammatical. A paired sample *t*-test was performed to compare the 19 participants' scores before and after the study abroad experiences. Table 1 lists the mean scores of pre and post tests, and the results of the *t*-test. Even though the same sentences were used for pre and post tests, the mean score for judging decreased from 17.63 to 17.42, while the mean for correction increased from 12.58 to 12.95. However, these changes were not statistically significant, $t(18) = 0.29, p = .78$, and $t(18) = -.069, p = .50$, respectively.

Table 1 Results of grammaticality judgment tests

	Pre		Post		<i>T</i>	<i>df.</i>	<i>p</i>	<i>r</i>
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>				
Judging (<i>N</i> = 25)	17.63	3.72	17.42	4.41	0.29	18	.78	.07
Correction (<i>N</i> = 21)	12.58	4.10	12.95	4.50	-0.69	18	.50	.16

While the differences in pre and post test means were statistically insignificant, individual differences were observed as shown in Figure 2 and 3. Differences in pre and post test scores ranged from -5 to 7 in judging and -3 to 5 in correction. Participant 12, for example, judged seven more sentences correctly and corrected five more errors in the post test, while participant 19 correctly judged five sentences fewer and corrected two fewer errors.

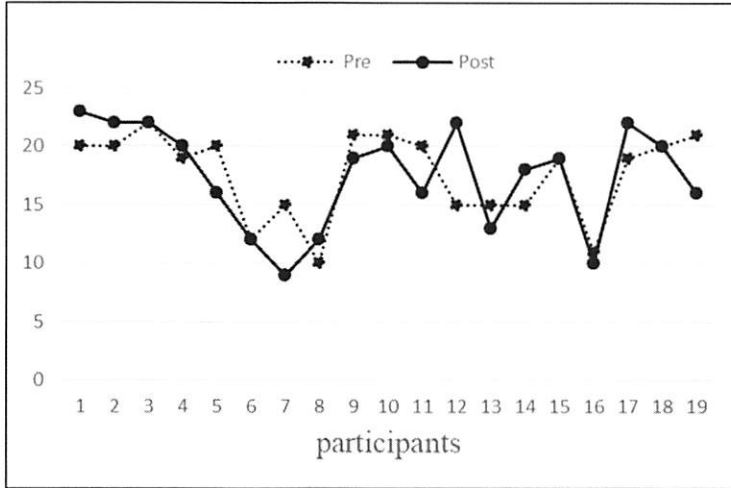


Fig. 2 Individual scores of GJT (Judging) in pre and post tests

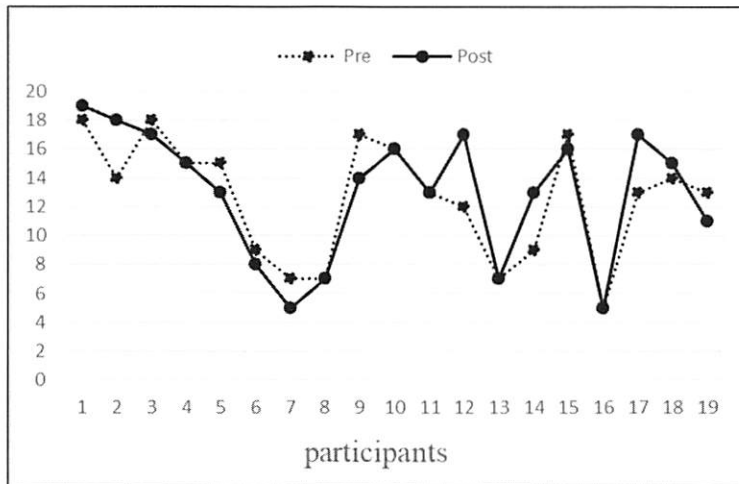


Fig. 3 Individual scores of GJT (Correction) in pre and post tests

3.2 Written Japanese to English Sentence Translation Tests

The Japanese to English sentence translation test had 13 sentences and was scored for 40 items covering 17 target grammar structures (Table 2). After the study abroad experience, the mean score decreased from 33.58 to 32.95, but the difference was statistically insignificant, $t(18) = 0.25, p = .25$. The decrease may have been caused by differences in target sentences, even though they were very similar in pre and post tests and covered the same grammar structures.

Table 2 Results of Japanese to English translation tests

	Pre		Post		<i>t</i>	<i>d.f.</i>	<i>p</i>	<i>r</i>
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>				
Target items (<i>N</i> = 40)	33.58	5.28	32.95	4.47	0.25	18	.25	.06

Individual score differences in pre and post tests (Fig. 4) show less movement than the GJT results. The largest gain was 3 by participants 8, 12 and 16, and the largest decrease was -5 by participant 10.

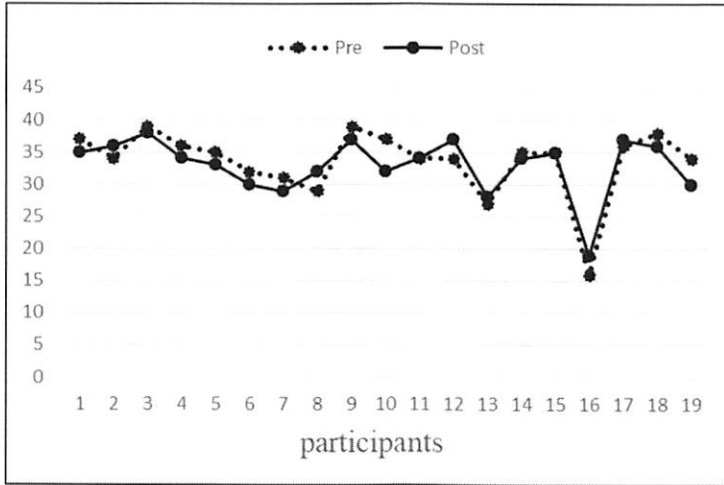


Fig. 4 Individual scores of the translation test in pre and post tests

3.3 Picture Description Tests

Pre and post picture description tests were analyzed for 11 measurements; number of syllables, syllables per second, length of the picture description from start to finish, number of unfilled pauses, total pause length, mean pause length, phonation time ratio, number of types of words used, number of tokens used, type-token ratio, and number of words per minute (Table 3). Praat software (Boersma & Weenink, 2018) was used to record the length of the description, as well as the lengths of pauses. Paired sample *t*-tests were conducted for each measurement, and only the phonation-time ratio showed a statistically significant difference with a medium effect size between pre and post tests, $t(18) = -2.99$, $p = .01$, $r = .58$.

Phonation-time ratio (PTR) is the ratio of filled (speaking) time within the whole length of utterance, and it is considered to indicate fluency. Even though the mean PTR has increased from .48 to .53, individual differences among the participants were large (Fig. 5). While seven participants' PTR increased more than .10, participant 3 and 7 decreased by .12 and .09 respectively.

Table 3 Results of picture descriptions tests

	Pre		Post		<i>T</i>	<i>d.f.</i>	<i>p</i>	<i>r</i>
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>				
# of syllables	102.84	31.33	112.21	49.59	-0.96	18	.35	.22
syllables/second	1.10	0.31	1.12	0.31	-0.40	18	.70	.09
length in seconds	104.00	52.53	106.51	51.52	-0.21	18	.84	.05
# of unfilled pauses	29.79	13.73	29.32	15.02	0.14	18	.89	.03
total pause length	60.03	43.51	50.58	29.44	1.20	18	.24	.27
<i>M</i> pause length	1.87	0.95	1.77	0.71	0.63	18	.53	.15
phonation time ratio	.48	.15	.53	.12	-2.99	18	.01*	.58
# of types	36.37	7.72	40.05	13.11	-1.44	18	.16	.32
# of tokens	75.53	23.41	83.00	38.55	-0.93	18	.36	.21
type token ratio	.50	.08	.50	.07	-0.12	18	.90	.03
words/minute	48.49	13.47	49.48	13.87	-0.37	18	.71	.09

**p* < .05

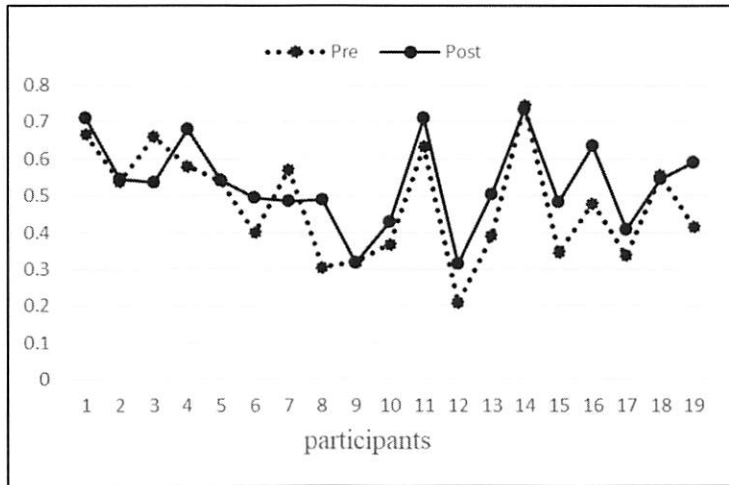


Fig. 5 Individual changes in phonation-time ratio

3.4 Standard Speaking Test (SST)

In the SST, each of the four sections (Warm-up, Picture description, Role-play, and Story-board) was rated for the six components (text type, comprehension, communication, accuracy, pronunciation, and fluency). For this study, average ratings of the four sections were used as personal ratings for the six components. Table 4 shows the mean ratings of the 19 participants for the six components, and the results of a *t*-test comparing before and after the study abroad experiences. The mean ratings increased for all components except for pronunciation, and the paired sample *t*-test showed statistically significant

increases with medium to large effect sizes in accuracy and fluency; $t(18) = -3.47, p = .00, r = .63$, and $t(18) = -2.20, p = .04, r = .46$, respectively.

Table 4 Results of Standard Speaking Tests

	Pre		Post		<i>T</i>	<i>d.f.</i>	<i>p</i>	<i>r</i>
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>				
Text type	3.49	0.38	3.61	0.54	-1.04	18	.31	.24
Comprehension	3.84	0.57	3.89	0.61	-0.46	18	.65	.11
Communication	3.16	0.55	3.27	0.67	-1.15	18	.27	.26
Accuracy	3.17	0.34	3.30	0.41	-3.47	18	.00*	.63
Pronunciation	3.54	0.52	3.37	0.41	1.90	18	.07	.41
Fluency	2.99	0.47	3.16	0.54	-2.20	18	.04*	.46

* $p < .05$

Pronunciation is rated lower in SST when it affects comprehensibility. The rater in this study, with more than 20 years' experience teaching in Japan, may have been too familiar with Japanese English pronunciation to assess consistently. Also, as pronunciation is not emphasized in Japanese English education, highly fluent and grammatically accurate English users can often have very non-native like pronunciation. Thus, the pronunciation ratings seemed likely to represent a separate construct of ability as the other items in SST. For these reasons, overall scores for each participant were calculated using average scores of the five components, excluding pronunciation, from all four sections.

Figure 6 shows the differences in the overall ratings before and after the study-abroad experiences. Participants 2, 12, 15, and 17 showed larger improvements than others, while some participants' ratings declined upon returning from studying abroad. A paired sample *t*-test showed significant improvements with a medium effect size between the overall performances of participants before ($M = 3.33, SD = 0.34$) and after ($M = 3.45, SD = 0.33$) the study abroad experiences; $t(18) = -2.17, p = .04, r = .46$.

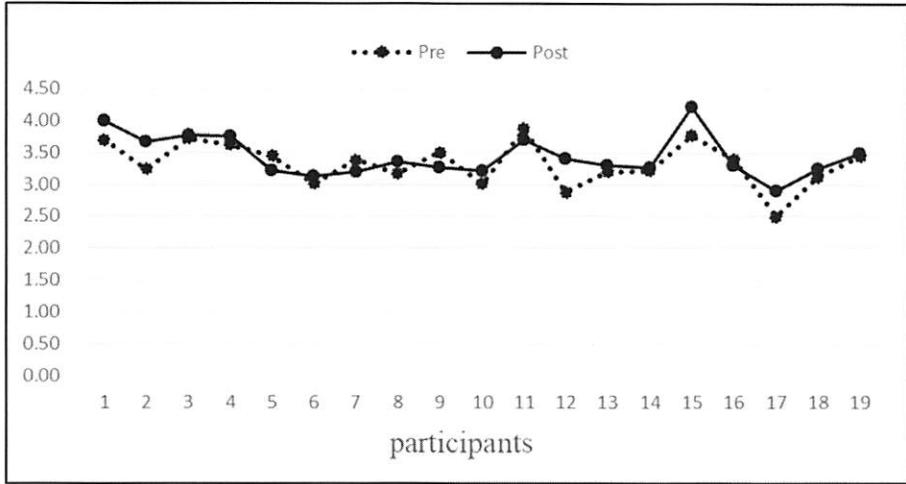


Fig. 6 Individual changes in SST mean ratings excluding pronunciation

3.5 Correlations

To answer the third research question, what factors seem to contribute to the language improvement, correlations between some of the information from pre and post surveys and improvements in test performances were analyzed. The factors investigated were; TOEIC IP scores (Listening, Reading and Total), EIKEN grades (if they held any), past travels and studies abroad experiences, past English studies outside schoolwork, current English use outside English classes, and accommodation, roommate, and classmate situations during the study-abroad.

Table 5 shows factors that had statistically significant correlations. *EIKEN* is participant's EIKEN grades (2nd grade, pre-2nd grade, 3rd grade, or none), *Homestay* means the participants stayed with a homestay family instead of at a school dormitory, *Non-Japanese roommate* means they had non-Japanese roommates at their homestay (all dormitory students had Japanese roommates), *Communication at home* means the participants actively communicated with homestay family or roommates in English, and *Classmates* means whether their classmates were all Japanese, more Japanese than non-Japanese, or more non-Japanese than Japanese. Staying with a homestay family showed correlations with the most performance results, followed by having non-Japanese roommates.

Table 5 Correlating factors

	<i>M</i> difference	<i>SD</i>	Factors				
			EIKEN	Homestay	Non-Japanese roommates	Communication at home	Classmates
GJT Correction	0.37	2.31				.55*	.49*
Description lengths	2.51	52.84	.51*				
Mean pause lengths	-0.10	0.66		.04*			
Total pause lengths	-9.45	34.24	.51*				
Syllables per second	0.03	0.26		.59**	.49*		
Words per minute	1.07	11.49		.57*	.54*		
SST text type	0.12	0.51		.56*			
SST comprehension	0.05	0.45		.53*	.48*	.49*	
SST fluency	0.16	0.33		.47*	.47*		.49*

* $p < .05$, ** $p < .01$

4. Discussion

This study had three research questions: whether short-term study-abroad experiences improve participants' English skills; if so, what aspects of English proficiency are improved; and what factors seem to contribute to the improvement. Pre and post grammaticality judgment tests on the 17 English grammar structures did not show statistically significant improvements. Written production accuracy of the same 17 grammatical structures assessed using Japanese to English sentence translation tests also did not show improvements. These results indicate that short-term study abroad experiences of three to four weeks did not affect participants knowledge of grammar or accuracy in controlled written production.

Phonation time ratio in the picture description test showed statistically significant increase with a moderate effect size, indicating improved fluency, which was also proved in the SST ratings. Even with the improved fluency, though, the participants in this study were not fluent users of English. In a study of 16 Hungarian English learners, Kormos and Dénes (2004) found that syllables per minute and phonation time ratio correlated strongly with perceived fluency by native and non-native teachers. In their advanced ($n = 8$) and higher intermediate ($n = 8$) groups, the average speech rates were 181.2 SPM and 115.9 SPM, and the average phonation time ratios were .70 and .50 respectively. In Tokunaga (2021), two native English speakers, both North American English teachers in Japan, narrated the same six pictures. Their average speech rate was 164.4 SPM, with the average phonation time ratio of .74. The participants in this study had the average speech rate of 1.12 syllables per second (67.2 SPM) and

the average phonation time ratio of .53 after studying abroad, indicating that even though they did not have long unfilled pauses, they spoke significantly slower than the learners in Kormos and Dénes (2004) and native English speaker teachers in Tokunaga (2021).

In the SST, accuracy ratings also had statistically significant improvement, not in accordance with the GJT and the translation test results. This is because SST rates accuracy not on specific grammar structures but on the types and severity of grammatical errors, whether the test-taker makes errors in simple grammar, or only in complex sentences, and how much they affect the communication.

Of course, performance changes had individual differences, and correlations were found between some of the factors and post-study-abroad improvements. Staying with a homestay family and having non-Japanese roommates seemed to have contributed to more language improvements. On an individual level, participant 12 showed improvements in all tests. This student had the highest TOEIC score of 710 to start with, and her scores on pre GJT and pre translation test were both high, indicating that she had a good proficiency of English before going abroad. During her study-abroad in the U.K., she did homestay with a single mother and her 2-year-old boy, and a Belgium roommate. In the post survey, she said she often talked with her host mother and spent a lot of time with the roommate, going to school and watching Netflix together. Also, she was placed in classes with many European students whom she said were kind and funny. Having a good English proficiency before going abroad probably helped her be proactive in communicating with her host family, roommate, and classmates. Participant 17, although having a low TOEIC score of 385, also improved her performances after studying in Australia for four weeks. She stayed with an older couple and had four roommates including one other Japanese person but not from the same school. According to the post survey, she often went out with her roommates, and the classes she attended had more non-Japanese than Japanese students. She noted that being among students from different countries helped her learn from other students' mistakes while they all tried to communicate with each other. Also, despite her low TOEIC score, she performed well on pre GJT and pre translation test, indicating a possibility that she had good base knowledge of English that she could rely on.

On the other hand, participants 5, 7, 9, and 11 had decreased scores and ratings in many of the post tests. Participants 5, 7 and 11 had lower TOEIC scores of 410, 375 and 450, while participant 9 had 620. Participant 5 did a homestay in Australia with an elderly woman but had a Japanese roommate from a different university. She spent most of her time at home speaking with her in Japanese. The school she attended had many Japanese students, and more than half of her classmates were Japanese. Participant 7 studied in Singapore and stayed with a family with three small children and a housekeeper. She had a Vietnamese roommate and talked with her until late every night. Her classes had less than half Japanese students for the first two weeks, but more in the following two weeks. Even though this participant had many opportunities to communicate in English, her low proficiency might have limited her improvement. In the post survey, she commented that she doesn't feel that her English improved. Participants 9 and 11 participated in the same four-week program in the U.S. In this program, a group of students from the university stayed in a school dormitory and studied in a non-integrated class. Even

though they had some opportunities to meet the local people on campus and in town, living and studying with only students from the same university in Japan did not seem to be an ideal environment to improve their English.

5. Conclusion

Summarizing the answers to the research questions, short-term study-abroad experiences did have limited effects on participants' English proficiency. The experience had a moderate effect on participants' spoken fluency, but not on accuracy. Having higher English proficiency before studying abroad, staying with a homestay family, having non-Japanese roommates and classmates, and making efforts to communicate with them seemed to have contributed to positive outcomes. In the post survey, some students commented on their language skills (translated from Japanese by authors):

I could not put my thoughts into words quickly.

My bad pronunciation caused communication difficulty.

People spoke too fast for me to understand.

I did not have enough knowledge or skill.

I don't think I learned anything new, but I became able to use what I already knew.

I learned to get my thoughts across with what I knew.

I became able to handle casual conversations.

I learned to use easy words to communicate.

I became to be able to talk without constructing whole sentences in my head first.

I couldn't make sentences, but I learned to communicate with words.

These comments seem to reflect the results that their knowledge and production accuracy did not improve, but they gained fluency by becoming able to use the knowledge they already had. Of course, what participants gained from the experiences were not only the language skills. As in many previous studies, the participants reported positive effects on their motivation and confidence in studying and using English (translated from Japanese by authors):

I can ask people for help without hesitation.

I am more willing to speak up.

I want to study English hard to maintain my English abilities.

I became more independent and confident.

My feeling to want to go abroad became stronger.

I learned the importance of trying without worrying about making mistakes.

I realized how low my English proficiency is, and I want to study harder.

The experience of being outside Japan made me want to experience more.

I had fun every day and could accept difficulties.

Although this study is limited to a small group of students in one university, the resulting information may be useful to other institutions and hopefully encourage further research to investigate outcomes of their programs. As we come out of the pandemic, more universities will hopefully resume

their study abroad programs. However, with travel agents, airlines and language schools trying to regain profit, and with possibilities of hotel quarantines, studying abroad may become more expensive. Therefore, it is all the more important for universities to help students get the most out of the study abroad experiences. As this study showed that short-term study abroad programs help participants get used to using what they already know, and participants with higher proficiency can get more out of the study abroad experience, pre-departure intensive English lessons may be helpful. Also, because study abroad programs can have good effects on participants' motivation to study English harder and communicate using English, providing more learning opportunities, such as elective classes, language practice lounge, or organizing presentation contests after their return may put their motivations into action before they fade.

Acknowledgements

This work was supported by Grant-in-Aid for Scientific Research (JSPS KAKENHI Grant Number 17K03041). Authors would like to thank Trevor Holster for conducting the SST interviews, and the two anonymous referees for the careful review of our manuscript.

Bibliography

- ACTFL-ALC Press. (1996). *Standard speaking test manual*. ALC Press.
- Boersma, P., & Weenink, D. (2018). *Praat: Doing phonetics by computer* [Computer software].
<http://www.praat.org/>
- DeKeyser, R. (2007). Study abroad as foreign language practice. In R. DeKeyser (Ed.), *Practice in a second language: Perspectives from applied linguistics and cognitive psychology* (pp. 208-226). Cambridge University Press.
- Ellis, R. (2005). Measuring implicit and explicit knowledge of a second language: A psychometric study. *Studies in Second Language Acquisition*, 27(2), 141-172.
<https://doi.org/10.1017/S0272263105050096>
- Ellis, R. (2009). Measuring implicit and explicit knowledge of a second language. In R. Ellis, S. Loewen, C. Elder, J. Erlam, J. Philp, & H. Reinders (Eds.), *Implicit and explicit knowledge in second language learning, testing and teaching* (pp. 31-64). Multilingual Matters.
- Fujioka, K., & Agawa, T. (2007). Motivating learners through special activities in a study abroad program. In K. Bradford-Watts (Ed.), *JALT 2006 Conference Proceedings*, 918-930.
- Geis, K., & Fukushima, C. (1997). Overview of a study abroad course. *The Language Teacher*, 21(11), 15-20.
- Godfroid, A., Loewen, S., Jung, S., Park, J., Gass, S., & Ellis, R. (2015). Timed and untimed grammaticality judgments measure distinct types of knowledge: Evidence from eye-movement patterns. *Studies in Second Language Acquisition*, 31, 269-297.
<https://doi.org/10.1017/S0272263114000850>

- Green, P., & Hecht, K. (1992). Implicit and explicit grammar: An empirical study. *Applied Linguistics*, 13, 168-184. <https://doi.org/10.1093/applin/13.2.168>
- Horness, P. (2014). A social narrative inquiry of three Japanese university participants in a short-term study abroad program. *Ryūgaku: Explorations in Study Abroad*, 7, 24 -37.
- Horness, P. (2018). Student survey responses from a short-term study abroad program. *Ryūgaku: Explorations in Study Abroad*, 11, 2 -27.
- Isa, M. (2016). Daigakusei no tanki kaigai kenshu no kouka: gakusei no jishinkan keisei yoin no kanten kara [Effects of short-term study abroad program on university students: Causes affecting students' confidence building]. *Okinawa Christian University Review*, 12, 36 - 49.
- Japan Student Services Organization (2019) *Heisei 30 nen do kaigai ryugaku keikensha tuisseki chousa houkokusho* [Follow-up survey results of study abroad program participants 2018]. http://ryugaku.jasso.go.jp/datas/master_link_details/pdf/220190904144829_bUFDx.pdf
- Japan Student Services Organization (2021), *2019 nen do nihonjin gakusei ryugaku joukyou chousa kekka* [2019 Survey results of Japanese students studying abroad]. https://www.studyinjapan.go.jp/ja/_mt/2021/03/date2019n.pdf
- Kimura, K. (2009). The influences of studying abroad on language proficiency and the use of language learning strategies: In the case of a three-week English program. *KATE Bulletin*, 23, 47-58.
- Kimura, K. (2011). How effective the short-term studying program in an English-speaking country is for Japanese university students: Focusing on the listening and writing proficiencies, *Shobi- Gakuen University Bulletin of Policy and Management*, 21, 17-30.
- Kimura, T., & Shimizu, H. (2016). Changes in students' attitudes toward cross-cultural understanding and learning motivation during the short study abroad program in the Philippines, *Kaetsu University Research Review*, 59(1), 75-99.
- Kobayashi, T. (1999). To what extent will Japanese EFL learners' English proficiency improve in intensive English programs abroad? *The Review of Liberal Arts*, 97, 83-100.
- Kormos, J., & Dénes, M. (2004). Exploring measures and perceptions of fluency in the speech of second language learners. *SYSTEM*, 32, 145-164. <https://doi.org/10.1016/j.system.2004.01.001>
- Kuramashi, Y. (2019). The effectiveness of short-term language programs for university students in the Philippines. *CASELE Research Bulletin*, 49, 1-10. https://doi.org/10.18983/casele.49.0_1
- Loewen, S. (2009). Grammaticality judgment tests and the measurement of implicit and explicit L2 knowledge. In R. Ellis, S. Loewen, C. Elder, R. Erlam, J. Philp, & H. Reinders (Eds.), *Implicit and explicit knowledge in second language learning, testing and teaching* (pp. 94-112). Multilingual Matters.
- Matsuda, Y. (2012). Result and significance of overseas study short program -The result of the questionnaires and the reports, *Journal of Nagoya Bunri University*, 12, 11 -16.
- Nishina, Y., Omotedani, J., & Morishita, M. (2017) A case study of effects of a short-term study abroad program on Japanese university students. *Journal of Kobe Gakuin University Center for Teacher Education*, 3, 1-16.

- Nonaka, T., & Seki, K. (2016). Change in English proficiency through a homestay program and a short-term study abroad program. *Niigata Seiryō University Junior College Bulletin*, 46, 89-97.
- Otsuka, K., & Negishi, J. (2009). The effect of a two-week study abroad program on speaking fluency and the relationship among fluency, English use anxiety, and English classroom anxiety. *The Bulletin of the Kanto-Koshin-Etsu English Language Education Society (KATE)*, 23, 59-70.
- Roehr, K. (2008). Metalinguistic knowledge and language ability in university-level L2 learners. *Applied Linguistics*, 29, 173-199. <https://doi.org/10.1093/applin/amm037>
- Ryugaku Journal (2019). *Ryugaku hakusho 2019* [White paper on study abroad 2019]. Ryugaku Journal.
- Sakai, H. (2008). Implicit and explicit grammatical knowledge of L2 English: Identification, correction, and provision of rules. *Annual Review of English Language Education in Japan*, 19, 91-100.
- Sasaki, N. (2017). Effects of short-term study abroad program for language training, *Bulletin of the University of Electro-Communications*, 29(1), 47-55.
- Shimada, K. (2010). Bunpousei handan tesuto ni okeru mondai teiji jikanseigen no umu to meijiteki anjiteki chishiki. [Timed/timed grammaticality judgment test and implicit /explicit knowledge]. *St. Andrew's University English Review*, 24, 41-53.
- Suzuki, R., & Hayashi, C. (2014). The effects of short-term study abroad programmes on students' English proficiency and affective variables. *The Bulletin of the Kanto-Koshin-Etsu English Language Education Society (KATE)*, 28, 83-96.
- Tateyama, E. (2002). The response of Japanese nursing students to a vacation English program in Australia, *Bulletin of the Japanese Red Cross Junior College of Akita*, 7, 89-97.
- Taura, H., Horii, K., Manishi, T., Okada, H., Shimizu, D., Kashiwamoto, E., & Tonari, T. (2009). A report on a three-week intensive English programme at Massey University, NZ. *Journal of Language and Culture: Language and Information*, 4, 1-22. <https://doi.org/10.24729/0000275>
- Tokunaga, M. (2017). Exploring English grammar knowledge and written production of Japanese EFL learners, *Studies in English Teaching and Learning in East Asia*, 6, 51-69.
- Tokunaga, M. (2021). Investigating fluency and accuracy of Japanese university EFL learners' spoken English production, *Journal of English Teaching*, 7(2), 163-178.
- Zhang, R. (2015). Measuring university-level L2 learners' implicit and explicit linguistic knowledge. *Studies in Second Language Acquisition*, 37, 457-486. <https://doi.org/10.1017/S0272263114000370>