Second Language Development Before, During, and After Study Abroad: A Longitudinal Study at a Japanese Women’s University

IIDA, Tsuyoshi  
Doshisha Women’s College of Liberal Arts

HERDER, Steven  
Kyoto Notre Dame University

Abstract
This paper describes a four-year longitudinal study, investigating how English teachers collaborate in their teaching before study abroad, how university students’ study overseas, and to what extent they develop both academic and general English skills. Twenty-seven female students out of ninety in a department of international studies participated. These students had to study overseas for about one year. Quantitative and qualitative analyses were used. Observations were adopted to examine collective teachers’ efficacy (Hattie, 2009). Students’ diaries were examined to see how they studied overseas. Preparation for the TOEFL iBT and TOEIC IP tests were employed respectively for academic and general English skills development throughout four years. Post study abroad interviews were carried out to examine students’ beliefs and attitudes about learning English. The results showed high collective teacher efficacy helped to develop students’ English skills. Students studied intensively not only before study abroad (BSA) but also during study abroad (DSA) and after study abroad (ASA). Standardized tests showed that there were significant differences in general reading and listening, but not in academic reading and listening before and after study abroad. The results suggest that the study abroad program helps learners develop their general English proficiency, but does not always help improve some parts of their academic English proficiency.

Keywords: study abroad, collective teacher efficacy, academic English, general English, longitudinal study

Introduction
This paper aims to report on the development of students’ English proficiency throughout four years of university, including one-year study abroad, by focusing on the interaction between teaching and learning as determined through a collaborative teaching approach among eight professors, because to the best of our knowledge, no previous research has explored it longitudinally within the study abroad literature. The paper documents improvements in English language proficiency of Japanese students studying English over four years, who have also participated in a year-long study abroad experience as part of the
requirements to graduate from a department of international studies at a women’s university in Western Japan. It also reports how teachers in this program showed high collective teacher efficacy (Hattie, 2009) not only with each other but also by collaborating with students to teach and learn in this program. English language proficiency gains are measured from two perspectives: 1) Academic English using the TOEFL iBT test and 2) General English using the TOEIC test, which were both administered throughout the students’ four years of study.

While there has been considerable interest in the development of English language proficiency among university students in Japan, such research tends to generally involve students staying in Japan for the duration of their major programs. However, a growing trend in English language education in Japan is the incorporation (particularly in private universities) of study abroad years into university curricula. However, the efficacy of such policies in terms of their impact on students’ overall English language proficiency over the course of a four-year program has not been explored in detail to date. The research presented here sought to address this gap by examining 27 students’ English language proficiency development using two indexes over their four-year course of studies.

Therefore, we started a basic study of students’ development of English proficiency by collecting their ongoing test scores and study skills. We focused on our students who study overseas because a growing number of departments that send students overseas are being created at present. We realized that we needed to describe how students learn English and to what extent they develop their language skills during their university life. The university department in this study aims to prepare students to be active and useful members of Japanese society, with confidence in academic critical thinking skills, strong communication skills and an ability to understand and solve problems proactively.

This paper not only tracks qualitative and quantitative data regarding English proficiency gains, but also introduces how a collaborative teaching approach focusing on the concept of collective teacher efficacy (Donohoo, 2017) and a balanced focus on Complexity, Accuracy, and Fluency (CAF) in the classroom helped attain the research study results. It is hoped this paper will also add useful ideas to the literature for departments with a study abroad component. It will be especially useful for those interested in exploring and expanding ways to have a positive influence on the academic and professional development of university students in Japan, offering objective data on how university students who experienced study abroad actually struggle to develop their English proficiency during university life.

**Background**

Study abroad is now commonly viewed as one of the best ways to improve one’s language skills (Fry, 1984; Kinginger, 2008), and Freed’s seminal paper in *The Interdisciplinary Journal of Study Abroad* (1998) provides a full background of issues related to language learning and study abroad from the 1960s to the 1990s, focusing on students from North America, Australia, Europe, and Scandinavia. While Japan had been mentioned as a destination for Western students studying abroad, it was not until Fry (1984) described young Japanese technicians sent abroad to acquire new technical skills during the early period of Japan’s modernization that Japanese students studying abroad are first introduced into the literature.

However, Serrano, Llanes, and Tragant (2011) contend that there has been little research...
confirming the superiority of study abroad over an intensive program at home except for a few areas, most notably oral fluency. This contradicts the common belief that learning a target language in the country where it is spoken is much better for language progress than L2 classroom learning in a country where it is not spoken. Some research demonstrated that after study abroad, students became more fluent than those in the intensive program at home (e.g., DeKeyser, 1991; Freed, 1995; Lafford, 2004; Serrano, et al. 2011). Other research showed that students with overseas experience gained more vocabulary than those who studied in an intensive program at home (DeKeyser, 1991; Dewey, 2004). Except for these language skills, however, students in a study abroad program were not always better than those in an intensive program at home (Collentine, 2004; DeKeyser, 1991).

There has been some research on Japanese students who studied abroad (e.g., Kimura, 2011; Sasaki, 2007); with one of the earlier papers describing Japanese study abroad students on a short, 15-week study abroad in New Zealand (Tanaka & Ellis, 2003), and focuses primarily on the relationship between student beliefs about language learning and their language proficiency as measured through the TOEFL ITP (paper-based test). Prior to sharing their own findings, Tanaka and Ellis convey conclusions from previous studies about linguistic gains from a general perspective and in relation to Complexity, Accuracy, and Fluency: First of all, students who participate in a study abroad make greater linguistic gains in English than students who study English only in their own country. Second, study abroad has a bigger linguistic impact on lower level students than on students with a higher language ability prior to study abroad. Next, overall oral language proficiency improves more than any one specific area such as reading, listening, or grammar. Finally, Complexity and Accuracy do not change noticeably, while Fluency gains are strong (2003).

Iida (2013) examined how Japanese university learners of English improved their English after a two-semester study abroad program by using TOEIC scores, a questionnaire, and interviews. In total, 92 female university students took the test before study abroad (BSA) and after study abroad (ASA) and answered a questionnaire ASA. They were divided into four groups based on the types of courses they took overseas: (a) Academic Course (AC) group refers to the group of students who took only academic courses, (b) Academic + ESL Course (AC+ESL) group means students who took both academic and ESL courses simultaneously, (c) ESL→AC stands for students who took ESL first and moved on to academic courses afterwards, (d) ESL-only means students who took only ESL courses throughout their time abroad. Results showed that the different groups developed reading and listening skills differently.

However, Iida’s study did not show the development of academic English skills through study abroad. Since university students should develop academically in addition to improving their L2 skills, it is necessary to consider not only L2 proficiency, but also students’ academic achievements when researching university students’ development of L2 proficiency. Obviously, it is true that we need a precise experimental design to research whether a study abroad program or an intensive program at home is better for students to acquire L2 proficiency, such as having two groups; one is students who study abroad, and the other is students who study intensively at home. However, such scientific research design is not always effective for practical purposes such as developing students’ L2 proficiency. It has been
almost fifteen years since such a study abroad program was created in Japan at Akita International University (Lehner, 2011) and there are certainly not enough studies of such programs and their outcomes. What is important in research design, as far as a study abroad program is concerned, is to examine and reflect on students' current development of English, learn more about their overseas learning, and continue to improve the curriculum based on data collected in the program, particularly considering the interaction between teaching and learning. Thus, the present study employed a longitudinal study with the following research questions.

**Research questions**

1. How does collective teacher efficacy influence students’ basic skills in English in an intensive English program BSA?
2. How do students adjust to studying in a foreign university?
3. How do university students in the program develop academic and general English skills over the course of four years of undergraduate study?

**Study Abroad Program**

To better understand the parameters of the study, it is important to provide a structural description of the study abroad program. The study was carried out in a department of international studies in a women’s university located in the western part of Japan. All students in this department have to study abroad in an English-speaking university for two semesters, the second-year autumn and the third-year spring semesters. The curriculum of this department has three stages: BSA, DSA, and ASA. There are two categories of courses: English skills courses (skills courses) aimed at developing students’ English skills, and content courses that focus on British and American studies, Japanese Studies, and International Studies. In principle, almost all courses are conducted in English. The first research question focuses on skills courses because English language teaching plays an important role in this department; in particular, skills courses are closely related to students’ development of English, especially BSA.

From the first-year spring semester to the end of the second-year spring semester (three terms), students study English intensively to obtain a high TOEFL iBT and overall GPA score, both of which are used to decide their target university in Canada, the US, the UK, New Zealand, or Australia. A placement test is administered during the freshman orientation period to create eight small skills classes streamed by level with a maximum of twelve students each. The aim of these skills classes is to have students begin studying with students at their same level. While students’ passive English knowledge may be somewhat similar, their active English ability is vastly different. For example, based on an online vocabulary test administered in 2016, there was virtually no difference in passive vocabulary knowledge among all eight classes. However, there was an obvious difference in the amount of experience that students have in actually accessing and using English actively. By placing students in these small streamlined classes, teachers can pay closer attention to students’ English level and help develop their skills individually. The highest TOEFL iBT score they acquire will have an impact on where and how long they can study abroad.
The length of staying overseas is different from university to university: for instance, nine months for students who go to Britain and twelve months for students in Canada. Students have to stay in an English-speaking country for at least two semesters, or nine months. The department has a relationship with twenty-seven, English-speaking universities to which a student can apply. Their target university is decided based on their highest TOEFL iBT and GPA scores BSA. Through required courses known as Tutorial I & II, four teachers of content courses share the responsibility of contacting all students DSA by means of e-mail and an online virtual classroom. The students have to submit two reports each month: the first one to their assigned professor about their independent research required DSA and the second report to support staff at the home university about their daily lives and safety. They are required to prepare and present their research questions and method in a poster session BSA, then research and report on it every month DSA, and finally, produce results and submit a paper on their topic ASA. Through these Tutorial courses, teachers can monitor students' life, their English language learning in ESL courses, academic courses, and students' own research overseas.

After returning from study abroad, all students usually take one required English skill course, one content course, a discussion and debate course, and a graduation thesis seminar. Depending on the number of credits students obtained BSA and DSA, they may need to take several other courses ASA to graduate in four years. In the graduation thesis seminar, which is four-credits in each semester, they are required to choose their own themes to study, read academic materials, make presentations, and finally submit a 3000-word graduation thesis in English. Even if students have not gotten any overseas credits, almost all of them usually take courses in the home university and are able to graduate from the university within four years.

Method
Participants
Based on a university research grant budget, twenty-seven participants out of ninety who enrolled in the 2013 academic year took part in this study. The number of participants were small because the TOEFL iBT test is very expensive. Most of the grant budget was used for students’ taking the test. However, we selected participants carefully so that they represented the whole range of students in this program. Three or four participants were chosen randomly from each English skills class streamed by level, based on scores of a placement test at the beginning of the first year. At the beginning of the research, we planned to get data from 32 students who were randomly selected in each skill class streamed by level, asking them to take the test. Since the present study is longitudinal and the TOEFL iBT test is offered at specific test centers at fixed times, only twenty-seven students were able to submit scores.

Data Collection
Regarding research question 1, we observed and described features of English skill classes and the collective teacher efficacy approach to collaboration in 2016 and 2017, which was almost the same as in 2013. We observed teachers’ and students’ collaboration and interaction in 2016 and 2017 because of the convenience of collecting data. Also, students’ speaking and writing fluency were examined in each of their skills classes. Research question
used students’ diaries and monthly reports from their lives overseas. Eleven students agreed to cooperate in this study by submitting their diaries almost every month. One author also collected data through interviews, giving ten participants an interview about learning English and their motivation to study English. In particular, he questioned them about how they studied English in the intensive English program in the first- and second-year BSA, in ESL and academic courses DSA, and in academic courses ASA. Moreover, since teachers in this department had a chance to talk with students overseas through the online virtual classroom and a great number of opportunities to teach and talk with the participants from first year to fourth year, we could observe these students’ attitude toward learning at university carefully. The interviewees and journal writers were selected by considering students’ levels of English proficiency so that they represented the whole department. Research question 3 employed the TOEFL iBT and TOEIC IP tests as those of academic English skills and general English skills, respectively. All students in this department had to take the TOEFL iBT at least twice in their first year and the TOEIC IP once in the second year BSA. Furthermore, students were required to take the TOEIC IP ASA, but they did not have to take the TOEFL iBT test because they were no longer applying for admission to an overseas university. For us, it was quite difficult to get TOEFL iBT scores ASA because while they need the TOEIC score to prove their English proficiency for job hunting, they simply do not need the TOEFL iBT any longer. We were able to obtain students’ data of these tests by appealing to their own curiosity about how they would score after being abroad for about a year. As shown in Table 1, we obtained results of two TOEFL iBT tests and one TOEIC IP test BSA, and then the results of one TOEFL iBT test and one TOEIC IP test ASA.

Table 1
Tests employed in the study from 2013 - 2016

<table>
<thead>
<tr>
<th>Year</th>
<th>BSA</th>
<th>DSA</th>
<th>ASA</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1st year in June</td>
<td>1st year in December</td>
<td>2nd year in June</td>
</tr>
</tbody>
</table>

Data Analysis
To examine our research questions, both quantitative and qualitative analyses were adopted. First, skill teachers’ collaboration and teaching were observed and analyzed. Also, their weekly meetings were observed and analyzed. Next, authors examined eleven students’ diaries written DSA that they had been asked to keep before their departure and they tried to observe patterns of learning overseas. They read the diaries carefully, in particular focusing on how students DSA studied English. Two students’ initial and final journals entries were included in this paper. One student directly enrolled in academic courses without taking ESL courses, the other began in an ESL course and moved to academic courses. Finally, quantitative analyses were conducted. The means of total TOEFL iBT scores were compared from first-year in June, to first-year in December, and finally, to fourth-year to see if there were
significant differences among them by using a repeated-measures ANOVA, followed by comparing means of the reading, listening, speaking, and writing sections of the TOEFL iBT in these periods to see differences, employing a repeated-measures ANOVA. Next, means of the listening and reading scores of the TOEIC IP were compared to see differences between BSA and ASA by means of a paired t-test. With respect to qualitative analysis, we adopted the methods of interview, journal studies, and observation. The data collected by these methods were described and used as supplementary data for interpretation of academic and general English skills.

Results

Collective Teacher Efficacy and Skills Development

We begin with describing English teachers’ collaboration through a collective teacher efficacy approach. Bandura (1997) originally coined the term self-efficacy to mean “the conviction that one can successfully execute the behavior required to produce outcomes” (p. 193). Educators have dubbed teacher self-efficacy as a teacher’s belief that they can do what is necessary to have an impact on student learning (Donohoo, 2017). When teachers join forces in a school and share a belief that by working together, they can make a difference for students, it is called collective teacher efficacy (Donohoo, 2017).

Our small learning communities of 12 students offer many advantages to both the instructors and the students, including teacher-student rapport, individualized attention, ongoing formative feedback (that often pertains to most students in the class), emotional support among peers, friendly competition within the class and with other classes, or even with previous years’ classes (aiming to be the best “C” class in our department’s history), and through shared failures and successes. Small-sized skills classes are a distinct and purposeful part of the skills program design.

Skills classes are tailored to preparing students for specific questions they will face on the TOEFL iBT test. Besides practicing specifically for the reading, listening, speaking, and writing sections of the TOEFL iBT, a number of additional skills are focused on as core English skills: typing fluency, speed reading, extensive reading, speaking fluency, academic reading skills, and intensive listening strategies.

While students clearly struggle with both English input and English output for the first few months of the academic year, they certainly develop the invaluable traits of perseverance, grit, and resilience they will need while living abroad for a year in an English-speaking country.

To illustrate the focus on fluency skills at the start of the first year, Table 2 shows improvements in speaking and writing output fluency. As can be seen, there was an overall 65% gain in speaking speed from the first class of the year to the last class. Students were given equivalent topics: Talk about your friends in the pre-test and Talk about your family in the post-test. Students recorded a one-minute answer on their smartphone, then listened and counted words. They were given a second chance as well, then they submitted their average score. It was a meaningful activity for most students as they saw proof of their hard work throughout the year. Additionally, Table 2 shows a 42% mean average gain for the 10-minute writing (by hand) activity. Students did these speaking and writing tests at separate times and in separate classes, but the prompts were the same.
Furthermore, Table 3 indicates the gains that first-year students made in typing proficiency during 2016. It is clear that everyone improved their output fluency with a 74% average gain. However, the degree to which each class improves is the result of both individual effort and the team effort by each skills class. While fluency output is important for first-year students preparing to study abroad, it is not the only goal of linguistic development.

The skills teachers’ team determined that it is important for students to develop both general and academic English proficiency to some degree BSA. The total amount of English studied in the first three semesters, excluding students’ self-study, is about 500 hours. Besides skills classes, they study content subjects such as British studies, North American studies,
Japanese History, Japanese Culture, and International Relations as required subjects, and other courses as elective subjects, which are usually taught in English.

**Students’ Adjustment to Studying Abroad**

While they are overseas, students are divided into three sojourn types: direct entry into academic courses (AC), the course from ESL to academic (ESL→AC), and the ESL-only course. Students who gain a higher TOEFL iBT score that a target university set can take academic courses from the beginning. A student in the direct entry group (Student A) wrote in her diary when the new academic year started as follows:

**October 11^{th}**

*I am sick. Suffering from slight fever and headache last night, I feel better now except for the sore throat. The weather suddenly got cold these days. I really should’ve taken more care of myself. Moreover, this first academic week was both physically and mentally demanding that I feel tired now.*

**October 20^{th}**

*My first assignment for Gender and Women’s Studies is to write a 1000-word review of an article. Here’s the instruction,*

> Using one of the seminar readings from weeks 1–5, summarize the key points and critically assess the arguments being made.

*The problem with this assignment is that we have to “critically assess” the arguments. I’ve never done this kind of essay before, so this assignment instruction sounded complicated. After reading some of the tips for writing article review on the Internet, I now understand better than before.*

Although her English was excellent, she felt that the academic work was demanding and she did not have enough experience writing arguments critically. It is true that she learned critical reading at her home university, but such preparation was not sufficient when taking this academic class. Another student who commenced with ESL courses (Student B) described herself at the beginning of the course:

**September 3^{rd}**

*I took ESL test. I did not speak English well in the interview. At the interview, I told that I would like to start from AP4 or AP5 so that I could take academic courses. However, I was discouraged to hear that I was at AP3 level. I was anxious about the announcement on next Friday because I didn’t know results of writing and reading tests.*

Even if she had obtained sufficient scores required by the ESL course at the target university, she needed to take a placement test at the beginning of the ESL courses and was placed according to the test. The student wanted to know whether she could succeed in taking academic courses or not at the beginning of the ESL course because she had to return within one year. This is the reason why she got nervous about her class.
At the end of the course, Student A wrote about her class as follows:

May 11\textsuperscript{th}

\textit{Even before the exams, we are introduced with new topics in sociology lectures. Today’s lecture was about Marxism. The lecture itself was very approachable, but the seminar reading that I did after I went back to my room was a little bit complicated. We are required to read the first chapter of The Communist Manifesto by Karl Marx for this week’s seminar. Since I didn’t want to spend ages doing one seminar reading, I read the chapter in Japanese before reading it in English. However, it turned out that The Communist Manifesto was still difficult even in Japanese. I spent a lot of time reading this today…}

She sometimes read texts in Japanese first to understand the English version. Her overseas university assigned her many reading assignments. Student B reflected on her academic course:

March 13\textsuperscript{th}

\textit{I had an academic class in the morning. I submitted a draft paper. In the discussion period today, I put my opinions into shape before class so that I was able to state my opinion in the discussion. However, it was difficult to express it. When I tried to say it, a topic changed! I could not follow their speed of discussion. However, in the afternoon class I could say what I wanted to say to a teaching assistant! I have a lot of things to improve with my speaking skill.}

As mentioned earlier, she was placed in the course that did not allow her to take academic classes at the beginning of the course. However, she took several standardized academic tests that acknowledged her as an academic student, and finally, she was accepted as a university student. Still, she needed time to get used to discussion in the target language.

Although it is difficult to report on other students’ cases of learning overseas due to space limitations, almost all students tried to do their best to take academic courses at their English-speaking university. Whereas a year abroad at a foreign university is a significantly challenging and stressful experience, it is true that a very small number of students become sick or suffer from mental illness BSA or DSA. Finally, about five percent of the students enroll in academic courses directly. About ninety percent of students move from ESL to academic courses. These students make a considerable effort in ESL courses to obtain the English proficiency that each university establishes to allow them to enroll in academic courses, because according to the home university’s rules, students who only take ESL courses overseas cannot transfer the credits from those courses to the home university. In the end, about five percent of the students are not successful in enrolling in academic courses.

However, all students in this department can choose an independent study topic that relates to their own personal research questions; they research these questions, report on their research every month to teachers at their home university, have a Mid-term online presentation DSA (Iida, 2014), have a final presentation ASA, and finally submit a 1500-word report upon their return to Japan. This Tutorial course gives students eight credits. In other
words, even if students cannot take academic courses abroad, they are able to get at least eight credits overseas as long as they succeed in passing this course.

**Development of Academic and General English**

Moving from qualitative to the quantitative, Table 4 shows descriptive statistics of total scores of the TOEFL iBT. To examine research question 3, in the first place, the means of total scores of the TOEFL iBT were compared between the three periods by carrying out a repeated-measures ANOVA. The results show that differences between conditions were unlikely to have arisen by sampling error ($F(2, 52)=31.46, p<.001$); an overall effect size of 0.548 (partial $\eta^2$) showed that 55% of the variation in scores can be accounted for by the development of academic English. Pairwise comparisons using Bonferroni showed that there were differences in scores between first-year in June and first-year in December ($p<.001$) and between first-year in June and fourth-year, but not between first-year December and fourth-year ($p=.056$). The results show that learners develop their academic English proficiency BSA but not DSA or ASA.

Next, the development of the four skills of the TOEFL iBT (Table 4 and Figure 1) were examined separately, beginning with reading, and then, listening, speaking, and finally, writing. A repeated measures ANOVA with Greenhouse-Geisser correction determined that iBT reading mean did not show a statistically significant difference among the three periods.

<table>
<thead>
<tr>
<th>Table 4</th>
<th>Descriptive Statistics of TOEFL iBT totals</th>
</tr>
</thead>
<tbody>
<tr>
<td>Period</td>
<td>1st-year  2013</td>
</tr>
<tr>
<td>Mean</td>
<td>53.7</td>
</tr>
<tr>
<td>SD</td>
<td>16.2</td>
</tr>
</tbody>
</table>

*Figure 1. Descriptive Statistics of TOEFL iBT total.*
Table 5

*Descriptive statistics of four skills of TOEFL iBT*

<table>
<thead>
<tr>
<th></th>
<th>1st year June 2013</th>
<th>1st year December 2013</th>
<th>4th year August 2016</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Reading Mean</strong></td>
<td>13.2</td>
<td>14.4</td>
<td>15.3</td>
</tr>
<tr>
<td><strong>SD</strong></td>
<td>4.8</td>
<td>4.7</td>
<td>4.6</td>
</tr>
<tr>
<td><strong>Listening Mean</strong></td>
<td>12.0</td>
<td>16.0</td>
<td>15.6</td>
</tr>
<tr>
<td><strong>SD</strong></td>
<td>4.9</td>
<td>4.6</td>
<td>4.9</td>
</tr>
<tr>
<td><strong>Speaking Mean</strong></td>
<td>14.1</td>
<td>16.1</td>
<td>17.5</td>
</tr>
<tr>
<td><strong>SD</strong></td>
<td>4.7</td>
<td>3.8</td>
<td>3.4</td>
</tr>
<tr>
<td><strong>Writing Mean</strong></td>
<td>14.4</td>
<td>17.2</td>
<td>18.8</td>
</tr>
<tr>
<td><strong>SD</strong></td>
<td>4.6</td>
<td>4.0</td>
<td>3.9</td>
</tr>
</tbody>
</table>

*Figure 2.* Line graphs of four skills of TOEFL iBT in three periods.

$F(2, 51) = 3.06, p = .056)$. Regarding listening, a repeated measures ANOVA with Greenhouse-Geisser correction found that the iBT listening mean differed significantly among the three periods ($F(2, 40) = 14.78, p < .001$); an overall effect size: $0.36$ (partial $\eta^2$). Pairwise comparisons using Bonferroni showed that the difference in scores between first-year in June and first-year in December and between first-year in June and fourth-year are significant ($p < .001; p = .004$), but not between first-year in December and fourth-year ($p = 1.00$).

Next, the same method of analysis adopted in the case of iBT speaking showed significant differences among the three periods ($F(2, 45) = 24.6, p < .001$); overall effect size: $0.52$ (partial $\eta^2$); Pairwise comparisons using Bonferroni resulted in differences in scores among the three periods (first-year in June and first-year in December: $p < .001$; first-year in June and fourth-year: $p < .001$; first-year in December and fourth-year: $p = .026$).

Finally, the same statistical method was employed regarding iBT writing, finding that significant differences are found among the three periods: ($F(2, 48) = 27.6, p < .001$) with an overall effect size of $0.52$ (partial $\eta^2$). Both pairwise comparisons showed differences in scores between first-year in June and first-year in December ($p < .001$), between first-year in June and fourth-year ($p < .001$), and between first-year in December and fourth-year ($p = .035$).

To summarize the development of academic English over four years, it can be said that...
productive skills develop BSA, DSA, and ASA whereas receptive skills do not always develop before and after study abroad except in listening before study abroad.

Research question 2 addressed the development of general English proficiency. Table 6 and Figure 3 illustrate students’ general English skills before and after study abroad. A one-way repeated measures ANOVA showed a significant difference in listening skills between pre-study abroad and post-study abroad (F(1, 25)=51.87, \( p < .001 \)); the overall effect size was .68 (partial \( \eta^2 \)). Also, the same analysis was conducted regarding TOEIC Reading, resulting in a significant difference between them (F(1, 25)=10.99, \( p = .003 \)), with an overall effect size of 0.31 (partial \( \eta^2 \)).

Table 6

<table>
<thead>
<tr>
<th></th>
<th>BSA</th>
<th>ASA</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2nd-year</td>
<td>3rd-year</td>
</tr>
<tr>
<td></td>
<td>June 2014</td>
<td>November 2015</td>
</tr>
<tr>
<td>Listening</td>
<td>M</td>
<td>SD</td>
</tr>
<tr>
<td></td>
<td>328</td>
<td>79.7</td>
</tr>
<tr>
<td>Reading</td>
<td>M</td>
<td>SD</td>
</tr>
<tr>
<td></td>
<td>299</td>
<td>66.0</td>
</tr>
<tr>
<td>Total</td>
<td>M</td>
<td>SD</td>
</tr>
<tr>
<td></td>
<td>627</td>
<td>138</td>
</tr>
</tbody>
</table>

Figure 3. Line graphs of TOEIC listening and reading between two periods.

Discussion

Impact of Collective Teacher Efficacy

Two factors have a meaningful impact on the results of this study as far as Research question 1 is concerned. The first key factor is a departmental commitment to collaboration through a collective teacher efficacy approach. The skills classes are taught by six
experienced native English-speaking teachers (NESTs) and two Japanese teachers of English (JTEs). As a policy, all teachers are either full-time permanent faculty or full-time contract faculty, who have a clear focus on the program, treat each other as equals, and teach in as much of an English immersion atmosphere as possible. These teachers meet at a weekly lunchtime meeting to discuss a wide range of topics: student issues, class dynamics, reflections on lessons, new ideas for skills development, data on students’ output, and so on. A shared Google Drive Sheet keeps minutes of the meetings, useful links, ongoing student data, work deadlines, and other departmental information. The meeting is chaired by a skills coordinator, but any teacher can add topics to the weekly online agenda. Additionally, student handouts, student seating charts with pictures, lesson plans, and other resources are shared in an online Dropbox folder, which is accessible to all teachers. This level of collaboration nurtures trust, support, teamwork, motivation, and a sense of teaching satisfaction. While it certainly requires extra work and an increased level of patience and listening skill, the rewards far outweigh the effort required.

The second key factor is a balanced approach to addressing Complexity, Accuracy, and Fluency (CAF). These three measurable constructs can all indicate improvements in language proficiency: complexity gains are indicated by vocabulary acquisition, larger vocabulary usage, or increased sentence length in both written and spoken output; accuracy gains can easily be measured in written or spoken output; and finally, counting words per minute, measuring the number or the length of pauses, or recording extended discourse can all be indications of fluency development gains.

Unfortunately, for students planning to study abroad, the still ubiquitous Grammar Translation Method in Japanese high school English classes focuses almost completely on Complexity and Accuracy (Nishino, 2008, 2011; Sakui, 2004). This creates a huge disadvantage for students who need to score well on a test such as the TOEFL iBT, which demands speed and fluency. To address this handicap that many Japanese students have, an initial focus on fluency output can help students in many ways: First, as they see their output increasing, they gain some self-confidence, thus beginning a virtuous cycle: their efforts and incremental successes lead to motivation, which can result in confidence gains, which in turn results in more effort, etc. Additionally, fluency output promotes the idea that English is a communication tool, and therefore a focus on expressing meaning is actually more important than being 100% correct with grammar. Finally, increased fluency prepares students to be able to interact much more successfully in an international environment during their study abroad.

It should be highlighted that it does not take a long time for most students to acquire some sense of fluency. They fill in their output data for speaking, writing, speed reading, and typing on their online student profile several times a week. Within a few months, teachers can see an increase in student self-confidence. The students are then challenged to take the next step in learning—to refocus their energies on either their weaknesses in accuracy, or to focus on increasing complexity in their output. From one of the authors’ high school teaching experience, he noticed that most students up to low intermediate level choose to try to “fix” their output by focusing on their accuracy, while high intermediate and above students tend to challenge themselves to develop more complexity in their speaking and writing output. Some of the ways that a fluency approach is implemented will be presented in the next section.
Developing English Proficiency

The results of research question 3 showed that students develop not only general English proficiency but also academic English proficiency throughout four years at university. However, the increase in academic proficiency BSA is not as high as that of DSA and ASA; in particular, unlike productive academic skills, receptive academic skills do not improve significantly DSA and ASA. These results might be accounted for by the curriculum of the department, students’ experiences overseas, or students’ motivation to study English. In particular, due to students’ intrinsic motivation to study abroad and teachers’ involvement in the program BSA, as well as students’ experiences learning in the English-speaking countries DSA, and the curriculum and students’ motivation to study English ASA, might all have an influence on their general and academic English skills. Based on these reasons, discussions of the results of general and academic English proficiency during university are deemed valuable.

Almost all students in this program study English much more than regular university students in Japan, especially BSA because of the curriculum requirements, students’ strong motivation to study abroad, the small learning community, and teachers’ involvement in this program. The curriculum requires students to study all four English skills intensively. Many students enter the program based on fulfilling a dream to have a significant study abroad experience, and a more general desire to improve their English ability for some kind of future career. It is not clear whether they have any specific interest in international studies or pursuing a career using English when they first enter the program. However, they do have interest in English itself and they are somewhat aware of the challenges they face in preparing to study abroad. First and foremost, they need to secure a student visa by attaining a minimum TOEFL iBT score (45/120), TOEFL ITP (450) or an IELTS band averaged at level 5.0, depending on the particular entry benchmark established by each of the twenty-seven universities that the department sends students to each year.

This challenging requirement has a great advantage for teachers of this department that students enter the program not only highly motivated but also all collectively working towards the same clear goal to achieve as high a score as possible on these standardized tests. Students begin the academic year in April and take their first practice online practice TOEFL iBT test in May. Then, they take the real TOEFL iBT first in June and then again in December. Many students also take the TOEFL iBT test or IELTS test at various times of their own volition.

We are convinced that the small learning community approach as well as the use of information and communication technology (ICT) in the classroom helps students improve their English proficiency. On the first day of the program, some 80–90 students take a placement test to measure their reading, listening, grammar structure, and writing skills. Based on their results, students are placed into a skills class from the highest level “A” to the lowest level “H”. In addition to the class size, as shown before, use of ICT such as typing programs, online websites, and Google Drive are indispensable for students’ development of the target language mainly because they can get immediate and ongoing feedback regarding their improvements, and thus see their ongoing development of ICT skills in class.
Considering the fact that it takes time for students to realize an improvement on the TOEFL iBT test BSA or that some students are discouraged by the results of the test in spite of their great efforts to improve their English, the development of the ICT skills shown in Tables 2 & 3 may help almost all students feel some sense of accomplishment in class BSA.

**Adjustments to Study Abroad**

According to diary accounts and interviews with students, almost all students led a full life DSA as mentioned earlier. They fulfilled the requirement set by the department and obtained admission to an English-speaking university before their departure. As explained earlier, about five percent of students went directly into academic courses abroad and about ninety percent of students took ESL courses in the first semester and moved to academic courses and foundation courses, while about five percent stayed in ESL courses throughout their time abroad. In other words, except for students who could enter academic courses directly, most students needed to take ESL courses in the beginning of the overseas program, which suggests that students might need time to strengthen the four skills more to gain higher academic proficiency. English-speaking universities usually assign students a great deal of reading homework and have much more time for discussion in class than Japanese universities. It is true that most English-speaking universities in which students in this department enroll, demand students to tackle such assignments. However, according to students’ interviews, not all universities always give them a lot of materials to read or have discussions in class. They said that compared to learning at their home university, there were some universities or courses that demanded less reading assignments and discussion than they did BSA. This might be related to the fact that most students have taken introductory level courses DSA. The results of their academic receptive skill gains suggest that they would have needed to study at least two more semesters in the academic courses if they had wanted to improve academic English proficiency.

**Development of Academic and General English**

However, with respect to general English proficiency, students’ great exposure to the target language communities may improve their proficiency as shown in Table 6. Students’ daily life on campus and in a homestay in English-speaking countries automatizes listening and reading skills. Interviewing several students who began to study in the ESL courses reveals that they enjoyed communicating with ESL learners from different non-English speaking countries because they got exposed to and learnt about different cultures. They have less speaking anxiety communicating with other ESL learners, but their level of speaking anxiety rises when they communicate with native speakers of English. They were afraid of making grammatical errors and having an embarrassing moment when they could not express themselves. Almost all students said that at the very beginning it was difficult to discuss topics given by professors with native speakers of English in the classroom due to their natural English-speaking speed. Gradually, they became accustomed to such an environment. Although students who shared a room with a native speaker of English in the dormitory or actively joined some club activities made good friends with native speakers, students who did not have such experiences had difficulties in communicating with native
speakers. In general, students tend to develop more general English proficiency than academic English proficiency.

Finally, the post-study abroad program, job-hunting, and students' idea about their future career might influence students' development of academic English and motivation to study academic English. Since by the end of their third year, most students have obtained enough credits to graduate except for senior-year required courses, seniors are usually only required to enroll in a senior seminar, which aims at writing at least a 3000-word thesis in English. They usually read not only Japanese but also English references and papers to write the thesis, but it is difficult to say that all students read as many papers and books as they are expected to by the professor. One of main reasons for taking a small number of courses is that students want to find a job before anything else. The career program held by the career support center at the university begins soon after they return home in their third year. In fact, most companies begin recruiting before the fourth academic year. Students are usually busy attending company explanation meetings, taking tests, and being interviewed for a job. They sometimes go all the way to Tokyo twice a week to have interviews because many companies have their head office in the capital city. This job-hunting period usually continues until summer holidays, but sometimes until the end of December or February of the following year. Knowing the severity of job-hunting, they tend to avoid taking many courses when they become seniors even if the department provides English courses to further develop language proficiency.

Moreover, only a few students in this department want to study at graduate school in Japan or overseas. These students study academic English hard to get the required English proficiency for graduate school, but other students do not study academic English because they only want to work after graduation. Actually, their interest in English turns away from academic English to general English during university life because they believe that the higher their TOEIC scores, the better job they think they can find. According to Onda (2017), students' motivation to study English in their department changes during the four years; in particular, most of them found it difficult to continue to study academic English. Such dynamic changes of motivation to study English and their study abroad experiences, and job hunting might result in the statistically non-significant development of academic English proficiency.

Conclusion

The research found that the key factors are a departmental commitment to collaboration through a collective teacher efficacy approach and balanced approach to addressing CAF. It also clarified that university students who studied overseas develop general English skills, but the development of academic English is not as high as we expected. The present study abroad program seems to be effective for the development of academic English because of the curriculum requirement, teachers' involvement in the program, and students' motivation to study English. However, it appears to be difficult for students to develop academic English proficiency because of insufficient period of studying introductory academic courses at English-speaking universities, their job-hunting challenge, and less motivation to continue to study at graduate schools overseas. One thing that we
should keep in mind is “The responsibilities of educational authorities, qualifying examining bodies and teachers cannot simply be confined to the attainment of a given level of proficiency in a particular language at a particular moment in time, important though that undoubtedly is” (Council of Europe, 2001, p. 5). In this paper, we did not concentrate on the attainment of English proficiency through standardized tests but employed them as measures to evaluate how students achieve proficiency of a second language throughout the whole study abroad program.

Although the present study adopted a longitudinal study approach because of the limited research model options for researchers who study an intensive English program in a department where students are required to study abroad, the disadvantages of case studies cannot be denied, such as a lack of generalization of research findings. Therefore, further similar case studies need to be carried out. Not only studies on linguistic gains but also ones on other advantages of study abroad, for instance, exploring homestay, unplanned informal encounters, service encounters, and the overseas classroom DSA might be necessary for further studies.

**Acknowledgement**

This research was funded by a grant of Doshisha Women’s College of Liberal Arts (DWCLA) in 2016. We appreciate three anonymous reviewers’ invaluable comments and suggestions in this paper, as well as the considerable efforts made by students throughout the four years of their studies.

**References**


