

Readiness, Language Contact, and Oral Performance Development

During a Study-Abroad Program

by

MIKAMI Hitoshi

DISSERTATION

Submitted in Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy

GRADUATE SCHOOL OF INTERNATIONAL DEVELOPMENT

NAGOYA UNIVERSITY

Approved by the Dissertation Committee:

Toru KINOSHITA

Junko YAMASHITA

Shuji OZEKI

Approved by the GSID Committee: December 21, 2016

## **Acknowledgements**

I am indebted to Prof. Toru Kinoshita, Prof. Junko Yamashita, and Prof. Shuji Ozeki for their continuous support of my doctoral studies. I would also like to express my gratitude to the 46 participants of this study and Natalie-Anne Hall.

## Preface

This dissertation is an original work by the author, Hitoshi Mikami, submitted for the degree of Doctor of Philosophy at Nagoya University.

One part of this study (i.e., some descriptions in 1.1 and 1.2 (pp.10-28) and the data reported in study one (pp. 37-96)) is based on the author's own past work and has been reported in two academic journals (see, Mikami, 2014b, 2014d in the reference list). The copyright holders of these papers, The Japan Association of College English Teachers (JACET) and The Japan Association for Language Education & Technology (LET) respectively, gave permission to reuse these works for this dissertation (JACET: 20 May 2016 via e-mail; LET: 15 January 2016 via e-mail). In relation to this, it must be noted that there are minor differences between the numeric data reported in study one of this dissertation and those reported in Mikami (2014b, 2014d). These differences arose for the following reasons. First, in his 2014 studies, the author made a few errors in the calculation of descriptive statistics and the transcription of the data. The former relates to calculations in Table 1 (Mikami, 2014d) and the index labelled as the declarative knowledge index (Mikami, 2014d). The latter relates to Appendix A of Mikami (2014b) (the TOEFL iBT® score of participants No. 17, 23, and 24 in particular). These errors were corrected accordingly in this dissertation. It must be

emphasized that, as we will see in what follows, none of these corrections change the general findings of the 2014 studies. Second, while the author used the raw scores of each index for the statistical calculations in Mikami (2014d), he utilized the z-scores of the same indexes for this dissertation (i.e., the tests in 5.3.1, 5.4, 6.2, and Table 4 in particular). The employment of z-scores resulted in slightly different statistical test results from Mikami (2014d). Note that the general findings reported in Mikami (2014d) remain the same regardless of the above-stated modification.

Also, some descriptions in the literature review section (pp. 18-23) were borrowed from the author's own past work, the data of which is stored in the database of the University of Warwick. As the status of the above paper is unpublished (E. Ushioda, personal communication, February 19, 2015), the reuse of the above work for another project (i.e., this paper) is justifiable (B. Bassetti in the MA dissertation report, 11 September, 2015).

## **Table of Contexts**

<b>1. Introduction .....</b>	1
<b>    1.1. Literature Review .....</b>	10
<b>    1.2. Focus of This Study .....</b>	24
<b>2. Research Questions .....</b>	29
<b>3. Method (Study One).....</b>	37
<b>    3.1. Participants.....</b>	38
<b>    3.2. Design of SA Program.....</b>	41
<b>    3.3. General Procedures and Description of Materials.....</b>	42
<b>        3.3.1. Oral performance indexes .....</b>	43
<b>        3.3.2. Grammatical knowledge index .....</b>	44
<b>        3.3.3. International posture.....</b>	46
<b>        3.3.4. Language contact profile .....</b>	48
<b>        3.3.5. Interviews .....</b>	50
<b>4. Data Analysis (Study One) .....</b>	51
<b>    4.1. Research Question One.....</b>	52
<b>    4.2. Research Question Two.....</b>	53
<b>    4.3. Research Question Three.....</b>	56

<b>4.4. In-Class TL Contact Data .....</b>	58
<b>5. Results (Study One) .....</b>	59
<b>5.1. Summary of Participants' Information .....</b>	59
<b>5.2. Research Question One.....</b>	64
<b>5.3. Research Question Two.....</b>	65
<b>5.3.1. Simultaneous multiple regression analysis.....</b>	65
<b>5.3.2. Interview data analysis.....</b>	68
<b>5.4. Research Question Three.....</b>	82
<b>6. Discussion (Study One) .....</b>	83
<b>6.1. Research Question One.....</b>	83
<b>6.2. Research Question Two.....</b>	84
<b>6.3. Research Question Three.....</b>	91
<b>7. Educational Implications (Study One) .....</b>	93
<b>8. Method (Study Two).....</b>	97
<b>9. Data Analysis (Study Two) .....</b>	98
<b>9.1. Research Question Four .....</b>	98
<b>9.2. Research Question Five .....</b>	99
<b>10. Results (Study Two) .....</b>	100

<b>10.1. Summary of Participants' Information .....</b>	100
<b>10.2. Research Question Four .....</b>	102
<b>10.3. Research Question Five .....</b>	106
<b>11. Discussion (Study Two) .....</b>	108
<b>11.1. Research Question Four.....</b>	110
<b>11.2. Research Question Five.....</b>	124
<b>12. General Discussion.....</b>	133
<b>13. Conclusion.....</b>	144
<b>References .....</b>	149
<b>Appendix .....</b>	170
Appendix A: Participant information .....	170
Appendix B: Descriptive statistics on two groups' TL contact .....	178
Appendix C: Original question items in the language contact profile.....	179
Appendix D: The definition of terminology .....	180

## **1. Introduction**

It is a widely held belief that language learners can effectively improve their oral performance in their target language (TL) in a study-abroad (SA) context<sup>1</sup> (Allen, 2010; DeKeyser, 2007b, 2010; Kinginger, 2011; Magnan & Back, 2007). This assumption is actually true for the learners who are able to make the most use of learning resources in an SA context. Such learning resources are called language contact (hereafter TL contact) in SA research and this encompasses the variety of language activities (a) from which SA participants can learn knowledge and skills in their TL (i.e., written and spoken inputs, including feedback for one's production), and (b) where SA participants can practice their newly learned knowledge and skills in the TL for more proficient oral performance (Mikami, 2014b, 2014c, 2014d). When language learners are capable of utilizing TL contact, it allows

---

<sup>1</sup> This study uses the term ‘study-abroad’ as follows: learning one’s TL as a main subject in a host country where it is spoken by inhabitants as their first language.

them to construct a spontaneous TL learning cycle; this self-investment widens the opportunity for TL leaning/practice; and, as the learning cycle continues, the chance of TL oral development will also increase (Allen, 2010; Freed, Segalowitz, & Dewey, 2004; Hernández, 2010; Llanes & Muñoz, 2013; Sasaki, 2011).

The above-stated chain of events also in turn indicates the existence of less successful scenario. That is, the development of TL oral skill will be more or less limited when language learners cannot benefit from TL contact and therefore fail to build a spontaneous and continuous TL learning cycle in an SA context (Allen, 2010; Davidson, 2010; DeKeyser, 2010; Trentman, 2013).

As shown in the above two different scenarios, language learners' TL contact behavior is of great importance to TL learning in SA contexts. It is natural, therefore, that the following questions (and alike) have been two of the most central issues in SA research: (a) what are the attributes of different TL contact behavior in SA learning contexts? and (b) how

and to what extent is language learner's TL contact behavior connected to the development of TL oral skill? Regarding these questions, SA studies have so far revealed that there are at least two different aspects that affect SA participants' TL contact behavior and TL oral skill development.

On one hand, there are a number of factors that an SA participant has little or no control over, and yet limit and/or decide the amount and characteristics of TL contact (Davidson, 2010; Hernández, 2010; Kinginger, 2011; Sasaki, 2011; Segalowitz & Freed, 2004; Trentman, 2013). As an example, the timeframe and/or design of an SA program is often pre-fixed by SA administrators in the case of sponsored SA programs (e.g., a student exchange program between two affiliated educational institutions) (e.g., Isabelli-García, 2010). Let us consider two different SA groups that join sponsored SA programs. That is, learners who join an eight-month interaction-oriented SA program and those who participate in a two-month instruction-oriented SA program. Here, it is clear that (a) the two SA groups

go through quite different TL learning experiences (i.e., an eight-month SA vs. a two-month

SA; an interaction-oriented program vs. an instruction-oriented program) (DeKeyser, 2007b;

N. Taguchi, 2008) and (b) yet the main attributes of these differences lie in the structure of

the SA program, not in learner's factors, such as each individual's intensity of TL learning

motivation. Such structural difference, especially the time frame of SA, brings about

difference in the linguistic gains of SA participants: as a general tendency, the longer the SA

time frame becomes, the greater the TL skill developments SA participants will show

(Davidson, 2010; Sasaki, 2011). Indeed, as these data suggest, the structure of SA programs

(and alike, such as the design of immersion program) has been the interest of a number of

studies (Davidson, 2010; DeKeyser, 2007b; Freed, Segalowitz, et al., 2004; Llanes & Muñoz,

2013; Sasaki, 2011; Yashima & Zenuk-Nishide, 2008; Yashima, 2009)

Important as the investigation of SA program structure is, learner's factors also play a

significant role in changing SA participants' TL contact behavior and their linguistic

achievement. Learner's factors in this context can be paraphrased as the psychological attributes (Borsboom, Cramer, Kievit, Scholten, & Franic, 2009) relevant to the pursuit or avoidance of TL contact. In SA literature, learner's factors are not limited to (a) learners' state of the mind (e.g., types or the intensity of learning motivation) (Allen, 2010; Hernández, 2010) but also encompass (b) their mental skills that affect the efficiency of TL processing and change the consequences of TL contact events (DeKeyser, 2010; Golonka, 2006; Mikami, 2014a). The latter aspect—the mental/linguistic skills—may need further explanation. For instance, if a learner fails to comprehend a TL text due to the lack of grammar skill, it first prevents him/her from effectively learning from a given TL contact event. Also, one's psychological reaction to the given TL contact event will be negative under this circumstance; as such negative experience continues, one's learning motivation and/or the self-efficacy belief in the domain of TL reading will be by and large damaged; and the decrease in learning motivation and self-efficacy both can trigger the disruption of

self-investment in the domain of TL reading (Mikami, Leung, & Yoshikawa, 2016; Schunk

& Zimmerman, 2007; Ushioda, 1998). In that sense, language learners' linguistic skill forms

part of the learner's factors as it is responsible for changes in the effectiveness and

continuance of TL contact.

The findings of some prior SA studies have clarified the intriguing characteristics of

these learner's factors. That is, some learner's factors behave as predictors of success or

failure of TL learning in SA contexts (Allen, 2010; Davidson, 2010; Dekeyser, 2010; Glonka,

2006; Hernández, 2010; Mikami, 2014a, 2014b, 2014c). These predictors can be paraphrased

as the readiness for productive SA learning. As its name indicates, readiness in this context

stands for the sufficiency or insufficiency (or existence or lack thereof) of specific learners'

factors at a pre-SA stage. In a nutshell, the more readiness one equips oneself with, the greater

the benefit of TL contact, and thus the more TL development one will achieve. Although a

number of psychological attributes can potentially constitute the readiness for productive SA

learning, the findings of prior SA studies showed the significant value of two types of psychological attributes for the construction of spontaneous and continuous TL learning cycle in an SA learning context. The first psychological attribute is the degree of TL grammatical knowledge (Davidson, 2010; DeKeyser, 2010; Mikami, 2014b, 2014d); and the second one is the type and vividness of TL learning motivation (Allen, 2010; Hernández, 2010; Mikami, 2014c; Tanaka & Ellis, 2003; Yashima & Zenuk-Nishide, 2008).

Because readiness factors are under each individual's control and change the effectiveness of SA learning, mapping out the impact of initial readiness on TL contact behavior and the development of TL oral performance has significant value for both researchers and practitioners (e.g., SA administrators). Yet, we still have little understanding of how two different types of readiness—grammatical skill and motivation—jointly affect SA participants' TL contact behavior during an SA visit and change the outcome of SA learning, due to a lack of empirical data.

In response to these problems, the author decided to clarify the relationships between initial readiness, TL contact behavior, and the development of TL oral performance. To this end, the author sampled 46 Japanese university students studying English as their foreign language (JEFL university students) who joined a one-semester SA program, and documented their initial readiness, TL contact behavior during their SA participation, and the development of their TL oral proficiency. Through the analyses of these data, the author will show the detailed relationship between SA participants' initial readiness, their TL contact behavior during the SA visit, and their gains in TL oral proficiency.

The main body of this study consists of 13 chapters. In what follows, the author first briefly reviews relevant SA literature so as to show how the two different types of readiness—grammatical skill and motivation—are conceptually and empirically connected to the pursuit or avoidance of TL contact, and the development of TL oral performance. Then, the author will narrow down the focus of this study. Following that (i.e., in chapter two), the

five research questions will be introduced. For the convenience of data analyses and discussions, the author has divided his investigations into two studies—study one and study two. Chapters three to six are allocated to the methodology (chapter three), data analysis procedure (chapter four), results (chapter five), and discussion (chapter six) of study one. Educational implications derived from the results of study one will be shown in chapter seven. Chapters eight to 11 deal with the methodology (chapter eight), data analysis procedure (chapter nine), results (chapter 10), and discussion (chapter 11) of study two. In chapter 12, the author will sum up the general findings of this study and discuss future directions for SA studies. The conclusion of the whole study is given in chapter 13.

Appendices will directly follow the main body of this study. Appendix A is the complete set of participant information; Appendix B is the descriptive statistics of the TL contact data; Appendix C is the original question items used in this study; and Appendix D provides the definition of terminology used in this study.

## **1.1. Literature Review**

First of all, initial grammatical knowledge is an important factor for TL learning/practice during an SA program. From the novice to even some advanced stages, adult language learners cannot comprehend nor produce their TL spontaneously without reflecting on their “consciously known, verbalizable grammar rules (a form of declarative knowledge)”; (DeKeyser, 2010, p.84) (see also Appendix D). In theory, if learners equip themselves with an ample amount of declarative knowledge in addition to sufficient vocabulary, they can decode TL inputs and also assemble TL outputs without a problem. However, to conduct real-time language activities, such as conversation, learners are required to apply complex knowledge of TL for both comprehension and production, and simultaneously must deal with time and psychological pressures, such as anxiety (Allen & Herron, 2003; MacIntyre, Noels, & Clément, 1997). Under these circumstances, consciously referring to every relevant rule needed for language activity will not only be cognitively demanding, but also an inefficient

and ineffective way to perform language activity instantaneously (Ellis, 2009; Krashen, 1985).

To solve this problematic situation, language learners need to transfer their consciously

known linguistic rules to situation-specific rules (i.e., behavioral routines) through initial

and/or repeated domain specific practice (e.g., oral practice is indispensable for behavioral

routinization in oral performance). Those specific rules are also known as procedural

knowledge. Because procedural knowledge is routinized within learners' language systems, it

can be applied to specific language performance much faster, and with a lower error rate and

cognitive load (see also Appendix D). After this, through a sufficient amount of practice,

specific rules will be synthesized into more general rules applicable to a broader range of TL

performance. This transitional process is called automatization. As automatization proceeds,

learners can achieve further "spontaneous, effortless, fast and errorless (DeKeyser, 2007a,

p.3)" performance within TL activity (Robinson, 2007; Skehan, 2002) (see also Appendix D).

From this, it is clear that appropriate and sufficient practice is an indispensable factor for the

development of TL performance.

Nevertheless, second and foreign language learning contexts often do not afford a satisfactory amount of language practice needed for knowledge routinization and eventual automatization, particularly in the domain of speaking (Freed, Segalowitz, et al., 2004; Llanes & Muñoz, 2013; Muñoz, 2006). In contrast, when we think of language learning in an SA environment, it seems largely self-evident that learners will attain ample access to TL practice through a wide variety of TL contact. Considering this is widely believed to be an advantage of the SA learning context, it is not so surprising that in many cases TL learners participate in an SA program expecting considerable improvement in their TL oral performance (Allen, 2010, pp. 27-28). However, prior SA studies have revealed that, in order to learn from TL inputs and conduct oral practice effectively, learners need to equip themselves with an adequate level of grammatical knowledge before the start of their SA (Brecht, Davidson, & Ginsburg, 1995; Davidson, 2010; Golonka, 2006; Mikami, 2014b). Without this, SA

participants will be overwhelmed by the amount, speed and complexity of TL inputs (Segalowitz & Freed, 2004), and fail to both learn from them and conduct sufficient oral practice needed for knowledge proceduralization (i.e., the initial step of knowledge automatization). Even though SA participants can compensate for their lack of grammatical readiness and make room for oral practice relying on scaffold language learning during an SA program, the amount of oral practice provided by the strong-scaffolding will be too small to realize considerable oral development. This is because of its passive and less interactional nature compared with the weak-scaffolding or out-of-class TL contact (DeKeyser, 2007b, 2010). This disadvantage in the lack of readiness is considered to apply particularly to participation in shorter SA visits rather than longer ones, because less-prepared SA participants can spare more time for less-structured TL classrooms or out-of-class TL contacts within a one-year SA participation than a half-year abroad program, for example (e.g., Davidson, 2010).

Despite the above assumption, the link between grammatical readiness, TL contact and development of TL oral performance has not been fully supported by data. For instance, Brecht et al. (1995) and Davidson (2010), two of the largest sample studies in SA literature, quantitatively proved that initial grammatical knowledge level can be the predictor of oral performance development during SA participation (see also Golonka, 2006); however, they only speculated that the reason for this probably lies in the difference in the efficiency of TL learning. Moreover, even when SA participants' TL contact data was recorded in the form of observation (or questionnaire) data, the reason for the pursuit and avoidance in TL contact rooted in the existence or lack of initial grammatical readiness was rarely reported. In one of the few exceptions, DeKeyser (2010) found that, during a six week SA program, participants with less grammatical knowledge were overloaded by TL input given by native speakers from the first stage of their sojourn. This inadequacy in comprehension not only prevented these participants from learning from inputs, but also demotivated them to pursue further

spontaneous TL interaction. Even though this learning experience prompted these learners to re-acknowledge the value of grammar knowledge, they wound up their SA program with discouraging results in the development of TL oral performance.

The above findings shed light on the importance of capturing SA participants' affective factors related to their TL learning so as to illustrate a clearer relationship between grammatical readiness and the development of TL oral performance during an SA program.

As touched on above, SA participants have to experience a sufficient amount of TL contact in order to facilitate the knowledge proceduralization and automatization required for the development of TL oral performance. This is to say, SA participants need to maintain a positive orientation toward their investment, even if sometimes their learning behavior fails to deliver positive learning results (Allen, 2010; DeKeyser, 2010; Hernández, 2010; Mikami, 2014c). For example, especially in the first stage of their sojourn, SA participants might fail to communicate their opinions, be unable to understand interlocutors' utterances and receive a

negative evaluation or discouraging feedback for their TL performance. As a number of motivational studies in applied linguistics have documented, these negative learning experiences by and large damage language learners' sense of control over their learning process (Dörnyei, MacIntyre, & Alastair, 2014; Dörnyei, 2001). Although it is possible for (language) learners to repair their motivation and reorient themselves toward reinvestment even after unfavorable learning experiences (Ushioda, 1996, 1998), repeated exposure to negative outcomes can be a very strong factor that diminishes learners' belief in self problem-solving (Bandura & Schunk, 1981; Dörnyei, 2001; Schunk & Zimmerman, 2007). Thus, when SA participants' emotional state is seriously damaged by frequent exposure to negative learning experiences, they start to lose control over their ability to master TL through self-investment or the methods that they are currently using. We can see examples of these negative beliefs in cases where SA participants are convinced that TL contact outside the classroom (i.e., unaided TL learning/practice) is completely out of their reach. When these

negative beliefs toward one's aptitude, ability and/or external factors progress to a serious stage and become chronic, they lead to disruption of learning behavior (Abramson & Seligman, 1987; Dörnyei, 2001).

When we reanalyze less-prepared participants in DeKeyser (2010) within this theoretical frame, we can demonstrate a reasonable connection between grammatical readiness, TL contact behavior and TL oral development. Inadequacy in initial declarative knowledge first causes frequent exposure to negative learning outcomes; negative learning outcomes corrode the participants' control over their ability to master TL through spontaneous TL contact; and a sense of a lack of control orients the participants toward avoidance of TL oral practices with native TL speakers (stated as "because they were too painful", p. 89). On the contrary, in this framework, it is also credible to hypothesize that well-prepared SA participants will be able to protect or even improve their positive orientation toward spontaneous TL contact because they are more capable of learning from inputs and applying

their TL knowledge for speech production. In other words, they are able to attain more encouraging learning outcomes for their investments (or we can say that they can learn their TL more effectively). As education literature has demonstrated, positive learning experience enhances language learners' self-efficacy and self-value toward pursuing their learning objectives, and prompts motivated learning behavior (Allen, 2010; Dörnyei, 2009; T. Taguchi, Magid, & Papi, 2009; Zimmerman, 2006). Finally and most importantly, it is expected that their motivated learning behavior in turn facilitates development of TL oral performance through expanding the ground for knowledge proceduralization and automatization.

Prior SA studies also reported the intriguing role of initial learning motivation in SA participants' pursuit and avoidance in TL contact. First, Allen (2010) conducted a follow-up survey using the questionnaire and interviews. His participants ( $n = 6$ ) were American university students studying French as their TL and participating in a six-week SA program (their TL proficiency was intermediate). In his study, Allen explored the types of motivation

which are connected to his participants' spontaneous TL learning behavior throughout SA participation. Allen conducted individual interviews with his participants multiple times during the SA program and documented the reasons for the increase or decrease in his participants' out-of-class TL contacts. In Allen's study, the participants were required to respond to the questionnaire first and then asked to explain why they were engaging in the language activities that they answered in the questionnaire. The results of Allen's study indicated that the participants who had strong linguistic motive from the beginning of their SA (e.g., one joined an SA program mainly because he/she wanted to improve their TL skill) made a greater amount of TL contact and improvement in their oral TL performance than those who expected more pragmatic benefits from their SA experience (e.g., future career development). According to Allen, the above results were caused by the difference in vividness of personal goals. The participants who had strong linguistic motive could specify small goals that they needed to achieve in order to accomplish their final goal when they

encountered their shortcomings in their daily life situations. For instance, one could notice that he/she needs to master how to use the past tense so as to improve his/her speech accuracy when he/she failed to in their communication with others. In other words, if the improvement of TL skills is the main target of the SA, an ample amount of TL contact allows SA participants to notice the gap between their current TL skills and the level that they want to reach (see also, e.g., Long, 2014). This also makes it possible to set small personal goals through which they can come closer to their ideal TL proficiency. On the other hand, those who expected pragmatic benefits from their SA experience could not imagine what type of TL skills they should train or what kind of weakness they had for future career development (note that they had no previous working experience before their SA visit). Thus, unlike linguistically motivated students, the students with pragmatic motive failed to detect problems that they should deal with, and could not construct a productive learning cycle.

Second, sampling 20 American Spanish learners, Hernández (2010) revealed that a

specific type of initial motivation promoted the SA participants' out-of-class TL contact during a one-semester SA program. In his study, integrative motivation—an interest in communicating with the target language group as well as positive attitudes toward TL native speakers and their culture (p.601; see also Masgoret & Gardner, 2003)—showed a strong connection with the amount of TL contacts and eventual gains in TL oral proficiency, whereas instrumental motivation (see also Appendix D) did not.

Third, Mikami (2014c) applied the findings of Allen (2010) and Hernández (2010) to JEFL learners. In his study, Mikami sampled 24 Japanese JEFL university students who participated in a one-semester SA program. The affective factor employed in Mikami's study is called international posture and it represents “a tendency to see oneself as connected to the international community, to have concerns for international affairs and a readiness to interact with people other than Japanese (Yashima & Zenuk-Nishide, 2008, p. 567)”. He used this index instead of integrative or instrumental motivation for the following reasons. First, as a

large amount of language learning motivational literature points out, it is difficult for foreign language learners to picture a solid and tangible image of specific TL groups or cultures, especially if learners are physically distanced from TL-speaking countries and/or their target language is one spoken worldwide such as English (e.g., Dörnyei, Csizér, & Németh, 2006). Similarly, foreign language learners' motivational orientation is not clearly separable as integrative or instrumental (Yashima, 2000, 2002). For foreign language learners, the objective of language learning is inevitably affected by scholastic requirements, as well as enjoyment and self-values included in the language learning process. For these reasons, more and more recent second and foreign language motivational studies consider language learning motivation as a learner's self-related cognitive system which continuously interacts with one's learning environment (Dörnyei et al., 2014; Dörnyei, 2009). International posture is one reflection of the above self-related motivation, in that, under its conceptualization, TL learning is considered a window to access a world in which learners want to participate for

communicational and practical reasons, and therefore stronger international posture represents greater incentive for TL learning (Kormos & Csizér, 2008; Yashima, 2002). Based on these reasons, Mikami (2014c) investigated the impact of pre-SA international posture on the amount of spontaneous TL contact during the SA visit<sup>2</sup>. The results of his study revealed that international posture accounted for 36% of the variance of his participants' spontaneous TL contact behavior during the SA visit, and showed the usefulness of this index as the predicting variable of spontaneous TL contact in the case of JEFL university students.

---

<sup>2</sup> Mikami (2014c) was conducted as one of pilot studies of this study. The same applies to Mikami (2014b).

## **1.2. Focus of This Study**

The above-mentioned studies paved the way for the current study. First, although DeKeyser (2010) conducted exhaustive observation for his study, the six-week SA program may be too short in duration to realize a practice/proceduralization cycle and development of TL oral performance. With regard to this, SA investigations suggested a one-semester (i.e., six months) SA program as a possible benchmark for TL oral performance improvement (Freed, Segalowitz, et al., 2004; Golonka, 2006; Hernández, 2010; Mikami, 2014c; Segalowitz & Freed, 2004).

Second, the main focus of prior SA studies was on the individual impact of initial grammatical skill and motivation on the pursuit or avoidance of TL contact and/or the development of TL oral performance. To put it another way, so far little is known about the combined effect of these aspects on the amount of spontaneous TL contact (spontaneous TL learning effort), and the development of TL oral performance during an SA visit.

Third, it is often hard for SA studies to discuss the generalizability of their research findings (J. D. Brown, 2001, 2006). The major reason for this lies in the sampling issues shared among SA investigations. Most SA studies sample the students of a single educational institution (e.g., the students of X University) who join the same SA program (Allen, 2010; DeKeyser, 2010; Dewey, 2004; Freed, Segalowitz, et al., 2004; Hernández, 2010; Kohro, 2001; Magnan & Back, 2007; Sasaki, 2011; N. Taguchi, 2008; Trentman, 2013). With this sampling method, SA studies can minimize the risk of including confounding variables (Hoshino & Maeda, 2006; Rosenbaum & Rubin, 1983) in their data as researchers can assume the following: (a) the participants' pre-SA TL proficiency was by and large the same and (b) they experience the same (or at least similar) in-class TL learning during their SA visit as the content of in-class TL activities is often predetermined in the case of sponsored SA programs. Eliminating the influence of these between-participant differences from the data analysis is of importance particularly when we are focusing on the impact of spontaneous (i.e., out-of-class)

TL contact on the outcome of SA learning (personal communication with one of the

anonymous reviewers of JACET journal vol.57).

As one might expect, however, it is hard for researchers to attain ample sample size

with this sampling method. As an example, if an implementation of SA program is annual in a

given university and we sample 25 participants a year, it takes at least two years to attain 50

participants (we also should bear in mind that the response rate can hardly be 100% in the

case of longitudinal studies) (cf., Dörnyei & Taguchi, 2010). Sampling participants from

many different educational institutions can be a solution to this problem; however, in this case,

adjustment of between-participant difference becomes much harder. It is not surprising,

therefore, that SA studies have often been conducted with small numbers of participants as a

result of the protection against confounding. In particular, to the author's knowledge, a sample

size of less than 30 is common in SA studies (Allen, 2010; Dewey, 2004; Freed, Segalowitz,

et al., 2004; Hernández, 2010; Isabelli-García, 2010; Kohro, 2001; Magnan & Back, 2007;

Mikami, 2014b; Trentman, 2013).

This traditional sampling method, however, has two major limitations. First, if we sample our participants from a single educational institution, it is highly plausible that the results of our study would merely reflect the characteristics of the learners in a given educational institution, not that of the target population (J. D. Brown, 2001, 2006). Second, the smaller the sample size becomes, the harder to generalize the findings of a given study because yielding significant  $p$ -value, sufficient power ( $1 - \beta \geq .80$ ), and non-zero confidence interval (CI) in statistical tests will be harder along with the decrease in sample size (Chernick & LaBudde, 2011; Chernick, 2008; Faul, Erdfelder, Buchner, & Lang, 2009; Faul, Erdfelder, Lang, & Buchner, 2007; Larson-Hall & Herrington, 2010; Mizumoto & Plonsky, 2015; Plonsky, Egbert, & Laflair, 2015; Tufféry, 2011). These disadvantages attached to prior SA studies have prevented them from illustrating a clear statistical link between initial readiness, the pursuit or avoidance of TL contact (i.e., spontaneous TL learning effort), and

the outcome of SA learning. Therefore, this study decided to overcome the problem of prior SA studies by increasing the sample size and collecting data from wider varieties of SA participants.

The above-stated reasons prompted the author to investigate (a) the combined impact of grammatical and motivational readiness on the pursuit or avoidance of TL contact, and (b) the impact of spontaneous TL contact behavior on the development of TL oral performance.

## **2. Research Questions**

There are two different paths to investigate the link between initial readiness, the spontaneous TL learning effort, and the development of TL oral performance. The first way is to investigate the link between grammatical and motivational readiness, the sum of spontaneous TL contact, and the development of TL oral performance (Hernández, 2010; Mikami, 2014d). In this case, we consider the sum of SA participants' out-of-class (i.e., spontaneous) TL contact as the intensity of self-investment. Here, spontaneous TL contact represents the domain-general learning effort in that the sum of all TL activities makes up the total amount of TL contact. When we quantify the spontaneous TL contact data this way (i.e., as a single variable), the questions that we are trying to answer are as follows: (a) does initial readiness impact on the general amount of spontaneous TL learning effort? and (b) does the fluctuation in the general amount of spontaneous TL learning effort affect the development of TL oral performance?

If we yield positive answers to these questions, this means that initial readiness is of importance for more productive SA learning as readiness affects the outcome of SA learning through changing SA participants' spontaneous TL contact behavior. Knowing the existence of such a link is particularly helpful for SA administrators in that they can expect greater outcomes from SA programs if they conduct a pre-SA training program through which prospective SA students can enhance their TL grammar skill and/or international posture (Davidson, 2010; Golonka, 2006; Mikami, 2014c; Yashima & Zenuk-Nishide, 2008; Yashima, 2009).

As a downside, when we treat the spontaneous TL contact as a single variable, the following information will remain unknown: (a) the impact of initial readiness on the domain-specific TL contact behavior; and (b) the relative contribution of each particular type of TL contact (i.e., reading, listening, writing, and speaking) on the development of TL oral performance. With these data, we can clarify the detailed functions of initial readiness in

spontaneous TL contact behavior during an SA visit (Allen, 2010; DeKeyser, 2010; Hernández, 2010) and the impact of each particular type of TL contact on the development of specific TL performance (Freed, Segalowitz, et al., 2004). For researchers, the existence of detailed models has great importance because it accelerates the development of relevant theories in a given research field (Mislevy, 2009).

Therefore, dividing spontaneous TL contact into four types—reading, listening, writing, and speaking—and confirming the link between initial readiness, the amount of four types of TL contact, and the development of TL oral performance is another way to fulfill the aim of this study (Freed, Segalowitz, et al., 2004). As stated above, this type of investigation is useful when we want to reveal the relationship between initial readiness and the domain-specific TL contact behavior during an SA visit, and the role of each particular type of TL contact in the development of specific TL ability.

The difficulty of this approach lies in the increase in (a) the number of statistical tests

conducted in a single study, and (b) the number of dependent and independent variables included in the model. First, increasing the number of statistical tests increases the chance of making Type 1 errors (Hair, Anderson, Tatham, & Black, 1998; Howell, 2012). Second, splitting spontaneous TL contact into four types (and using them as the predicting variables) simply means that the number of predicting variables included in a model will be four. This is problematic for most SA studies because, as stated in chapter one (see, pp. 26–28), while sample size less than 30 is common in SA studies, even one of the most traditional and loosest rules of thumb of regression analysis states that “the ratio of subjects to predictors be at least 10:1 (Maxwell, 2000, p.434)”. Thus, it is not possible for most SA studies to use this method<sup>3</sup>. For these reasons, thus far only a few quantitative SA studies have ever divided the spontaneous TL contact into four types and examined its relationship with the development of

---

<sup>3</sup> Although there are some possible countermeasures for this problem (e.g., the use of bootstrapping) (Chernick, 2008), most of these methods are not yet widely adopted or known in the field of applied linguistics (Larson-Hall & Herrington, 2010; Plonsky et al., 2015).

TL skills (e.g., Davidson, 2007; Freed, Segalowitz, et al., 2004; Golonka, 2006).

In order to deal with the above-stated issues, the author decided to use both types of investigation methods stated above and answer the following research questions.

1. Is one-semester of SA long enough for SA participants to improve their TL oral proficiency?

2. Do SA participants' initial grammatical skill and international posture jointly impact on

their general spontaneous TL contact behavior? If so, how do these two different aspects

of readiness affect SA participants' spontaneous TL contact behavior and to what extent

do they do this?

3. Does SA participants' general spontaneous TL contact behavior impact on the gains in

their TL oral proficiency?

4. Do initial grammatical skill and international posture impact on SA participants' spontaneous TL reading, listening, writing, and speaking behavior?
  5. Do SA participants' spontaneous TL reading, listening, writing, and speaking behavior impact on the development of TL oral proficiency?
- First of all, this study will confirm whether or not a one-semester SA is long enough for SA participants to improve their TL oral performance (Hernández, 2010; Magnan & Back, 2007; Segalowitz & Freed, 2004). If as a group the participants of this study succeed in improving their TL oral proficiency within a one-semester SA program, then it will be meaningful to investigate how and to what extent the two different types of readiness explain the variance in SA participants' spontaneous TL contact behavior and gain in TL oral performance. On the other hand, if a one-semester SA program is too short in duration to improve TL oral performance, the significance of readiness in individual TL oral development

will also be limited (at least the value of this study for practitioners, e.g., SA administrators, will be decreased as readiness makes little difference in the final outcome of SA learning in this case) (Hernández, 2010).

Research question two and three focus on (a) the contribution of initial readiness to the general spontaneous TL contact behavior in SA learning contexts and (b) that of the general spontaneous TL contact behavior on the gains in TL oral performance. The answers to these questions help us to understand the role of the two types of readiness in the development of TL oral performance in SA learning contexts. The above information provides SA administrators with the basic data for future pre-departure TL trainings.

Research question four and five clarify (a) the impact of initial readiness on the four types of spontaneous TL contact (reading, listening, writing, and speaking) and (b) the relative contribution of each type of TL contact on the development of TL oral performance. The aim of these investigations is to provide future direction for SA studies.

Because research question two and three, and four and five have different goals, hereafter this study treats research questions one to three as study one, and the remainder—research questions four and five—as study two. From the next chapter to chapter seven, this study will deal with research questions one to three; from chapters eight to 11, this study will discuss research questions four and five. Chapter 12 is allocated to general discussion, and the conclusion of this study will be given in chapter 13.

### **3. Method (Study One)**

In this chapter, the author will first outline his sampling method, the characteristics of the participants, and the design of the SA program involved. Following this, he will explain the investigation procedure and the target attribute of each measurement used in this study.

### **3.1. Participants**

The snowball sampling method (Robson, 1993) was used in this study. First, the author personally recruited five JEFL university students. Second, these participants helped the author to identify other potential participants. As a result, 30 JEFL university students were initially recruited for the study. Criteria for retaining participants in the study were as follows (Segalowitz & Freed, 2004): (1) Japanese needed to be their first language; (2) they needed to have started their English learning as one of their school subjects at middle school and continued learning English under the Japanese schooling system<sup>4</sup>; (3) they needed to have never studied English abroad; (4) they must not have taken English as their academic major; (5) there must not be anyone who spoke English as their heritage language; and (6) they had to complete all components described below. Of the original sample, 24 met all criteria, and thus these 24 JEFL university students became the initial sample.

---

<sup>4</sup> Two to three 50-minute classes per week taught by Japanese English teachers.

Following this, the author repeated the snowball sampling method in order to increase the sample size of his study. The author recruited four JEFL university students. Then, as with the last time, these four participants helped the author to identify other potential participants. This time, 27 JEFL university students were initially recruited and 22 met the criteria of this study—these 22 students became the second sample. The number of retained students reached 46 at this point.<sup>5</sup>

This study used the data of all 46 retained students ( $N = 46$ ) ( $M = 19.23$  years old;  $Mdn = 19.16$ ;  $SD = 2.01$ ) (32 females, 14 males). All 46 participants experienced a one-semester SA program somewhere between 2012 and 2013. The difference between the first and second sample was the start dates of their SA programs. The first 24 participants started their SA program earlier than those who were included in the second sample ( $n = 22$ ), and therefore they completed the components of this study earlier than those who departed later. The

---

<sup>5</sup> The overall participant drop-out rate of this study was 19.30% (= 11 / 57).

participants of this study were sampled from six different Japanese universities. As with their home university, the participants' SA destinations were also diverse: the United States ( $n = 16$ ), the United Kingdom ( $n = 12$ ), Australia ( $n = 9$ ), Canada ( $n = 8$ ) and Ireland ( $n = 1$ ).

### **3.2. Design of SA Program**

All SA programs were sponsored programs and their main focus was to improve the participants' general command of English. To this end, taking more than three English as second language classes per week during the SA program was recommended. The class level of the participants was optimized to the intermediate level and depended on their pre-SA language test scores (intermediate:  $n = 8$ , intermediate-low:  $n = 38$ ). In addition, it was possible for the participants to take liberal arts classes at the recipient universities. Except for seven participants who took one communication-based English class as second language class (e.g., discussion class), the participants experienced instruction-based English as second language classes (e.g., grammar, academic writing, presentation, and test preparation class) during the SA. In that sense, the participants had limited opportunity for TL oral practice inside the classroom and had broad chances for TL practice outside the classroom during their stay (DeKeyser, 2007b; Isabelli-García, 2010).

### **3.3. General Procedures and Description of Materials**

Prior to participating in the SA program, the participants first took the TOEFL iBT® to measure their pre-SA English oral performance. Subsequently, they completed a grammatical task and a questionnaire. The questionnaire consisted of two parts: (a) personal information, and (b) motivational indexes. The first part of the questionnaire asked about participants' gender, age, years of schooling, academic majors, prior language experience and overseas experience. The participants' initial international posture was documented in the second part of the questionnaire. During their sojourn, the participants were interviewed twice. The second interview was followed by the completion of the language contact profile. Interviews and completion of the language contact profile were conducted over the Internet. Finally, participants took TOEFL iBT® within six weeks of their homecoming to record their post-SA English oral performance. The instruments used in this study will be described in what follows.

### **3.3.1. Oral performance indexes**

The participants' scores in the speaking section of TOEFL iBT® were applied to pre- and post-SA TL oral performance indexes. This test consists of six tasks: Two familiar topics; two campus situations; and two academic course topics (Educational Testing Service, 2008a, p. 18). Responses to all six tasks are scored by three to six different raters. The response for each task is rated on a scale of 0 to 4 based on delivery, language use and topic development (Educational Testing Service, 2008a, pp. 44-45). The average of all six ratings is converted to a scaled score of 0 to 30 (Educational Testing Service, 2008a, p. 26). The reliability and validity (Haertel, 2006; Messick, 1989) of this test is documented in a number of studies (A. Brown, Iwashita, & McNamara, 2005; Educational Testing Service, 2008b, 2010, 2011; Ockey, Koyama, Setoguchi, & Sun, 2014; Sawaki, Stricker, & Oranje, 2009)<sup>6</sup>.

---

<sup>6</sup> One can access more validity-related documents at:

<https://www.ets.org/toefl/research/topics/validity>

### **3.3.2. Grammatical knowledge index**

Section two of TOEFL® ITP was used for the documentation of grammatical readiness ( $\alpha = .95$ ). This section consists of 40 questions: 15 in the structure section and 25 in the written expression section. Question items were excerpted from a past test (Kanamaru & Educational Testing Service, 2012, pp. 96-102). The participants did 15 fill-in-the-blank tasks in the structure section. In this task, they were given four alternative words or phrases and required to choose one which correctly completed the given sentence. In the written expression section, the participants completed 25 error recognition tasks. In this part, they were required to spot one highlighted word or phrase which contained a grammatical mistake from four alternatives. One can confirm the sample questions of fill-in-the-blank tasks and error recognition tasks in Educational Testing Service (2013). As with the actual test, the time limit was set at 25 minutes for all questions. The participants had not taken the same test elsewhere before the investigation. Each correct answer was converted into 1 point and the

sum consisted of the total score—the minimum score was 0 and the maximum score was 40 in this index. The score on this index represented each individual's capability to refer to his/her knowledge of English grammar (Davidson, 2010; DeKeyser, 2007a, 2010; Golonka, 2006).

Each participant's overall score on this index is shown in Appendix A.

### **3.3.3. International posture**

International posture consisted of 15 items with three subscales ( $\alpha = .86$ )<sup>7</sup>. As stated in chapter one, international posture represents “a tendency to see oneself as connected to the international community, to have concerns for international affairs and a readiness to interact with people other than Japanese” (Yashima & Zenuk-Nishide, 2008, p. 567)”. Question items were excerpted from Yashima (2002) and translated into Japanese. Intergroup approach-avoidance tendency assessed participants’ tendency to approach or avoid non-Japanese in the domestic context (seven items). Interest in international vocation/activities measured participants’ intensity of interest in an international career and

---

<sup>7</sup> In the subscale, this study yielded almost the same level of tolerable alpha values as in prior investigations (display order: the present study; Yashima, 2002): interest in foreign affairs ( $\alpha = .64, .67$ ), interest in international vocation/activities ( $\alpha = .75, .73$ ) and intergroup approach-avoidance tendency ( $\alpha = .76, .79$ ).

living overseas (six items). Interest in foreign affairs reflected participants' interest in international issues (two items). Each item was answered on a 7-point Likert scale. Each participant's total score was converted into a scaled score of 15 to 105. Participants' score on this index is shown in Appendix A.

### **3.3.4. Language contact profile**

Two versions of the questionnaire were employed in this study: in-class and out-of-class language contact profiles. The latter was composed of nine items extracted from the 10-item language contact profile in Hernández (2010). In this questionnaire, the participants answered their weekly action time in each of following out-of-class TL activities: speaking with native or fluent speakers of English; reading e-mail and Internet/web-based content (i.e., the results of net searches or social network use), newspapers, novels and magazines; listening to TV and radio, movie/video, and music with or without English subtitles or lyrics; and writing e-mail, including short messages. Because the amount of these out-of-class TL activities predominantly depends on participants' discretion (Dörnyei et al., 2006; Hyland, 2004), the amount in these responses reflects an estimation of each individual's spontaneous TL contact behavior during the SA. Each participant's action times in general out-of-class TL contact, out-of-class TL reading, out-of-class TL listening, out-of-class TL

writing, and out-of-class TL speaking is cited in Appendix A.

The in-class language contact profile was employed to document the amount of in-class TL contacts during the SA program. This questionnaire consisted of two original sections and one item from Hernández (2010). The original questions examined the content and the amount of classroom instruction the participants experienced during their stay (see Table 2 and Appendix C for the questions). In addition, although the action time in writing homework is affected by one's discretion to some extent, this activity can also be considered as an extension of classroom education. For this reason, the one remaining item in Hernández (2010), writing homework assignments, was added to the in-class language contact profile in this study. Question items were written in Japanese.

### **3.3.5. Interviews**

Semi-structured interviews were conducted twice (3/6 of the SA:  $M = 12.93$  min.;  $Mdn = 12.01$  min.;  $SD = 3.36$ ) (5/6 of the SA:  $M = 15.27$  min.;  $Mdn = 15.11$  min.;  $SD = 2.73$ ). In both investigations, the author asked the participants to explain details of and reasons for their current TL contact (Allen, 2010; Trentman, 2013). The purpose of this investigation was to document the combined impact of two different types of readiness on the participants' learning behavior during the SA program from the qualitative aspect. This study particularly concentrated upon the participants' self-assessed changes in TL competence and personal feeling toward TL contact inside and outside the classroom during the SA program (Allen, 2010; DeKeyser, 2010; Hernández, 2010; Sasaki, 2011). Investigations were carried out in Japanese.

#### **4. Data Analysis (Study One)**

To address three research questions (research question one, two, and three), this study used a paired sample *t*-test, correlation analyses, regression analyses, cluster analysis, Mann-Whitney U tests, and interview analysis. In what follows, this study will explain the usage and details of each analysis. The normality (or lack thereof) of each index/variable was confirmed by the Kolmogorov-Smirnov test. To set the alpha level at .05, the adaptive false discovery rate procedure was employed (Benjamini & Hochberg, 2000). All reported significances in the results section were corrected for the false discovery rate. The descriptive statistics, *p*-values, effect sizes and 95% CIs of each statistical test was computed by SPSS<sup>®</sup> version 17.0 and the web application *langtest* (Mizumoto & Plonsky, 2015; Wilkinson & Task Force on Statistical Inference, 1999). The statistical power ( $1 - \beta$ ) of these tests were computed utilizing *G\*Power 3* (Faul et al., 2009, 2007).

#### **4.1. Research Question One**

A paired sample  $t$ -test was utilized in order to confirm research question one: whether or not a one-semester SA was long enough for SA participants to improve their TL oral performance. In this test, each participant's raw scores on the pre- and post-SA oral performance indexes were used.

#### **4.2. Research Question Two**

Correlation analysis, simultaneous multiple regression analysis, cluster analysis, Mann-Whitney U tests, and interview analysis were employed for research question two. First, this study checked whether or not the predictor and dependent variables showed meaningful correlation (in the field of applied linguistics) (Dörnyei & Ushioda, 2015). The Pearson's correlation was calculated using the standardized scores (i.e., z-scores) of the grammatical knowledge index, international posture, and the general out-of-class TL contact (i.e., the sum of TL reading, listening, writing, and speaking). After checking moderate to strong correlations between (a) the predictor variables (i.e., the readiness indexes) and (b) the dependent variable (the general out-of-class TL contact index), all three variables were entered into the regression model. The standardized scores (i.e., z-scores) of each index were used for the regression analysis.

After testing the statistical link between initial readiness and the amount of general TL

contact, this study explored qualitative support for the results of regression analysis using the interview data. The interview data were utilized for the following reasons. As stated in chapter two, the use of the sum of out-of-class/spontaneous TL contact data does not allow the author to investigate the statistical relationship between initial readiness and the domain-specific TL contact behavior (this exploration is saved for study two). A qualitative exploration is one way to deal with this problem (L. Cohen, Manion, & Morrison, 2007). Such exploration helped the author to discuss the results of study one and two from as many angles as possible (Allen & Herron, 2003; L. Cohen et al., 2007; Richards, 2005). Thus, the author utilized the interview data in his analysis.

To improve the interpretability of interview analysis, this study first divided the participants into two groups using a hierarchical cluster analysis (Ward's method). The participants were classified into the low readiness group (LRG,  $n = 25$ ) and the high readiness group (HRG,  $n = 21$ ) based on their scores on the grammatical knowledge index and

international posture. Mann-Whitney U tests were employed to confirm the appropriateness of this classification (the LRG vs. the HRG in their raw scores on grammatical knowledge index and international posture). In interview data analysis, this study particularly focused on the following three angles: (1) the impact of initial readiness on TL learning experience (2) the influence of TL learning experience on a sense of control over and orientation toward out-of-class TL contact; and (3) the connection between emotional transition toward TL learning and TL contact behavior during the SA program (Bandura & Schunk, 1981; DeKeyser, 2010; Dörnyei, 2001, 2009).

#### **4.3. Research Question Three**

Correlation analysis and single regression analysis were employed to answer research question three. First, the Pearson's correlations between the predictor variable—the general spontaneous out-of-class TL contact (i.e., the sum of TL reading, listening, writing, and speaking)—and its dependent variable (i.e., the gains in TL oral performance) were confirmed (note that the standardized scores (z-scores) of each variable were used for the computation). The correlation between these variables were strong (in the field of applied linguistics) (Dörnyei & Ushioda, 2015), and therefore these variables were entered into the regression model (note that the standardized scores (z-scores) of each variable were used for the computation).

Regarding the dependent variable, the participants of this study started their SA with dissimilar oral scores. As Magnan and Back (2007) stated, SA participants with lower initial proficiency often show more apparent TL gains than those who started their SA with higher

TL levels. This is to say, if the participants who start their SA with lower oral scores show far greater gains in oral score than more proficient participants, it is incorrect for this study to investigate the impact of spontaneous TL contact on development of TL performance using the index labelled as the gains in TL oral performance. Thus, to confirm the usability of this index, this study picked out 37 participants from the final sample based on their degree of oral score improvement and then divided them into two groups: the null gainers and the gainers (cf., Golonka, 2006). The null gainers consist of the participants who improved their oral score less than the average level (i.e., less than 3 points;  $n = 12$ ), and the gainers improved their oral score more than 3 points (i.e., 4 to 7;  $n = 25$ ) (see Appendix A). No problematic difference was found between the null gainers' initial TL oral performance ( $M = 16.00$ ,  $Mdn = 15.00$ ,  $SD = 1.68$ ) and gainers' ( $M = 15.64$ , median = 16.00;  $SD = 1.60$ ) initial TL oral performance.

#### **4.4. In-Class TL Contact Data**

This study concentrated only on the relationship among initial readiness, out-of-class TL contact and development of TL oral performance. This is because, as they participated in a sponsored SA program, the participants' classroom attendance was a reflection of a quasi-fixed school requirement rather than their motivational transitions. For this reason, in this study, in-class TL contact data was used only to describe its qualitative contribution to changes in learning motivation and self-perceived linguistic development, and to confirm the homogeneity in the amount of classroom attendance.

## **5. Results (Study One)**

### **5.1. Summary of Participants' Information**

First, the participants' overall score in the grammatical knowledge index ranged from 12 to 36 on a 40-point scale ( $M = 25.02$ ,  $SD = 4.91$ ,  $Skew = -0.37$ ); and the participants' score in international posture ranged from 50 to 99 on a scale of 15 to 105 ( $M = 79.02$ ,  $SD = 9.65$ ,  $Skew = -0.17$ ).

Second, as Figure 1 shows, the participants' TL oral performance score ranged from 13 to 21 on a 30-point scale ( $M = 15.87$ ,  $SD = 1.74$ ,  $Skew = 0.49$ ) at the pre-SA stage. At this stage, 39 out of 46 participants received a rating of limited level for their TL oral performance, and the remaining seven attained a fair level based on TOEFL iBT® criteria<sup>8</sup>. In the post-SA context, the participants' TL oral performance score ranged from 16 to 24 on the same scale

---

<sup>8</sup> TOEFL® score scales on speaking: Good (26–30); Fair (18–25); Limited (10–17); and Weak (0–9) (Retrieved from <http://www.ets.org/toefl/ibt/scores/understand>).

( $M = 19.57$ ,  $SD = 2.12$ ,  $Skew = 0.15$ ). This time, 38 out of 46 participants attained a fair level,

while eight remained at the limited level.

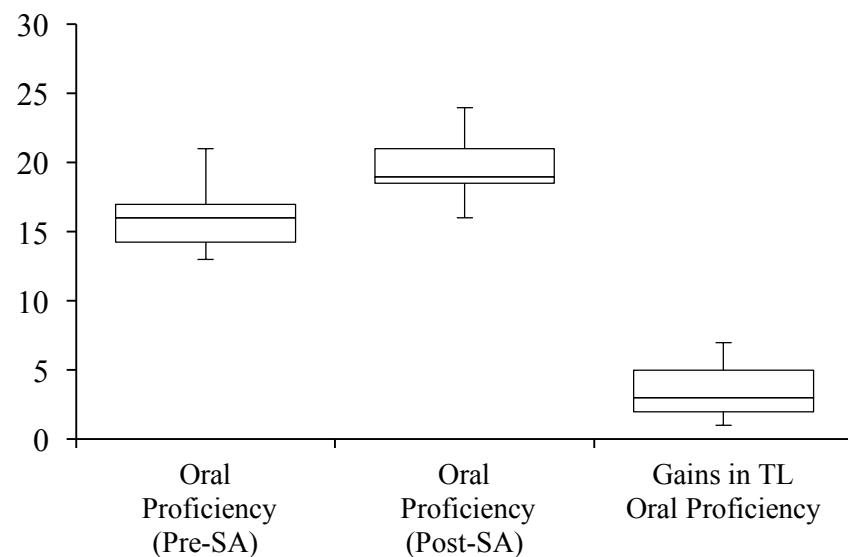


Figure 1. Summary of TL oral proficiency indexes, where the vertical axis indicates scores, the top and bottom of the box are the 75th and 25th percentiles, the line inside the box is the median, and the ends of the whiskers represent the minimum and maximum.

Third, the total hours of participants' weekly out-of-class TL contact ranged from 14.00 to 72.00 hours ( $M = 46.83$  hours,  $SD = 14.07$ , see Table 1 for the breakdown). Fourth, the total hours of participants' weekly in-class TL contact ranged from 15.00 to 39.00 hours ( $M =$

25.54 hours,  $SD = 5.31$ ). Table 2 shows that, on average, the participants took almost seven English as second language classes ( $M = 10.35$  hours,  $SD = 1.54$ ) and two liberal arts classes ( $M = 3.12$  hours,  $SD = 1.54$ ), and engaged in homework assignments for over 10 hours a week ( $M = 11.28$  hours,  $SD = 2.10$ ) during the SA program.

**Table 1**

Descriptive Statistics on the Out-of-Class TL Contact

Item No.	Description of item	<i>M</i>	<i>SD</i>	<i>Skew</i>
1	Speaking English with Native speakers and fluent TL speakers	13.85	7.98	0.40
5	Reading email or Internet content in English	8.08	3.48	0.02
6	Listening to English TV and radio	4.87	3.44	0.89
7	Listening to English movies or videos	5.55	3.96	1.29
8	Listening to English music	3.75	2.10	1.82
2	Reading English newspapers	3.20	1.87	0.96
4	Reading novels in English	2.46	2.03	1.21
10	Writing email in English	2.61	1.97	2.18
3	Reading English language magazines	2.47	2.30	1.92

*Note.* Descriptive statistics were computed using the raw scores.

**Table 2**

Descriptive Statistics on the In-Class TL Contact

Item No.	Description of item	M	SD	Skew
9	Writing homework assignments in English	11.28	2.10	0.85
12	English as Second Language class participation	10.35	1.54	-0.70
11	Liberal arts class participation	3.12	1.54	0.30

*Note.* Descriptive statistics were computed using the raw scores.

Table 3 shows the descriptive statistics on the variables used in regression analyses (research question two and three).

**Table 3**

Descriptive Statistics on the Independent and Dependent Variables

Variable	<i>M</i>	<i>SD</i>	<i>Skew</i>
Grammatical knowledge index	25.02	4.91	-0.37
International posture	79.02	9.65	-0.17
Out-of-class TL contact (General)	46.83	14.07	-0.29
Gains in TL oral performance	3.70	1.63	0.01

*Note.* Descriptive statistics were computed using the raw scores.

## **5.2. Research Question One**

All participants improved their TL oral performance score from one to seven points compared to the pretest scores ( $M = 3.70$ ,  $SD = 1.63$ , see also Figure 1). A paired-samples  $t$ -test was conducted to determine whether the participants statistically improved their TL oral performance compared to the pretest results. The result of the  $t$ -test was statistically meaningful ( $t = 15.24$ ,  $df = 45$ ,  $p < .001$ ,  $1 - \beta > .80$ ,  $d = 1.85$  [95% CI = 1.44, 2.25]). This result indicates that the participants positively developed their TL oral performance during the SA.

### **5.3. Research Question Two**

#### **5.3.1. Simultaneous multiple regression analysis**

To test the relationship between initial readiness and the amount of general out-of-class TL contact during the SA program, multiple regression analysis was conducted. Table 4 shows the Pearson's correlation between the variables: Both the grammatical knowledge index and international posture showed statistically meaningful relationships with the amount of general out-of-class TL contact. Thus, in this model, participants' standardized scores on the grammatical knowledge index and international posture were entered as the predictor variables and their standardized scores on the general TL contact—the sum of four types of TL contact—became the dependent variable.

**Table 4**

Correlation Matrix One (Observed Variables)

Variables	2	3	4
1. Grammatical knowledge index	.54 **	.68 **	.41 **
2. International posture	-	.65 **	.39 **
3. Out-of-Class TL contact (General)		-	.71 **
4. Gains in TL oral performance			-

Note. \*\* =  $p < .01$ , 2-tailed tests.

The regression model was statistically meaningful:  $R^2 = .57$  (adjusted  $R^2 = .55$ ),  $F(2, 43) = 28.97$ ,  $p < .001$ ,  $1 - \beta > .80$ . This result indicates that two different types of initial readiness jointly account for 55% of the variance of the participants' out-of-class TL contact.

Variance inflation factors (VIF) was 1.42 in this model. A VIF of 10.00 or less is considered the adequate range for a lack of multi-collinearity (Hair, Anderson, Tatham, & Black, 1998).

Also, grammatical knowledge ( $\beta = .46$  [95% CI = .22, .70]) and international posture ( $\beta = .40$

[95% CI = .17, .64]) emerged as the predictors of the general out-of-class TL contact behavior (i.e., the beta values and 95% CI estimates of both predictors were adequate).

### 5.3.2. Interview data analysis

As stated above, this study divided the participants into two groups using a hierarchical cluster analysis (Ward's method) so as to improve the interpretability of interview analysis. The participants were classified into the low readiness group (LRG,  $n = 25$ ) and the high readiness group (HRG,  $n = 21$ ) based on their scores in grammatical knowledge index and international posture. As shown in Table 5, there were considerable differences in both aspects of readiness between the two groups prior to participating in the SA program.

**Table 5**

Descriptive Statistics of Two Groups on Readiness and Results of Mann-Whitney U test

Index	LRG			HRG			Mann-Whitney U test			
	<i>M</i>	<i>Mdn</i>	<i>SD</i>	<i>M</i>	<i>Mdn</i>	<i>SD</i>	<i>z</i>	exact <i>p</i>	$1 - \beta$	<i>r</i>
Grammatical knowledge	23.68	24.00	5.39	28.00	29.00	4.66	2.71	.006	> .80	.40
International posture	71.84	74.00	5.77	87.57	88.00	5.47	5.80	< .001	> .80	.86

*Note.* LRG = Low Readiness Group ( $n = 25$ ), HRG = High Readiness Group ( $n = 21$ ).

According to the first interview data, the participants' most serious problem in TL oral performance, particularly during the first stage of their sojourn, was that they could not perform speech production instantaneously and spontaneously. Likewise, at this stage, all the participants also made special mention that they felt difficulty in communicating with experts (i.e., English native speakers and fluent speakers of TL). Under this circumstance, participants could not comprehend the experts' utterances immediately, and therefore they neither responded to an interlocutor in a proper way nor had leeway to pay extra attention to their oral performance.

At the same stage, however, 76.19% (16 out of 21) of highly ready participants first mentioned that, although they had to pay more attention, they were able to understand the general ideas and structures of the experts' utterances. In addition, those 16 participants also asserted that their problem in communicating with the experts did not harm their learning orientation because they were convinced that their current problem in TL comprehension and

speech production was improvable through their own effort. The following excerpt from the first interview with a member of the HRG—Shizuka (no. 5 in Appendix A)<sup>9</sup>—is an example of such comments.

The author: So...you find it difficult to communicate in English.<sup>10</sup>

Shizuka: Yes. But...umm, don't get me wrong. I can communicate with them, and my friends are very supportive of me. But...umm...I need to try hard to follow their conversation and need to try...harder to convey my ideas in English.

[Omission]

The author: Did you try to improve your situation after you realized your problems in English speaking?

---

<sup>9</sup> Note that (a) all participant names provided in this study are pseudonyms and (b) investigations were originally carried out in Japanese (the author translated the conversations between him and the participants into English).

<sup>10</sup> “...” in each excerpt stands for a pause of more than 1 second.

Shizuka: Yes. I...thought that I needed more exposure to English. So I asked my

friends to introduce me to good sources for learning English in my spare time.

The author: Such as?

Shizuka: Dramas, movies...and radio programs. Oh, and I started to make my written

assignments as long as possible, so that I could receive a lot of comments.

[Omission]

The author: Do you think these activities contribute to better English ability?

Shizuka: Yes I think so. Actually, everything does...yes. In comparison, now I can

understand native speakers' utterances so much better than before.

The author: How about your speaking ability?

Shizuka: Umm...I am not so convinced that I have dramatically improved my speaking

ability yet. I am still making a...huge effort when I speak. But...yeah, again, in

comparison, I think my speaking ability is much better than before the SA.

Similar to the case of Shizuka, the remaining 15 HRG members also stated that they increased the amount of out-of-class TL contact on top of classroom participation after they realized their problem in TL ability. Throughout their extra learning attempts, the 16 HRG members not only oriented themselves to speaking activities, but tried to learn practical and new TL knowledge through more self-paced TL activities (i.e., reading and listening).

At the second interview, 80.95% of the HRG (the above 16 plus one of the remaining five participants) judged that their TL oral performance as well as their comprehension ability had improved compared to the midpoint of their SA program. They stated that their investments in TL oral practice had rewarded them in several ways, such as reducing their anxiety about speaking with the experts (11 out of 17), giving them greater self-efficacy in TL speaking (15 out of 17), and resulting in favorable comments about their oral performance from the experts (nine out of 17). These participants reflected that those positive learning

experiences encouraged them to increase their investment in TL learning and practice. Using

their words, they strengthened their motivation to “speak better” (no. 4 in Appendix A)

because the more practice they did, the more “progress” (no. 12 in Appendix A) they could

feel, and the more “improvable aspects” (no. 22 in Appendix A) they realized in their L2 oral

performance. The following excerpt from the second interview with Hikari (no.6 in Appendix

A) briefly summarizes the relationship between their positive learning experience and their

motivation to further self-investment.

Hikari: So...as I mentioned, there were many positive outcomes...and a lot of useful

advice for my speaking ability.

The author: Did these experiences encourage your English learning?

Hikari: Sure. It felt really good to see...good results that allowed me to feel my

progress. These results convinced me that I was on the right track.

The author: Yes.

Hikari: So I did more practice aiming for better English.

In contrast, at the first interview, 48.00% (12 out of 25) of less-prepared participants stated that they could not capture even the general meaning of the experts' utterances during the first stage of their sojourn. Moreover, 60.00 % of less-prepared participants (the above 12 plus three of the remaining 13 participants) asserted that, as they repeatedly failed to understand the experts' utterances and convey their opinions in TL, they gradually lost their willingness to communicate with the experts until halfway through their sojourn. The below excerpts from the first interview with Tomo (no.15 in Appendix A) and Kaoru (no.8 in Appendix A) demonstrate how the lack of readiness demotivated these participants and decreased willingness to speak with the experts in the extramural context.

Tomo: Well...in conversation, I barely understood what they [the experts] said

and...and it made me really upset. Since I couldn't understand, I couldn't respond in  
the right way. Every time things went like that.

The author: How did you feel about that?

Tomo: I was really...well, discouraged by the fact that I couldn't understand them. I  
simply couldn't. I thought that communicating with native speakers was...umm...  
difficult.

The author: Did you try to improve your situation?

Tomo: Umm...not really. In fact, the more conversation I made, the more I realized  
that my English was not good enough.

Kaoru: Speaking with the experts did not help my learning much.

The author: Can you explain to me why?

Kaoru: Well, in the classrooms, I can ask teachers to explain to me the content that I do not understand.

The author: Yes.

Kaoru: But outside the classroom, I cannot ask people to be my teacher. It was so

unpleasant for me constantly not being able to understand what they said and... and

seeing their confused face every time I tried to respond to them.

The first interview further revealed that, even though they realized their problem in TL

ability, the above 12 less-prepared participants were not encouraged to conduct the extra

self-paced TL activities. This is because, in these activities, those 12 participants could not

decode the contents of TL inputs without consulting the Internet constantly. The following

statement of the participants in the LRG (Makoto, no. 1 in Appendix A) at the first interview

is an example of these 12 participants' view of conducting TL contact as their independent

learning.

Makoto: Umm...these activities [reading and listening] take many hours. When I read books...or watch TV, I have to check the Internet so many times to understand the content, what is good or bad or funny about them. That is tiresome, it is not fun.

Nevertheless, at the time of the second interview, even those 12 participants asserted that they gradually increased the number of self-paced TL activities in comparison to the midpoint of their sojourn. They reflected that participation in English as second language classes gave them good opportunities to improve their comprehension ability, so that they became able to learn TL knowledge even outside the classroom. According to the information from the second interview, classroom settings provided these six participants with a good place to learn “practical English knowledge” (no. 20 in Appendix A) and to feel their “progress by completing the curriculum step by step” (no. 2 in Appendix A). As their

self-perceived L2 competence increased, the six participants started to invest more time in extramural reading and listening, and completing homework assignments because they “could understand the meaning better than before” (Tomo) and “extra effort is not as troublesome as it used to be” (Makoto).

On the other hand, although the above-mentioned 60.00% (15 out of 25) of participants also felt gradual growth in their TL oral performance compared to that of the first stage of their SA, they did not overcome their unwillingness to speak with the experts outside the classroom even at the 5/6 stage of their SA program.

Figures 2 and 3 show descriptive statistics of the LRG’s and HRG’s TL contacts inside and outside the classroom (see also Appendix B)<sup>11</sup>. As shown in Figure 2, with regard to the in-class TL contacts, there was no remarkable distributional difference between the two

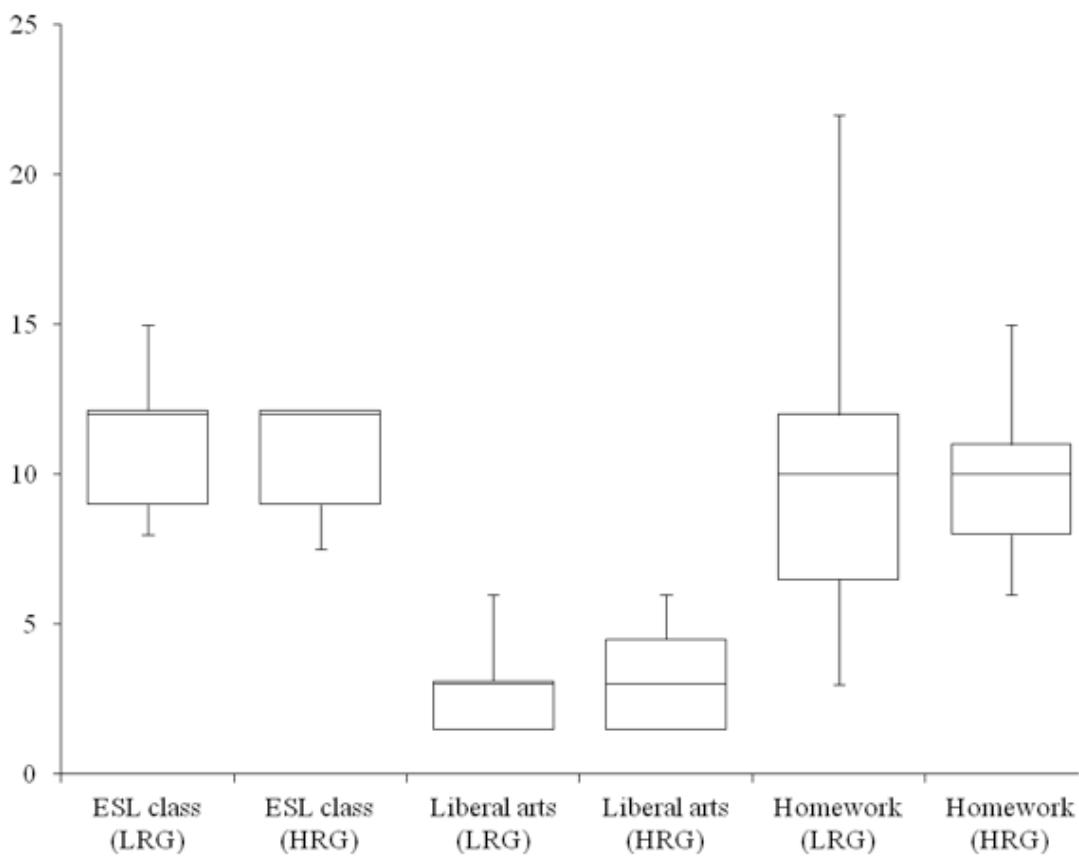
---

<sup>11</sup> In Figure 3 and Appendix B, four reading activities and three listening activities in out-of-class TL contact (see Table 1) were combined as one variable.

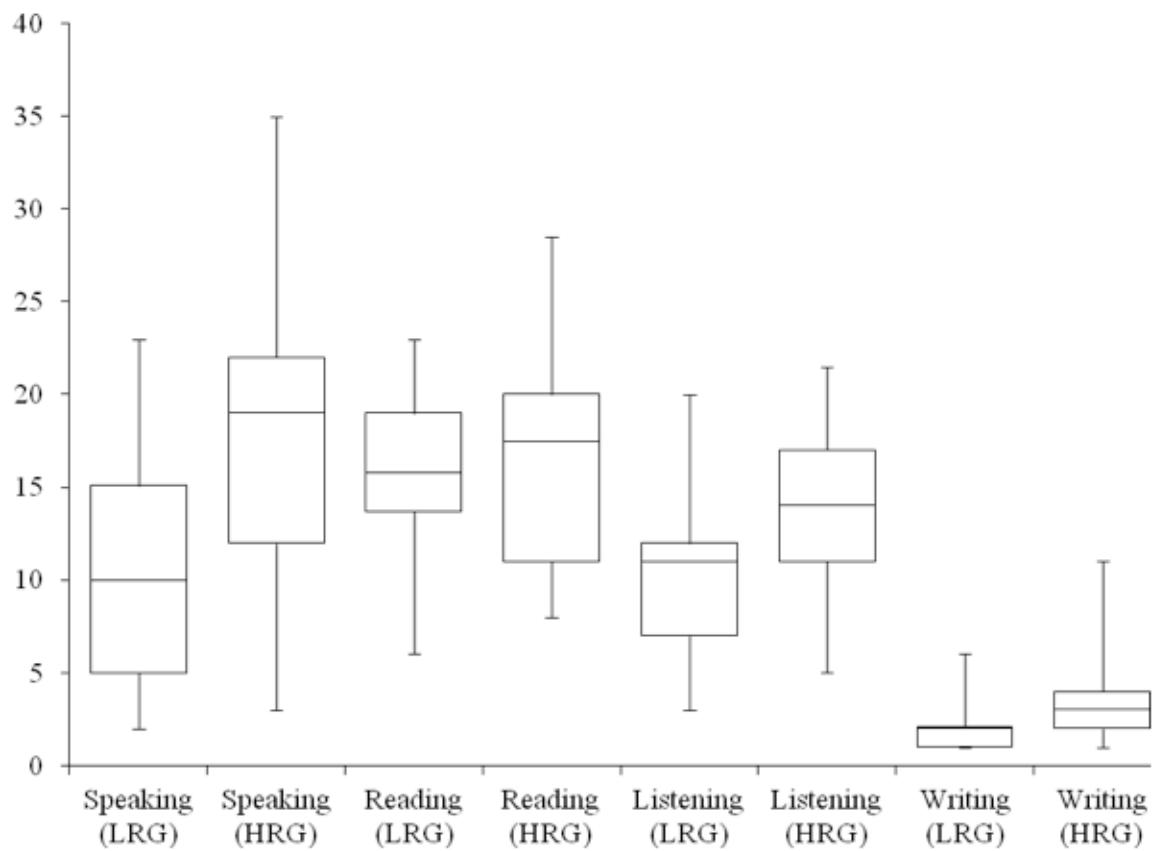
groups (also refer to Appendix B). In Figure 3, there were two distinct distributional differences between the two groups (see also Appendix B). First, in speaking activity, the LRG's 75th percentile (15.00 hours) was close to the 25th percentile of the HRG (12.00 hours), and the LRG's weekly action time ( $M = 10.70$ ,  $CV = 0.59$ )<sup>12</sup> was roughly seven hours shorter than that of the HRG ( $M = 17.60$ ,  $CV = 0.46$ ). Next, in listening activity, the LRG's 75th percentile (12.00 hours) was similar to the 25th percentile of the HRG (11.00 hours), and the LRG's weekly action time ( $M = 9.88$ ,  $CV = 0.41$ ) was about four hours shorter than that of the HRG ( $M = 14.22$ ,  $CV = 0.35$ ).

---

<sup>12</sup>  $CV = \text{Standard Deviation} / \text{Mean}$ .



*Figure 2.* Comparison of two groups' in-class TL contact, LRG = low readiness group ( $n = 25$ ), HRG = high readiness group ( $n = 21$ ), where the vertical axis indicates hours, the top and bottom of the box are the 75th and 25th percentiles, the line inside the box is the median, and the ends of the whiskers represent the minimum and maximum.



*Figure 3.* Comparison of two groups' out-of-class TL contact, LRG = low readiness group ( $n = 25$ ), HRG = high readiness group ( $n = 21$ ), where the vertical axis indicates hours, the top and bottom of the box are the 75th and 25th percentiles, the line inside the box is the median, and the ends of the whiskers represent the minimum and maximum.

#### **5.4. Research Question Three**

Regression analysis was conducted to investigate the relationship between the amount of general out-of-class TL contact and gains in TL oral performance. In this model, the predictor variable was the standardized scores on the general out-of-class TL contact index; the dependent variable was the gains in TL oral performance (= the post-SA test score – pre-SA test score). As shown in Table 4 (p. 66), the Pearson's correlation between these variables was .71 (two-tailed test,  $p < .001$ ). The regression model was statistically meaningful:  $R^2 = .51$ ,  $F(1, 44) = 45.60$ ,  $p < .001$ ,  $1 - \beta > .80$  ( $\beta = .71$  [95% CI = .50, .93]). This result indicates that the amount of general out-of-class TL contact explains 51% of the variance of pretest to posttest gains in TL oral performance.

## **6. Discussion (Study One)**

### **6.1. Research Question One**

The result of the *t*-test showed that the participants succeeded in improving their TL oral performance within a one-semester SA program (see *p*-value and power or 95% CI of the test). This result indicates that at least something existing in the SA contexts had positive impact on the development of participants' TL oral performance. The data provide this study with enough reason to test the role and the contribution of initial readiness and TL contact behavior on the development of TL oral performance.

Furthermore, the result of the *t*-test showed consistency with the findings of prior SA studies in that the participants could improve their linguistic skill within a one-semester SA (Hernández, 2010; Magnan & Back, 2007; Mikami, 2014b; Segalowitz & Freed, 2004). The result of this study, then, adds empirical evidence to the following statement: A one-semester SA program can be a benchmark for TL skill development.

## **6.2. Research Question Two**

As touched on in chapter one, the SA literature hitherto have investigated the individual impact of grammatical and motivational readiness on TL contact behavior and/or development in TL oral performance. For example, grammatical readiness was often considered a factor that affects the efficiency of language learning (Davidson, 2010; Golonka, 2006), and motivational readiness was regarded as a factor related to persistence in TL learning (Allen, 2010; Hernández, 2010; Yashima & Zenuk-Nishide, 2008). The results of this study, however, demonstrated the combined impact of two different types of readiness on SA participants' TL contact behavior during the SA program.

First, the results of simultaneous multiple regression analysis confirmed that the model used for research question two stands (see *p*-value and power of the test). Moreover, both independent variables—the degree of initial grammatical knowledge and international posture—emerged as predictors of the amount of out-of-class TL contact during the SA program.

These results first gave the statistical answer to research question two: Grammatical and motivational readiness indeed affects one's general out-of-class TL contact behavior during the SA program (Allen, 2010; DeKeyser, 2010; Hernández, 2010; Mikami, 2014b). Regarding the impact of initial readiness, adjusted  $R^2$  of this model showed that two different types of readiness jointly explain 55% of the variance in the participants' out-of-class TL contact behavior. Also, the contribution of grammatical knowledge ( $\beta = .46$  [95% CI = .22, .70]) and international posture ( $\beta = .40$  [95% CI = .17, .64]) to the amount of general out-of-class TL contact behavior was equally meaningful (Dörnyei & Ushioda, 2015).

This study conducted interview analysis so as to illustrate how the two different aspects of readiness affect SA participants' general TL contact behavior. In this analysis, the author focused on the participants' self-assessed changes in TL competence and orientation toward TL contact inside and outside the classroom. In the case of the HRG, 76.19% of group members were able to comprehend the general contents and structure of the experts'

utterances from the first stage of their sojourn. This means that to a greater or lesser degree they could utilize their knowledge of grammar within listening activities (cf., Davidson, 2010; DeKeyser, 2010; Golonka, 2006). In addition, even though these participants could not smoothly communicate with the experts during the first stage of their sojourn, this experience did not diminish their positive orientation toward TL oral practice. Instead, they attributed their problems in TL ability to a lack of learning/practice and were oriented toward extra learning in addition to classroom participation. It is correct to say that these participants' reaction to their shortcomings in TL abilities is a reflection of their mental readiness to toil through the TL learning process in order to be a member of English speaking communities (Yashima & Zenuk-Nishide, 2008) and their linguistic readiness to retain a certain sense of control over their TL learning process (DeKeyser, 2010; Dörnyei, 2001; Schunk & Zimmerman, 2007; Zimmerman, 2006). Within their learning behavior, because they were capable of using their declarative knowledge particularly in self-paced/pressure-free TL

activities (i.e., reading, listening and writing) (cf., Ellis, 2009; Krashen, 1985), those participants first succeeded in learning from TL inputs (including feedback) provided in both classroom education and independent learning. Simultaneously, owing to their high aspiration to use TL for communication and desire to improve oral competence, these participants expanded their opportunities to conduct TL oral practice outside the classroom.

On the other hand, in the case of the LRG, in total 60% of the members lost their positive orientation toward out-of-class speaking activity by the midpoint of the SA. In addition, of 25 members, the 12 individuals in the LRG were not oriented toward self-paced out-of-class TL activities either, even at the halfway stage of their sojourn. The reason for their unwillingness to communicate and learn TL outside the classroom was their lack of linguistic and mental backup needed for problem-solving. That is, they were not ready to decode TL inputs using their knowledge of grammar or manage their fear of speaking TL without a scaffold (DeKeyser, 2007b; Krashen, 1985), which they have in classroom learning

contexts.

Although even the above participants in the LRG asserted that they increased the amount of their self-paced TL activities along with their self-perceived growth in TL comprehension ability, descriptive statistics indicated that a half-year SA program was probably not long enough for less-prepared participants to increase their out-of-class TL contact to the level of the HRG.

With regard to the amount of TL input, the distinct difference between the HRG and the LRG was found in listening activities. This disparity most likely reflected the intensity of international posture of each group. The participants in the HRG were better motivated to learn conversation-applicable TL expressions for greater oral competence from the initial point. Moreover, in a series of studies, Robinson reported that grammatically well-prepared language learners try to detect structural regularities of TL input even in the explicit learning condition (cf., Robinson, 2007, p.261). Taking this finding and the two groups' initial

grammatical knowledge level into account, the disparity in the amount of listening activities is largely responsible for the different amounts of newly learned TL knowledge between the HRG and the LRG. Next, in speaking activities, the LRG's weekly action time was on average seven hours shorter than that of the HRG. Considering that declarative knowledge can be transferred to behavioral routine through domain specific practice, the dissimilarity in opportunities for practicing newly learned TL knowledge within speech production is a strong factor leading to the different learning outcomes in TL oral performance. Although the participants of this study also had chances for oral practice inside the classroom, as DeKeyser (2007b) aptly summarized, the amount of TL oral practice that SA participants attain mainly relying on classroom attendance within a limited SA duration is probably insufficient owing to its passive nature. Finally, the LRG caught up with the HRG only in reading activities. The reason for this probably lies in the nature of reading as the most private activity (MacIntyre et al., 1997). Unlike other more public activities, language learners can reread and interpret

written input entirely at their own pace with minimal risk of input misuse and embarrassment (MacIntyre et al., 1997, pp. 279-280), and this is likely to be the reason that the LRG succeeded in attaining positive motivation toward this activity earlier than the other TL activities. In short, it appears that lack of readiness will be particularly problematic for SA participants in learning from spoken TL inputs and conducting sufficient oral practice needed for TL oral development.

### **6.3. Research Question Three**

The above findings provided a good reason to accept the assumption that two different types of readiness jointly affect the amount of general TL contact behavior during an SA program. Hence, the next logical step which this study has to take is to determine whether the amount of general out-of-class TL contact shows a positive connection with gains in TL oral performance. The result of regression analysis revealed that the amount of general out-of-class TL contact was indeed linked to gains in TL oral performance: The amount of general TL contact explains 51% of the variance of pretest to posttest gains in TL oral performance.

The results of the series of regression analyses revealed that grammatical and motivational readiness facilitated knowledge proceduralization and automatization through increasing the amount of general out-of-class TL contact. According to DeKeyser (2007a), knowledge proceduralization and automatization typically results in faster and more accurate

TL speech even under multi-task conditions (see also Skehan, 2002). As we have seen in the

methodology section, those are also the key factors for increasing the overall score in the

speaking section of the TOEFL iBT® rating: its evaluation is based on the delivery, language

use and topic development of given speech (Educational Testing Service, 2008a, pp. 44-45).

This is to say, increasing limited aspects of TL oral performance cannot lead to an increase in

overall score. Given the above, it is safe to accept the following conceptual flow: The more

initial readiness one has at a pre-SA stage, the more opportunities for TL knowledge

proceduralization and automatization one attains during his/her SA visit (Allen, 2010;

Davidson, 2010; DeKeyser, 2010; Golonka, 2006; Hernández, 2010; Yashima &

Zenuk-Nishide, 2008).

## **7. Educational Implications (Study One)**

As demonstrated above, if SA participants equip themselves with the capability to refer to their knowledge of grammar rules in TL comprehension, and even production to some extent, this not only increases the efficiency of independent language learning/practice outside of classroom education, but also reinforces SA participants' sense of control over their learning process. In addition, provided that SA participants have strong aspirations for TL communication as part of a high level of international posture, they can not only protect their positive orientation toward TL learning/practice, but also effectively expand the grounds for knowledge proceduralization, the initial step of knowledge automatization. Moreover, with readiness, SA participants can avoid merely improving limited aspects of oral performance, such as politeness formulas, routine requests and routine inquiries (DeKeyser, 2007b, p.213)—again, TOEFL iBT® rating is based on the delivery, language use and topic development of given speech (Educational Testing Service, 2008a, pp. 44-45). Taking these results, this study

suggests the following recommendations for SA administrators (Mikami, 2014b, 2014d).

First, in Yashima and Zenuk-Nishide (2008), Japanese high school students who were taught a content-based language learning program developed their international posture along with their TOEFL® ITP score and frequency of TL communication inside and outside the classroom. This result indicated that if prospective SA participants receive the opportunities to be taught international content in their target language or chances to convey their ideas in their TL within pre-departure training, they will realize what they can/want to do using TL and become aware of the gap between their objective and their actual TL level (Keck, Iberri-Shea, Tracy-Ventura, & Wa-Mbaleka, 2006). For instance, imagine that a lecturer who uses the same mother tongue as prospective SA participants and has a good command of TL conducts content-based pre-departure SA classes. If this lecturer uses international content such as TV shows, news, music and cultural programs as teaching materials and conversation topics, prospective SA participants can relate to international affairs and issues, and will be

able to have a tangible image of their ideal L2 level through practice and interaction with the lecturer. As Dörnyei (2001) suggested, communicating with achievers who share similar learning backgrounds with language learners can foster their positive belief in the TL learning process (i.e., observation and modeling) (see also Bandura & Schunk, 1981; Zimmerman, 2006).

In addition, after prospective SA participants have realized the gap between their objective level and their actual TL level, it is also important to nurture their ability to use their knowledge of grammar within TL activities. For this, it would be worth trying to record the TL speech production of prospective SA participants and giving them explicit feedback on their outputs. By being given correct forms of the TL expressions they wanted to convey, they will be able to increase their knowledge of grammar in the relevant domain. Similarly, it may also be productive to listen to TL broadcasts and write down the form of those inputs using fill-in-the-blank tasks, for example (cf., DeKeyser, 2010). Experiencing this type of training

first contributes to accustoming prospective SA participants to spoken input, and can increase

the referable grammar rules within their listening activities (cf., Davidson, 2010).

## **8. Method (Study Two)**

The aim of study two is to illustrate the impact of initial readiness on the particular types of out-of-class TL contact—reading, listening, writing, and speaking—(research question four) and the relative contribution of the amount of these TL contacts on the gains in TL oral performance (research question five) (Freed, Segalowitz, et al., 2004). To fulfill this aim, the author divided the general out-of-class TL contact into reading, listening, writing, and speaking activities, and reanalyzed the link between initial readiness and out-of-class TL contact behavior, and that between out-of-class TL contact behavior and the gains in TL oral performance. The data gleaned from the same participants and instruments introduced in chapter three were utilized in these analyses.

## **9. Data Analysis (Study Two)**

### **9.1. Research Question Four**

Correlation analysis and simultaneous multiple regression analyses were employed to answer research question four. This study first checked whether or not the predictor and dependent variables showed meaningful correlation (J. Cohen, 1992; Dörnyei, 2007; Field, 2005). The Pearson's correlation was calculated using the standardized scores (z-scores) of the grammatical knowledge index, international posture, and the four types of out-of-class TL contacts (i.e., reading, listening, writing, and speaking). To minimize the risk of type-1 error, the author ran regression analysis only when he confirmed the meaningful correlation between at least one of two predictor variables (i.e., grammatical knowledge index and international posture indexes) and the candidates for the dependent variable (i.e., out-of-class TL reading or listening or writing or speaking). The standardized scores (i.e., z-scores) of each index were used for the regression analyses.

## **9.2. Research Question Five**

Correlation analysis and simultaneous multiple regression analysis were employed to answer research question five. First, the Pearson's correlations between the predictor variable—the four types of out-of-class TL contacts (reading, listening, writing, and speaking)—and the dependent variable (i.e., the gains in TL oral performance) were confirmed using the standardized scores (z-scores) of each variable. Of four candidates (for the predictor variable), only the variables that had meaningful correlation with the gains in TL oral performance (i.e., the dependent variable) were entered into the regression model. The standardized scores (i.e., z-scores) of each index was used for the regression analysis.

## **10. Results (Study Two)**

### **10.1. Summary of Participants' Information**

The difference between study one and two is whether the author treats the out-of-class TL contact as the single variable (study one) or the four variables (study two). Thus, the only new information that should be provided in this section is the descriptive statistics on the four types of out-of-class TL contact (see Table 6). As one might notice, the average length of out-of-class TL writing activity is much smaller than that of the other three types of out-of-class TL contact (reading, listening, and speaking activities). As explained in the methodology section (in particular, 3.3.4), this is simply because this study categorized writing home-work assignment as in-class TL contact. As the standardized scores (z-scores) were used for computation of the following regression analyses, such difference does not affect the results of this study.

**Table 6**

Descriptive Statistics on the Independent and Dependent Variables

Variable	<i>M</i>	<i>SD</i>	<i>Skew</i>
Grammatical knowledge index	25.02	4.91	-0.37
International posture	79.02	9.65	-0.17
Out-of-class TL contact (Reading)	16.21	4.70	0.05
Out-of-class TL contact (Listening)	14.17	6.55	0.26
Out-of-class TL contact (Writing)	2.61	1.97	2.18
Out-of-class TL contact (Speaking)	13.85	7.98	0.40
Gains in TL oral performance	3.70	1.63	0.01

*Note.* Descriptive statistics were computed using the raw scores.

## **10.2. Research Question Four**

As indicated in Table 7, the out-of-class TL reading contact index had no meaningful correlation<sup>13</sup> with the predictor variables (i.e., the grammatical knowledge index and international posture). Thus, the out-of-class TL reading contact index was excluded from further analyses.

---

<sup>13</sup> In particular, no meaningful correlation in the field of applied linguistics (Dörnyei & Ushioda, 2015).

**Table 7**

Correlation Matrix Two (Observed Variables)

Variables	2	3	4	5	6	7
1. Grammatical knowledge index	.54 **	.09	.56 **	.37 *	.59 **	.41 **
2. International posture	-	.17	.51 **	.43 **	.53 **	.39 **
3. Out-of-class TL contact (Reading)		-	-.12	-.17	.02	.18
4. Out-of-class TL contact (Listening)			-	.32 *	.51 **	.43 **
5. Out-of-class TL contact (Writing)				-	.45 **	.26
6. Out-of-class TL contact (Speaking)					-	.74 **
7. Gains in TL oral performance						-

Note. \* =  $p < .05$ , \*\* =  $p < .01$ , 2-tailed tests.

The relationship between initial readiness and the remaining three types of TL contact

indexes (i.e., listening, writing, and speaking) was tested using multiple regression analyses.

In each analysis, the grammatical knowledge index and international posture were entered as

the predictor variables; and the dependent variable was one of the three types of TL contact

index (note that standardized scores were used for the computations).

The results of regression analyses are shown in Table 8. The VIF of each model was

1.42 (Hair et al., 1998). As shown in Table 8, of three regression models, two of them were

statistically meaningful (in terms of their *p*-values and statistical powers): Initial readiness

jointly predicted 34% and 38% of the variance of the out-of-class TL listening and speaking.

It also became clear, however, that this time the predicting power of grammatical knowledge

index was clearly stronger than that of international posture in both models (see Table 8). In

particular, while the beta value of grammatical knowledge index was .40 [95% CI = .11, .69]

in relation to the out-of-class TL speaking, that of international posture was .29 [95% CI

= .01, .58]; and when it comes to out-of-class TL listening, the beta value of grammatical

knowledge index was .43 [95% CI = .15, .71] and that of international posture was .30 [95%

CI = .01, .58].

**Table 8**

Results of Multiple Regression Analyses

Dependent variable	<i>F</i>	<i>p</i>	power	$R^2$ (Adjusted $R^2$ )	$\beta$		95% CI of $\beta$	
					GKI	IP	GKI	IP
Listening	12.81	< .001	> .80	.37 (.34)	.40	.29	[ .11 , .69]	[ .01 , .58]
Writing	5.80	.006	.70	.21 (.18)	.19	.33	[-.14 , .51]	[ .01 , .66]
Speaking	14.76	< .001	> .80	.41 (.38)	.43	.30	[ .15 , .71]	[ .01 , .58]

*Note.*  $df=2, 43$  for each model, GKI = grammatical knowledge index, IP = international posture.

### **10.3. Research Question Five**

Table 7 shows that there was no meaningful correlation between the out-of-class TL reading index and the gains in TL oral performance; the same goes for the relationship between the out-of-class TL writing index and the gains in TL oral performance. For these reasons, the out-of-class TL reading and writing indexes were excluded from the further data analysis<sup>14</sup>. The relationship between the remaining two out-of-class TL contact indexes—listening and speaking—and the gains in TL oral performance was tested by multiple regression analysis. The predictor variables were the standardized scores (z-scores) on the two types of out-of-class TL contact indexes (i.e., listening and speaking) and the dependent variable was the standardized scores on the gains in TL oral performance. The regression model was statistically meaningful:  $R^2 = .55$  (adjusted  $R^2 = .53$ ),  $F(2, 43) = 26.28, p < .001$ ,

---

<sup>14</sup> It was checked that these variables, the out-of-class TL reading and writing indexes, would not act as suppression variables even if they were entered to the simultaneous multiple regression model as this study initially planned.

$1 - \beta > .80$ . The VIF of each variable was 0.74. While the out-of-class TL speaking index

emerged as the predictor of the gains in TL oral performance in terms of the adequate beta

value and its 95% CI estimates ( $\beta = .70$  [95% CI = .46, .94]), the impact of out-of-class TL

listening activity on the linguistic gain was inconclusive (i.e., the beta value was not only

small but also included zero at the 95% CI level) ( $\beta = .07$  [95% CI = -.17, .31]).

## **11. Discussion (Study Two)**

The purpose of study two was to confirm (a) the association between initial readiness

and the amount of four types of out-of-class TL contact (i.e., reading, listening, writing, and

speaking), and (b) that between the specific out-of-class TL contact behavior and the gains in

TL oral performance. The results of correlation analysis and the series of regression analyses

first showed that grammatical and motivational readiness played a limited role in the

individual differences in the pursuit or avoidance of out-of-class TL reading and writing

activities, while these readiness factors jointly explain 34% and 38% of the variance in the

time spent on out-of-class TL listening and speaking activities (a week). Second, it was also

found that, of four out-of-class TL contact indexes, only out-of-class TL speaking functions as

a meaningful predictor of gains in TL oral performance.

In this chapter, the author will first discuss the alterations or modifications that should

be made in future model building so that future SA studies can capture a cleaner relationship

between initial readiness and the four types of out-of-class TL contact. Following this, the relationship between the amount of out-of-class TL contacts and the gains in TL oral performance will be discussed referring to the results of the correlation analysis and those of the series of regression analyses.

### **11.1. Research Question Four**

The relationship between initial readiness and the amount of out-of-class TL reading and writing were statistically inconclusive. In particular, first, the amount of out-of-class TL reading had no meaningful correlation with its predictor variables—the grammatical knowledge index and international posture. Second, the model including the amount of out-of-class TL writing as the dependent variable was underpowered (i.e.,  $1 - \beta \leq .80$ ) (J. Cohen, 1992). On top of this, the influence of grammatical knowledge on the amount of out-of-class TL writing could be either negative or positive at the 95% CI level ( $\beta = .19$  [95% CI =  $-.14, .51$ ]. With these data in mind, in what follows the author will explore the possible attributions of these inconclusive results.

Hypothesizing the weak relationship between the grammatical and motivational readiness and the spontaneous learning effort made in the domain of out-of-class TL reading and writing is the first option. When it comes to the amount of out-of-class TL reading, this

hypothesis fits to the data shown in Figure 3 (see p.81). In this Figure, the behavior of the LRG and HRG indicates that, at the 5/6 stage of their SA visit, both groups were engaging in a similar amount of out-of-class TL reading regardless of their initial readiness. To put it another way, in Figure 3, initial readiness did not function as a determining factor in the amount of self-investment in the domain of TL reading.

Provided that the data shown in Figure 3 reflect the true relationship (in the population), then, what type of countermeasures can be taken in future SA investigations? Judging from the recent findings of language learning motivation studies, replacing international posture with TL reading-specific motivational indexes seems to be a realistic approach (Kondo-Brown, 2006; Mikami et al., 2016; Takase, 2007; Yamashita, 2013). In language learning research, more and more researchers accept the domain specialty of psychological attributes related to TL learning, including motivation (Cheng, 2004; Grabe, 2009; MacIntyre et al., 1997; Saito, Garza, & Horwitz, 1999). This is to say, the development of specific TL

skills (e.g., TL reading skill) will be facilitated or inhibited by the existence (or lack thereof) of domain-specific psychological attributes (e.g., TL reading motivation). For instance, some prior studies have revealed that intrinsic self-value toward FL reading and self-efficacy in FL reading are significant predictors of TL reading skill development (Kondo-Brown, 2006; Matsui & Noro, 2010; Yamashita, 2013). In motivation studies, intended learning effort is considered to be the key that connects learning motivation with skill developments (Bandura & Schunk, 1981; Dörnyei, 2009; Mikami et al., 2016; Ryan & Deci, 2000; Zimmerman, 2006). Thus, the findings of these studies support the idea that future SA studies will find stronger association between motivational readiness and the amount of out-of-class TL reading when they employ the index specialized to measuring TL reading motivation instead of international posture.

The method of measuring the amount of out-of-class TL reading should also be the subject of reconsideration. As stated in the research methodology section, in this study, the

time spent on spontaneous TL reading per week was treated as the amount of out-of-class TL reading, i.e., the author followed the framework of Freed, Dewey, et al. (2004). Regarding this point, however, some prior second and foreign language reading studies showed that the volume of TL reading (i.e., the number of words or pages that each individual has read in a certain timeframe) can also function as a measurement of the amount of spontaneous TL reading (Fujita & Noro, 2009; Matsui & Noro, 2010; Takase, 2007). As an example, Takase (2007) reported that intrinsic self-value toward FL reading was, among other factors, the best predictor of the volume of spontaneous TL reading. The above data back up the criterion-related validity of the volume of TL reading and thus support the practicality of this index. One thing we should bear in mind is, however, that it is not yet definitive whether volume of spontaneous TL reading always functions as a better measurement compared to time spent on spontaneous TL reading. For instance, in Tanemura (2014), it was documented that the learners' belief indexes showed a stronger relationship with the time spent on

spontaneous TL reading than with the volume of reading. When we combine the findings of

Takase (2007) and Tanemura (2014), it is reasonable to think that the different type of

psychological attribute affects the different facets of learning effort spent on TL reading.

Theoretically speaking, of two types of readiness, grammatical knowledge can have a stronger

relationship with the volume of TL reading because this type of readiness is supposed to make

a difference to the efficiency of learning or the sense of control over TL reading (Brecht et al.,

1995; Davidson, 2010; Jeon & Yamashita, 2014) (support for this claim can also be seen in

the interview data, see 5.3.2). On the other hand, the time spent on TL reading activity may be

more strongly connected to motivational readiness. As an example, if SA participants equip

themselves with high degree of intrinsic self-value toward FL reading at a pre-SA stage (i.e.,

they read TL for enjoyment) (Kondo-Brown, 2006; Mikami et al., 2016), they may allocate a

number of hours to TL reading activity regardless of learning efficiency. Where this is the

case, the amount of time spent on out-of-class TL reading will show stronger connection with

motivational readiness than grammatical readiness. Considering these potentially more productive research scenarios, there is little reason for future SA studies not to treat both the time and volume of out-of-class TL reading as measurements of the amount of out-of-class TL reading and then test the association between initial readiness and these variables.

The first possibility—that grammatical knowledge and international posture at the pre-SA stage have little impact on the amount of out-of-class TL reading and writing—does not hold water for writing activity, however. As we can see in Figure 3 (p.81) and Appendix B (p.178), here the learner groups with higher readiness (i.e., the HRG) seemingly made more self-investment in out-of-class TL writing compared with the less readied group (i.e., the LRG). In other words, with regard to the amount of out-of-class TL writing, readiness seemingly made a difference to participants' behavior. These data lead the author to the second possibility, that is, the sample size of this study was insufficient to yield significance, adequate power, or non-zero 95% CI when it comes to out-of-class TL reading or writing.

This hypothesis fits particularly well with the data on out-of-class TL writing. A priori analysis of *G\*power 3* (Faul et al., 2009, 2007) indicated that, provided that the other conditions (e.g., effect size and *p*-value) remain the same, 57 participants will be required to yield sufficient power from the regression model that includes the amount of out-of-class TL writing. Although collecting almost 60 participants for a single study is not a common practice in SA studies (see pp.25–28), this sample size is by no means unattainable. Thus, it is safe to say that, at least in the case of out-of-class TL writing, it is too early to accept its weak relationship with the target readiness of this study—it is a realistic scenario that the statistically meaningful relationship between initial readiness and the amount of out-of-class TL writing will emerge if we sample more than 57 participants.

On the contrary, it seems unrealistic to confirm the relationship between initial readiness and the amount of out-of-class TL reading by simply increasing sample size. A priori analysis showed that, provided that the other conditions remain the same, more than

485 participants will be needed to attain adequate statistical power from the regression model including the amount of out-of-class TL reading. Unlike the case of out-of-class TL writing, yielding more than 480 participants will be a huge challenge for SA studies. Thus, when it comes to out-of-class TL reading, replacing the motivational readiness index and adding the volume of TL reading to the measurement of the amount of TL reading would be a realistic approach for future SA studies.

The last possible reason for the inconclusive relationship between initial readiness and the amount of out-of-class TL reading can be found in the interview data. Here, 16 out of 21 (76.19%) SA participants with high readiness tried to learn practical and new TL knowledge through out-of-class TL reading at the 3/5 stage of their SA. On the other hand, the 12 less readied SA participants (48.00% of the less readied group) were not encouraged to conduct such self-paced out-of-class TL activities as their lack of readiness did not allow them to decode the contents of TL inputs without constantly consulting the Internet. It was the 5/6

stage of their SA when these 12 participants stated that they tried to read, listen or write TL more outside the classroom. As one might notice, the timing where almost 50% of the less readied SA participants started to increase their out-of-class learning effort corresponded to when they responded to the language contact profile (see 3.3). It is plausible to think, then, that the less prepared participants succeeded in boosting the amount of their out-of-class TL reading to the level of more readied SA participants somewhere between the 3/6 and 5/6 stage of their SA visit<sup>15</sup>. To deal with this problem or confirm the plausibility of this hypothesis, future SA studies need to conduct the language contact profile survey more than once (N. Taguchi, 2008) and illustrate the transition in the amount of out-of-class TL contact (see, e.g., Sasaki, 2011). When we compare a situation in which researchers conduct the language contact profile survey multiple times with a large sample investigation requiring at least 485

---

<sup>15</sup> As stated in the discussion of research question two, the least anxiety provoking nature of TL reading may have something to do with this catch-up (MacIntyre et al., 1997; Mikami et al., 2016).

participants, it is clear that the former option is more realistic for most researchers.

To sum up, regarding the relationship between the grammatical and motivational readiness and the amount of out-of-class TL reading and writing, this study recommends future SA studies to (a) alter the measurement of motivational readiness, (b) add the volume (words and pages) that SA participants have read in a certain timeframe to the measurement of the amount of out-of-class TL reading, (c) set multiple data collection points (e.g., 2/6, 4/6, and 6/6 of the SA visit), and, if possible, expand their sample size up to 60.

Unlike the amounts of out-of-class TL reading or writing, that of out-of-class TL listening and speaking showed a statistically meaningful relationship<sup>16</sup> with both forms of initial readiness. First, initial readiness—grammatical knowledge and international posture—jointly explained 34% of variance of the participants' out-of-class TL listening behavior. This result was not only statistically meaningful (J. Cohen, 1992; Dörnyei & Ushioda, 2015), but

---

<sup>16</sup> In terms of their *p*-value, statistical power and 95% CI.

was also consistent with the data shown in Figure 3 (p.81) in that initial readiness makes a

difference to SA participants' spontaneous TL learning behavior in the domain of listening.

The same findings applied to the relationship between initial readiness and the amount of

out-of-class TL speaking (adjusted  $R^2$  was .38 for this model).

However, the predicting power of international posture on the amount of out-of-class

TL listening and speaking becomes much weaker compared to the situation where we used

the amount of general out-of-class TL contact as its dependent variable. In particular, 95% CI

of  $\beta$  fell between .01 and .58 for both out-of-class TL listening and speaking whereas the

estimate was from .17 to .64 for the amount of general out-of-class TL contact. Although the

lower limits of the former does not include zero, the author must note that .01 was extremely

marginal. This is to say, it is not a remote possibility that future SA studies (or even perfect

replication studies) would fail to find a link between international posture and the amount of

out-of-class TL listening and/or speaking (Cumming, 2013).

What could these results possibly be attributed to? One potential explanation lies in the functional difference between general TL learning motivation and domain-specific TL leaning motivation. Language learning motivation studies have shown that the degree (or the type) of general TL learning motivation has a strong relationship with the intensity of general learning effort that one makes (Dörnyei, 2001, 2009; Masgoret & Gardner, 2003). For instance, as introduced in chapter one (pp.20-21), Hernández (2010) showed that the degree of integrated motivation index was a predictor of the amount of general out-of-class TL contact. Here, general out-of-class TL contact is synonymous with the domain-general TL learning effort in that the sum of the four types of out-of-class TL activity makes up the index. Because this index represents the domain-general TL learning effort that one made in a certain timeframe, it is durable against the individual differences in personal preference for TL activity. As an example, 10 hours of out-of-class TL reading and 10 hours of out-of-class TL speaking are treated as the same 10 hours of TL contact behavior. When considered this way, the stronger

association between international posture and the amount of general out-of-class TL contact

ceases to be so puzzling as the characteristics of international posture are closer to

domain-general learning motivation than domain-specific motivation (Kormos & Csizér,

2008; Yashima & Zenuk-Nishide, 2008; Yashima, 2002). Then, in order to capture a clearer

association between TL learning motivation and domain-specific TL contact behavior, we

should, just as in the case of out-of-class TL reading, replace international posture, the

domain-general motivational index, with more domain-specific TL learning indexes.

Regarding the 95% CIs of the grammatical knowledge index, it was confirmed that

this index functions as the factor explaining variance in the amount of out-of-class learning

behavior in listening and speaking (see Table 8, p.105). In particular, for every 1 standard

deviation increase in the grammatical knowledge index, there will be .40 standard deviation

increase in the amount of out-of-class TL listening (the value will fall somewhere between .11

and .69 at the 95% CI level) and .43 standard deviation increase in the amount of out-of-class

TL speaking [95% CI = .15, .71]. These results added the valuable data to the findings of SA

studies for the following reasons. Prior SA studies only assumed that grammatical readiness

makes changes in the efficiency of or the amount of effort made in SA learning (Brecht et al.,

1995; Davidson, 2010; Golonka, 2006); in this study, however, it was confirmed that

grammatical readiness does impact on the amount of learning effort made in the domain of

out-of-class TL listening and speaking.

## **11.2. Research Question Five**

As shown in the last chapter, of two independent variables, the out-of-class TL speaking index emerged as the only predictor of the pretest to posttest gains in TL oral performance ( $\beta = .70$  [95% CI = .46, .94]). This effect size (and its 95% CI) is considerable as it suggests that, for every 1 standard deviation increase in the amount of out-of-class TL speaking, there will be .70 standard deviation increase in the pretest to posttest gains in TL oral performance amount (the value will fall somewhere between .46 to .94 at the 95% CI level). To understand this surprising result, we need to recall the premise of the skill acquisition theory. That is, initial and/or repeated domain specific practice is a crucial factor in the behavioral routinization and eventual automatization of the specific (TL) performance (Anderson & Lebiere, 1998; DeKeyser, 2007a; Segalowitz, 2003). When this study applies this premise to its 46 participants, the role of out-of-class TL speaking in the development of TL oral performance will be clear because, among the four types of out-of-class TL contact

indexes, only the out-of-class TL speaking contact provides direct opportunity for TL oral practice.

On the other hand, despite its adequate correlation (see Table 7, p.103) (Dörnyei & Ushioda, 2015), the impact of out-of-class TL listening on the gains in TL oral development was inconclusive ( $\beta = .07$  [95% CI = -.17, .31]). Although it is possible for future studies to attain narrower (or non-zero) 95% CI with larger sample size (e.g.,  $N = 60$ , see, p. 116), from the point estimation of beta coefficient (i.e.,  $\beta = .07$ ) we can judge that the association between the amount of out-of-class TL listening and the gains in TL oral performance was weak (J. Cohen, 1992). It will be, therefore, beneficial for this study (and future SA studies) to consider a theoretical explanation for this result before blindly increasing sample size.

What we can firstly hypothesize is the indirect contribution of out-of-class TL listening to the development of TL oral performance. As we can see in Table 1 (p.61), the question items regarding out-of-class TL listening only include self-learning activities, that is,

watching films and television, and listening to the radio and music in TL. Through these activities, SA participants can learn new vocabulary, grammatical rules, and formulaic and creative expressions applicable to their spontaneous TL speech outside the classroom (the out-of-class TL speaking in this study represents TL conversations with native and fluent TL speakers). Yet, until SA participants test the appropriateness or accuracy of their TL knowledge in an actual conversation, it remains unclear whether or not these newly attained vocabularies, grammar rules or expressions correctly convey their mental representations or not in a given situation (Mikami, 2014b). If SA participants' attempts result in positive reactions or feedback—meaning, their speech made sense to their interlocutors or they receive some positive reaction or feedback for their speech—then they will strengthen their confidence in the way that they apply their newly attained TL knowledge to their TL speech. Even if SA participants' attempted speech failed to yield positive results, they would realize that they need different TL knowledge from what they already have in order to correctly

convey their mental representations in a given situation.

Also, as N. Taguchi (2008) pointed out, it is important to take the frequency of occurrence of TL inputs into account in the TL knowledge proceduralization. The proceduralization of some words, grammar rules or expressions takes longer time than others because some language uses occur more frequently than others even in SA contexts. As an example, it will be relatively easy for SA participants to master how to greet (as the occurrence of greetings is frequent and these are typical formulaic expressions) (DeKeyser, 2007b) compared to the mastery of less frequent expressions, such as indirect refusals and irony (N. Taguchi, 2008). When considered this way, it is natural to consider the amount of out-of-class TL listening to be of importance for more effective TL oral practice as increasing TL oral input allows SA participants to encounter larger amounts (and probably wider varieties) of testable TL knowledge in their TL oral practice. In that sense, even though a greater amount of out-of-class TL listening provides SA participants with further oral TL

input, it is difficult for SA participants to enhance their general TL oral performance (i.e., the delivery, language use and topic development of their TL speech) without testing their newly attained TL knowledge in real-life interactions (i.e., oral practice). Support for this hypothesis can be found in the interview data (see 5.3.2) and the adequately strong correlation between the amount of out-of-class TL listening and speaking ( $r = .51$ , see Table 7, p.103) (Dörnyei, 2007). Thus, it is possible that out-of-class TL listening indirectly contributes to the development of TL oral performance.

When the author takes the above hypothesis and the interview data into account, there is little reason not to apply the above relationship (the indirect contribution of passive TL contact to the development of TL oral performance) to out-of-class TL reading. Theoretically speaking, just as with out-of-class TL listening, out-of-class TL reading provides SA participants with new and testable TL knowledge; SA participants can test their understanding in actual conversations; they will receive some kind of feedback for their attempts; it helps

SA participants to update, proceduralize, or automatize their TL knowledge in the domain of TL speaking. In that sense, it is also possible to hypothesize the indirect contribution of out-of-class TL reading for the development of TL oral performance.

One obvious difference between out-of-class TL listening and reading is the form of TL input—the former gives SA participants oral TL input and the latter provides them with written TL input. This may explain why the correlation between out-of-class TL reading and speaking is weaker than that between out-of-class TL listening and speaking ( $r = .02$  vs.  $r = .51$ , see Table 7, p.103). SA participants who wish to enhance their TL oral performance may invest more learning effort to TL listening as both listening and speaking are conducted through the same medium (i.e., speech sound). As touched on in the last section, however, the quantifying method of out-of-class TL reading should be the subject of reconsideration. Until the methodology is reconsidered and results are confirmed, it is not possible to conclude which type of input (written vs. oral) contributes more to the development of oral TL

performance. Reconsidering the measurement of out-of-class TL reading is of importance in

that sense too.

Next, as we can see in Table 7 (p.103), even though the value was not negligible (Dörnyei, 2007), correlation between the amount of out-of-class TL writing and gains in TL

oral performance was merely moderate ( $r = .26$ ). This moderate correlation can be clearly

understood with reference to the result of Freed, Segalowitz, et al. (2004). In Freed and her

colleagues' study, the amount of out-of-class TL writing was a predictor of the gains in their

participants' oral fluency. According to the authors, this oral fluency development was

induced by the increase in the memorized TL expressions attained through writing practice. In

the framework of skill acquisition theory, TL knowledge attained through passive TL contact

(i.e., reading and listening) can also be proceduralized and eventually automatized in the

domain of TL writing through writing practice. Also, in the case of TL writing, SA

participants can retain their outputs and feedback from their correspondents in the form of

explicit texts, and are able to reread and explore the reasons for failure or success of their attempts in their spare time. This characteristic of TL writing may have great importance in the memorization of words, grammar rules, or expressions in a variety of situations, and thus help SA participants to develop their TL oral fluency. Provided that Freed and her colleagues' data reflect a true relationship, it is possible for this study to interpret a moderate correlation between out-of-class TL writing and the gains in TL oral performance as follows. The correlation was not so strong because the amount of out-of-class TL writing had affected the participants' fluency of TL speech—one facet of TL oral performance (DeKeyser, 2007b, 2010; Kormos, 2006). As stated in the methodology section (see 3.3.1), TL oral performance in this study is evaluated by the delivery, language use and topic development of one's TL speech (Educational Testing Service, 2008a, pp. 44-45). To put it another way, the fluency of TL speech is only one of the target attributes of the speaking section of TOEFL IBT® (Kane, 2012) and thus the influence of out-of-class TL writing on the development of TL oral

performance was also limited.

What one should also bear in mind regarding the relationship between the amount of out-of-class TL writing and the gains in TL oral performance is that this study employed a different qualification method for the out-of-class TL writing from that used in Freed and her colleagues' study. While Freed and her colleagues' study included writing homework assignments in out-of-class TL writing, this item was excluded from the out-of-class index in the current study (see 3.3.4 for the justification of this decision). Due to this methodological difference, confirmation of the above hypothesis is open to future SA studies. It will be beneficial for future studies, then, to confirm whether the impact of out-of-class TL writing is limited to the development of TL oral fluency or whether it also impacts on, for instance, the accuracy or complexity of TL speech using the measurements employed in Freed and her colleagues' study or this study.

## **12. General Discussion**

First of all, the results pertaining to research question one were consistent with those of prior SA studies. That is, one-semester of SA can be a benchmark for the development of TL oral skill (Freed, Segalowitz, et al., 2004; Hernández, 2010; Magnan & Back, 2007; Segalowitz & Freed, 2004). This result is to some extent foreseeable because, as stated in the method section, the main purpose of each SA program (that the 46 participants of this study joined) was to develop the general TL proficiency of student. To that end, the participants took some TL classes and liberal arts classes during their SA. Even though the chances to practice TL oral skill are limited in such classes, as recorded in the interview data, it is safe to assume that this in-class participation helped SA participants to build a more spontaneous and productive learning cycle until the 5/6 stage of their SA visit. The remaining question here is, then, whether or not initial readiness affects the participants' TL contact behavior outside the classroom, and the speed of TL oral performance development.

The results of regression analyses, Mann-Whitney U-tests, descriptive statistics, and interview analysis confirmed the positive link between grammatical and motivational readiness, the amount of general out-of-class TL contact, and the gains in TL oral performance. In a nutshell, these links mean that initial grammatical and motivational readiness allow SA participants to increase the amount of general out-of-class TL contact (Hernández, 2010; Masgoret & Gardner, 2003; Mikami, 2014b), and an increase in the amount of general out-of-class TL contact helps SA participants to develop their TL oral performance.

Regarding the former, the model used in this study showed that two different types of readiness jointly account for 55% of the variance of the participants' general out-of-class TL contact behavior. Possible explanations for this result were shown in the interview data—initial grammatical knowledge supports SA participants to decode or learn from TL inputs and compose TL outputs (DeKeyser, 2010; Robinson, 2007) in the out-of-class

context; and motivational readiness helps SA participants make a greater amount of self-investment in TL learning (Allen, 2010; Hernández, 2010; Yashima & Zenuk-Nishide, 2008). Although the transferability (J. D. Brown, 2001) of the results of interview analysis must be confirmed in future SA studies, the interview data of the well-prepared participants (i.e., the HRG) and less-prepared ones (i.e., LRG) adequately explains the working mechanism behind the results of regression analyses (see also the descriptive statistics cited in Appendix B). Based on these data, the author accepted the idea that grammatical and motivational readiness indeed made behavioral difference in the amount of general TL learning effort made outside the classroom.

In the successive regression analysis, 51% of the variance of pretest to posttest gains in TL oral performance was explained by the amount of general out-of-class TL contact. This model not only confirmed the link between the domain-general TL learning effort made outside the classroom and the development of TL oral performance, but also connected the

initial readiness with the outcome of SA learning (Brecht et al., 1995; Davidson, 2010; Golonka, 2006): In this study, initial readiness could explain 28.05% of the variance of the gains in TL oral performance if we calculate the direct impact of initial readiness on the gains in TL oral performance using the (adjusted) coefficient determinations of two models ( $= .55 \times .51$ ).

The reason for the dissimilar learning results among 46 participants can be attributed to the difference in the amount and effectiveness of out-of-class TL learning. It is clear from Figure 3 (p.81), that even at the 5/6 stage of the SA program, the well-readied participants had considerably more out-of-class TL contact than the less-readied ones with the exception of out-of-class TL reading. Considering this data with the results of interview analysis, it is safe to assume that this gap in the amount of out-of-class TL learning effort has existed throughout the SA visit (or it is also plausible to hypothesize that the gap was more remarkable before the 5/6 stage of their SA visit). Because grammatical readiness is expected to affect the efficiency

of TL learning in SA contexts (Davidson, 2010; DeKeyser, 2010; Golonka, 2006) and this assumption was backed by the interview data, it now becomes possible to accept the following relationship: When the amount and efficiency of out-of-class TL learning is different and this situation continues for a semester, it indeed affects the developmental speed of TL oral performance (Allen, 2010; Brecht et al., 1995; Davidson, 2007; Golonka, 2006; Hernández, 2010; Segalowitz & Freed, 2004).

The following were confirmed in study two: (a) initial readiness is the key to increase the amount of out-of-class TL listening and speaking; (b) of the above two variables, only the amount of out-of-class TL speaking showed a clear association with the gains in TL oral performance; (c) the amount of out-of-class TL listening and reading may indirectly contribute to the development of TL oral performance; (d) the amount of out-of-class TL writing may contribute to the development of one facet of TL oral performance (i.e., the oral fluency); (e) if we are interested in the relationship between initial TL learning motivation and

the TL contact behavior in a specific domain, it will probably be more productive to use domain-specific language learning motivation as an index; and (f) the method for measuring out-of-class TL reading should be reconsidered.

The first finding—initial readiness is the key to increase the amount of out-of-class TL listening and speaking—does not contradict the data shown in study one because it is reportedly natural for SA participants to wish to engage with TL activities through which they can attain more oral TL inputs and achieve greater TL oral performance (Allen, 2010; Magnan & Back, 2007; Mikami, 2014d; Trentman, 2013); and the results of regression analyses supported the claim that the more grammatical and motivational readiness SA participants have, the more oral TL inputs and outputs they can engage in SA contexts (Allen, 2010; DeKeyser, 2010; Mikami, 2014b, 2014c).

The second finding—the strong association between the amount of out-of-class TL speaking and gains in TL oral performance—was consistent with the premise of the skill

acquisition theory in that domain specific practice is the key to proceduralization and eventual automatization of specific TL performance (DeKeyser, 2007a, 2010; Mikami, 2014b; Segalowitz, 2003).

The third and fourth statements—the indirect and partial contributions of the amount of out-of-class TL listening and reading, and writing on the development of TL oral performance—were the hypotheses induced from the data observed in this study and the findings of prior SA studies. While it was true that the correlation between out-of-class TL reading and gains in TL oral performance was weak (Dörnyei, 2007) and the same went for the relationship between out-of-class TL listening and linguistic gain, the interview data documented that the participants with higher readiness acknowledged the contribution of these passive out-of-class TL activities to their TL oral practice (see also Mikami, 2014b).

Also, Freed, Segalowitz, et al. (2004) argued that the learning effort made in TL writing contributed to the development of TL oral fluency—one facet of TL oral performance in this

study. Based on this, it is not a huge jump for this study (or future SA studies) to hypothesize that (a) the amount of out-of-class TL reading and listening is indirectly connected to the development of TL oral performance through providing SA participants with TL input applicable to TL oral practice; and (b) the amount of out-of-class TL writing is particularly connected to the development of TL oral fluency rather than that of general TL performance.

Fifth, the impact of international posture on the amount of each of the four types of out-of-class TL contact was much weaker compared with the situation where the amount of general out-of-class TL contact was used as the dependent variable. Even though this was an unexpected result, the possible attributions of these decreases were hypothesized borrowing from relevant theories and the findings of prior motivation studies. Recent motivation studies in applied linguistics have revealed that the domain specific psychological attributes are responsible for increases and decreases in domain-specific language leaning effort (Cheng, 2004; Grabe, 2009; Kondo-Brown, 2006; MacIntyre et al., 1997; Mikami et al., 2016;

Yamashita, 2013). As an example, some studies succeeded in illustrating the relationship between reading-specific TL motivation and spontaneous TL reading behavior (Fujita & Noro, 2009; Matsui & Noro, 2010; Takase, 2007; Yamashita, 2013). These findings support the possibility that future SA studies will benefit from replacing international posture with domain-specific motivational indexes when their research interest is domain-specific spontaneous learning effort. In other words, the use of international posture, a domain-general TL motivation index (Kormos & Csizér, 2008; Yashima & Zenuk-Nishide, 2008; Yashima, 2002), is more appropriate when our research interest lies in the relationship between general TL learning motivation and general learning effort, as was the case with study one of the current study.

Lastly, measurement of the amount of out-of-class TL reading should be refined. As stated in the discussion section of study two, recent reading studies in applied linguistics have shown two ways of documenting the amount of spontaneous TL reading. One way is to

treat the time spent on TL reading as the amount of TL reading and the other is to use the volume of TL reading (the number of words or papers). Because of the mixed-picture provided by prior studies (Takase, 2007; Tanemura, 2014) and the theoretically complex association between initial readiness and the amount of spontaneous TL reading, at this point, it is advisable to employ both methods of documentation. Also, the timing of data collection should be the subject of reconsideration as it was likely to be yet another factor that prevented this study from illustrating a clear relationship between initial readiness and amount of out-of-class TL reading, and between the amount of out-of-class TL reading and gains in TL oral performance. As stated in the methodology section, the language contact data was gleaned at the 5/6 stage of the SA program in this study. At that point, initial readiness seemingly had nothing to do with the amount of out-of-class TL reading (see Figure 3, p.81). As suggested in the interview data, however, it is plausible that the participants with higher readiness conducted more out-of-class TL reading than those with

lower readiness earlier than this data collection point. For these reasons, it would be better if future SA studies treated both the time spent on out-of-class TL reading and the volume of out-of-class TL reading as the amount of out-of-class TL reading, collect data multiple times, and (a) confirm the behavior of each type of out-of-class TL reading index using statistical testing, and (b) check whether or not the multiple data collection provides a different picture from the results of this study.

## **13. Conclusion**

To sum up, study one clarified the role of grammatical and motivational readiness in the development of TL oral performance in a one-semester SA program. Initial readiness was connected to the development of TL oral performance thorough boosting the amount of general out-of-class TL contact. The explanation for these linkages was explored using interview data and descriptive statistics. The results of these analyses demonstrated that the difference in the amount of TL learning effort and efficacy of TL learning are the keys to understanding variance in the outcome of SA learning. Based on these findings, this study proposed several kinds of pre-departure training which could help prospective SA participants enhance their TL grammatical skill and international posture. Such training would benefit learners who wish to enhance their TL oral performance within a limited time period, because initial readiness would affect the amount and effectiveness of general out-of-class TL learning during SA participation. It is safe to say that the results of study one have high value for a

number of SA administrators and their pre-departure training construction.

The limitation of study one is simple: It did not (at least statistically) consider the impact of in-class TL learning on the development of TL oral performance. Although the influence of in-class TL learning upon TL oral performance is expectedly limited in this study (the justification is stated in 3.2, see, p.41), it is possible that a large portion of the remaining variances in the development of TL oral performance can be explained by the impact of initial readiness on the effectiveness of in-class TL learning/practice (cf., Keck et al., 2006; Robinson, 2007).

The results of study two first confirmed that (a) initial readiness is connected to the amount of out-of-class TL listening and speaking, and (b) the amount of out-of-class TL speaking was the key to TL oral performance development. The association between initial readiness, the amount of out-of-class TL speaking, and the development of TL performance was also backed by the relevant theory (i.e., the skill acquisition theory), and therefore

accepted in this study.

The following are the limitations of study two. First, compared with a situation in which TL contact data is integrated into a single variable (i.e., the amount of general out-of-class TL contact), the predicting power of international posture on the amount of four types of out-of-class TL contact was reduced. This study attributed the reason for this to the domain-generality of international posture. As pointed out by a large body of literature in applied linguistics, domain-general motivation is strongly connected to the general TL learning effort while domain-specific TL motivation is associated with the learning effort made in a specific domain. The behavior of domain-general and domain-specific out-of-class TL contact were consistent with this argument and the findings of prior studies. Based on these, this study proposed new directions for future SA studies: Measurement of TL learning motivation should be decided considering the characteristics of dependent variable(s) (domain-general vs. domain-specific).

Second, even though it was shown that out-of-class TL reading, listening and writing theoretically could facilitate the development of TL oral performance (see also the correlation between the amount of out-of-class TL listening, writing, and speaking in Table 7), this hypothesis is awaiting confirmation.

Third, it is highly plausible that the inconclusive relationship between initial readiness, out-of-class TL reading, and gains in TL oral performance stemmed from the measurement method and the data collection point of the amount of out-of-class TL reading (and, of course, the choice of motivational index). To deal with these problems, future studies should use both time and volume (i.e., the number of words or pages that have been read in a certain timeframe) to document the amount of out-of-class TL reading, and must also collect language contact data more than once so as to illustrate a clearer relationship between initial readiness, out-of-class TL reading, and gains in TL oral performance.

Given the fact that most SA participants need to improve their TL ability within a

limited time period, the benefit of shedding light on the key factors connected to productive learning outcomes is indisputable. Therefore, future SA studies need to answer the remaining questions of study one and two, and provide further comprehensive guidelines for future SA participants and administrators.

## References

- Abramson, L., & Seligman, M. (1987). Learned Helplessness in Humans: Critique and Reformulation. *Journal of Abnormal Psychology, 87*, 49–74.
- Allen, H. W. (2010). Language learning motivation during short-term study abroad: An activity theory perspective. *Foreign Language Annals, 43*, 27–49.
- Allen, H. W., & Herron, C. (2003). A mixed-methodology investigation of the linguistic and affective outcomes of summer study abroad. *Foreign Language Annals, 36*, 370–385.
- Anderson, J., & Lebiere, C. (1998). *The automatic components of thought*. Hove: Psychology Press.
- Bandura, A., & Schunk, D. H. (1981). Cultivating competence, self-efficacy, and intrinsic interest through proximal self-motivation. *Journal of Personality and Social Psychology, 41*, 586–598.
- Benjamini, Y., & Hochberg, Y. (2000). On the adaptive control of the false discovery rate in

multiple testing with independent statistics. *Journal of Educational and Behavioral Statistics*, 25, 60–83.

Borsboom, D., Cramer, A., Kievit, R., Scholten, Z., & Franic, S. (2009). The end of construct validity. In R. W. Lissitz (Ed.), *The concept of validity* (pp. 135–170). North Carolina: Information Age Publishers.

Brecht, R., Davidson, D., & Ginsburg, R. (1995). Predictors of foreign language gain during study abroad. In B. F. Freed (Ed.), *Second language acquisition in a study abroad context* (pp. 37–66). Amsterdam: John Benjamins.

Brown, A., Iwashita, N., & McNamara, T. (2005). *An examination of eater orientations and test-taker performance on English-for-Academic-Purposes speaking tasks*. Retrieved from <http://onlinelibrary.wiley.com/doi/10.1002/j.2333-8504.2005.tb01982.x/epdf>

Brown, J. D. (2001). *Using surveys in language programs*. Cambridge: Cambridge University Press.

- Brown, J. D. (2006). Generalizability from second language research samples. *Shiken: JALT Testing & Evaluation SIG Newsletter*, 10, 24–27.
- Cheng, Y.-S. (2004). A measure of second language writing anxiety: Scale development and preliminary validation. *Journal of Second Language Writing*, 13, 313–335.
- Chernick, M. R. (2008). *Bootstrap methods: A guide for practitioners and researchers* (2nd ed.). New Jersey: Wiley.
- Chernick, M. R., & LaBudde, R. A. (2011). *An introduction to bootstrap methods with applications to R*. New Jersey: Wiley.
- Cohen, J. (1992). A power primer. *Psychological Bulletin*, 112, 155–159.
- Cohen, L., Manion, L., & Morrison, K. (2007). *Research methods in education* (6th ed.). Abingdon: Routledge.
- Cumming, G. (2013). *Understanding the new statistics: Effect sizes, confidence intervals, and meta-analysis*. Abingdon: Routledge.

- Davidson, D. E. (2007). Study abroad and outcomes measurements: the case of russian. *The Modern Language Journal*, 91, 276–280.
- Davidson, D. E. (2010). Study abroad: When, how long, and with what results? New data from the Russian front. *Foreign Language Annals*, 43, 6–26.
- DeKeyser, R. M. (2007a). Situating the concept of practice. In R. M. DeKeyser (Ed.), *Practice in a second language: Perspectives from applied linguistics and cognitive psychology* (pp. 1–18). Cambridge: Cambridge University Press.
- DeKeyser, R. M. (2007b). Study abroad as foreign language practice. In R. M. DeKeyser (Ed.), *Practice in a second language: Perspectives from applied linguistics and cognitive psychology* (pp. 208–226). Cambridge: Cambridge University Press.
- DeKeyser, R. M. (2010). Monitoring processes in Spanish as a second language during a study abroad program. *Foreign Language Annals*, 43, 80–92.
- Dewey, D. P. (2004). A comparison of reading development by learners of Japanese in

- intensive domestic immersion and study abroad contexts. *Studies in Second Language Acquisition*, 26, 303–327.
- Dörnyei, Z. (2001). *Motivational strategies in the foreign language classroom*. Cambridge: Cambridge University Press.
- Dörnyei, Z. (2007). *Research methods in applied linguistics*. Oxford: Oxford University Press.
- Dörnyei, Z. (2009). The L2 motivational self system. In Z. Dörnyei & E. Ushioda (Eds.), *Motivation, language identity and the L2 self* (pp. 9–42). Bristol: Multilingual Matters.
- Dörnyei, Z., Csizér, K., & Németh, N. (2006). *Motivational dynamics, language attitudes and language globalization: A Hungarian perspective*. Bristol: Multilingual Matters.
- Dörnyei, Z., MacIntyre, P. D., & Alastair, H. (Eds.). (2014). *Motivational Dynamics in Language Learning*. Bristol: Multilingual Matters.
- Dörnyei, Z., & Taguchi, T. (2010). *Questionnaires in second language research*:

*Construction, administration, and processing* (2nd ed.). Abingdon: Routledge.

Dörnyei, Z., & Ushioda, E. (2015). *Teaching and researching motivation* (3rd ed.). London:

Pearson Education Limited.

Educational Testing Service. (2008a). *TOEFL iBT tips. How to prepare for the TOEFL iBT<sup>TM</sup>.*

Retrieved February 8, 2010, from

[http://www.ets.org/Media/Tests/TOEFL/pdf/TOEFL\\_Tips.pdf#search='TOEFL+iBT+TI](http://www.ets.org/Media/Tests/TOEFL/pdf/TOEFL_Tips.pdf#search='TOEFL+iBT+TI)

PS'

Educational Testing Service. (2008b). *Validity evidence supporting the interpretation and use*

*of TOEFL iBT<sup>TM</sup> scores.* Retrieved from

[https://www.ets.org/s/toefl/pdf/toefl\\_ibt\\_insight\\_s1v4.pdf](https://www.ets.org/s/toefl/pdf/toefl_ibt_insight_s1v4.pdf)

Educational Testing Service. (2010). *Linking TOEFL iBT<sup>TM</sup> scores to IELTS® scores.*

Retrieved February 3, 2011, from

[https://www.ets.org/s/toefl/pdf/linking\\_toefl\\_ibt\\_scores\\_to\\_ielts\\_scores.pdf](https://www.ets.org/s/toefl/pdf/linking_toefl_ibt_scores_to_ielts_scores.pdf)

Educational Testing Service. (2011). *Reliability and comparability of TOEFL iBT scores*.

Retrieved February 4, 2011, from

[http://www.ets.org/s/toefl/pdf/toefl\\_ibt\\_research\\_s1v3.pdf#search=%27Reliability+and+comparability+of+TOEFL+iBT+scores.%27](http://www.ets.org/s/toefl/pdf/toefl_ibt_research_s1v3.pdf#search=%27Reliability+and+comparability+of+TOEFL+iBT+scores.%27)

Educational Testing Service. (2013). *TOEFL ITP test taker handbook*. Retrieved from

[https://www.ets.org/s/toefl\\_itp/pdf/toefl\\_itp\\_test\\_taker\\_handbook.pdf](https://www.ets.org/s/toefl_itp/pdf/toefl_itp_test_taker_handbook.pdf)

Ellis, R. (2009). Implicit and explicit knowledge, learning and instruction. 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. 3–25). Bristol:

Multilingual Matters.

Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using

G\*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*,

41, 1149–1160.

- Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G\*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, 39, 175–191.
- Field, A. (2005). *Discovering statistics using SPSS* (2nd ed.). Thousand Oaks: Sage.
- Freed, B. F., Dewey, D. P., Segalowitz, N., & Halter, R. (2004). The language contact profile. *Studies in Second Language Acquisition*, 26, 349–356.
- Freed, B. F., Segalowitz, N., & Dewey, D. P. (2004). Context of learning and second language fluency in French: Comparing regular classroom, study abroad, and intensive domestic immersion programs. *Studies in Second Language Acquisition*, 26, 275–301.
- Fujita, K., & Noro, T. (2009). The effects of 10-minute extensive reading on the reading speed, comprehension and motivation of Japanese high school EFL learners. *Annual Review of English Language Education in Japan*, 20, 21–30.
- Golonka, E. M. (2006). Predictors revised: Linguistic knowledge and metalinguistic

awareness in second language gain in Russian. *The Modern Language Journal*, 90, 496–505.

Grabe, W. (2009). *Reading in a second language: Moving from theory to practice*.

Cambridge: Cambridge University Press.

Haertel, E. H. (2006). Reliability. In R. L. Brennan (Ed.), *Educational measurement* (pp. 65–110). Connecticut: Praeger Publishers.

Hair, J., Anderson, R., Tatham, R., & Black, W. (1998). *Multivariate data analysis* (5th ed.). New York: Prentice Hall International.

Hernández, T. A. (2010). The relationship among motivation, interaction, and the development of second language oral proficiency in a study-abroad context. *The Modern Language Journal*, 94, 600–617.

Hoshino, T., & Maeda, T. (2006). Applying propensity-score adjustment to social surveys with non-random sampling and a selection criterion for covariates. *Proceedings of the*

- Institute of Statistical Mathematics*, 54, 191–206.
- Howell, D. C. (2012). *Statistical methods for psychology* (8th ed.). Belmont: Duxbury Press.
- Hyland, F. (2004). Learning autonomy: Contextualizing out-of-class English language learning. *Language Awareness*, 13, 180–202.
- Isabelli-García, C. (2010). Acquisition of spanish gender agreement in two learning contexts: Study abroad and at home. *Foreign Language Annals*, 43, 289–303.
- Jeon, E. H., & Yamashita, J. (2014). L2 reading comprehension and its correlates: A meta-analysis. *Language Learning*, 64, 160–212.
- Kanamaru, T., & Educational Testing Service. (2012). *TOEFL® ITP test koushiki test mondai & gakusyuu guide. [The official TOEFL® ITP tests & guide]*. Tokyo: kenkyu-sya.
- Kane, M. T. (2012). Validating score interpretations and uses. *Language Testing*, 29, 3–17.
- Keck, C., Iberri-Shea, G., Tracy-Ventura, N., & Wa-Mbaleka, S. (2006). Investigating the empirical link between task-based interaction and acquisition: A meta-analysis. In J. M.

- Norris & L. Ortega (Eds.), *Synthesizing research on language learning and teaching* (pp. 91–131). Amsterdam: John Benjamins.
- Kinginger, C. (2011). Enhancing language learning in study abroad. *Annual Review of Applied Linguistics*, 31, 58–73.
- Kohro, Y. (2001). A pilot study of the linguistic impact of study abroad experiences on the writing of Japanese college ESL learners in America. *JACET Bulletin*, 33, 57–72.
- Kondo-Brown, K. (2006). Affective variables and Japanese L2 reading ability. *Reading in a Foreign Language*, 18, 55–71.
- Kormos, J. (2006). *Speech Production and Second Language Acquisition*. Abingdon: Routledge.
- Kormos, J., & Csizér, K. (2008). Age-Related Differences in the Motivation of Learning English as a Foreign Language: Attitudes, Selves, and Motivated Learning Behavior. *Language Learning*, 58, 327–355.

- Krashen, S. (1985). *The input hypothesis*. Harlow: Longman.
- Larson-Hall, J., & Herrington, R. (2010). Improving data analysis in second language acquisition by utilizing modern developments in applied statistics. *Applied Linguistics*, 31, 368–390.
- Llanes, À., & Muñoz, C. (2013). Age effects in a study abroad context: Children and adults studying abroad and at home. *Language Learning*, 63, 63–90.
- Long, M. (2014). *Second language acquisition and task-based language teaching*. New Jersey: Wiley-Blackwell.
- MacIntyre, P. D., Noels, K. A., & Clément, R. (1997). Biases in self-ratings of second language proficiency: The role of language anxiety. *Language Learning*, 47, 265–287.
- Magnan, S. S., & Back, M. (2007). Social interaction and linguistic gain during study abroad. *Foreign Language Annals*, 40, 43–61.
- Masgoret, A. M., & Gardner, R. C. (2003). Attitudes, motivation, and second language

- learning: A meta-analysis of studies conducted by Gardner and associates. *Language Learning*, 53, 123–163.
- Matsui, T., & Noro, T. (2010). The effects of 10-minute sustained silent reading on junior high school EFL learners' reading fluency and motivation. *Annual Review of English Language Education in Japan*, 21, 71–80.
- Maxwell, S. E. (2000). Sample size and multiple regression analysis. *Psychological Methods*, 5, 434–458.
- Messick, S. (1989). Validity. In R. L. Linn (Ed.), *Educational Measurement* (3rd ed., pp. 13–103). New York: American Council on Education and Macmillan.
- Mikami, H. (2014a). Extramural L2 contact in a study-abroad learning context. *Language Education & Technology Journal of Central Japan*, 25, 53–62.
- Mikami, H. (2014b). Motivation as a link between grammatical readiness and L2 oral proficiency development during a study-abroad program. *Language Education &*

- Technology*, 51, 111–140.
- Mikami, H. (2014c). *Motivational readiness and interaction in L2 oral proficiency development during a study-abroad program*. Paper presented at the International Applied Linguistics Association Conference 2014. Brisbane.
- Mikami, H. (2014d). Readiness, language contact and L2 oral proficiency development during a one-semester study-abroad program. *JACET Journal*, 58, 79–98.
- Mikami, H., Leung, C. Y., & Yoshikawa, L. (2016). Psychological Attributes in Foreign Language Reading: An Explorative Study of Japanese College Students. *The Reading Matrix*, 16, 49–62.
- Mislevy, R. J. (2009). Validity from the perspective of modeal-based reasoning. In R. W. Lissitz (Ed.), *The concept of validity* (pp. 83–108). North Carolina: Information Age Publishers.
- Mizumoto, A., & Plonsky, L. (2015). R as a lingua franca: Advantages of using R for

- quantitative research in applied linguistics. *Applied Linguistics*, Advance online publication.
- Muñoz, C. (2006). The effects of age on foreign language learning: the BAF project. In C. Muñoz (Ed.), *Age and the rate of foreign language learning* (pp. 1–40). Bristol: Multilingual Matters.
- Ockey, G. J., Koyama, D., Setoguchi, E., & Sun, A. (2014). The extent to which TOEFL iBT speaking scores are associated with performance on oral language tasks and oral ability components for Japanese university students. *Language Testing*, 32, 39–62.
- Plonsky, L., Egbert, J., & Laflair, G. T. (2015). Bootstrapping in applied linguistics: Assessing its potential using shared data. *Applied Linguistics*, 36, 591–610.
- Richards, L. (2005). *Handling qualitative data: A practical guide*. Thousand Oaks: Sage.
- Robinson, P. (2007). Aptitude, abilities, contexts and practice. In R. M. DeKeyser (Ed.), *Practice in a second language: Perspectives from applied linguistics and cognitive*

*psychology* (pp. 208–226). Cambridge: Cambridge University Press.

Robson, C. (1993). *Real world research: A resource for social scientists and practitioner researchers*. New Jersey: Wiley-Blackwell.

Rosenbaum, P. R., & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, 70, 41–55.

Ryan, R. M., & Deci, E. L. (2000). Intrinsic and extrinsic motivations: Classic definitions and new directions. *Contemporary Educational Psychology*, 25, 54–67.

Saito, Y., Garza, T. J., & Horwitz, E. K. (1999). Foreign language reading anxiety. *The Modern Language Journal*, 83, 202–218.

Sasaki, M. (2011). Effects of varying lengths of study-abroad experiences on Japanese EFL students' L2 writing ability and motivation: A longitudinal study. *TESOL Quarterly*, 45, 81–105.

Sawaki, Y., Stricker, L. J., & Oranje, A. H. (2009). Factor structure of the TOEFL

- Internet-Based test. *Language Testing*, 26, 5–30.
- Schunk, D. H., & Zimmerman, B. J. (2007). Influencing children's self-efficacy and self-regulation of reading and writing through modeling. *Reading and Writing Quarterly*, 23, 7–25.
- Segalowitz, N. (2003). Automaticity and second language acquisition. In C. Doughty & M. Long (Eds.), *The handbook of second language acquisition* (pp. 382–408). New Jersey: Wiley-Blackwell.
- Segalowitz, N., & Freed, B. F. (2004). Context, contact, and cognition in oral fluency acquisition: Learning Spanish in at home and study abroad contexts. *Studies in Second Language Acquisition*, 26, 173–199.
- Skehan, P. (2002). Theorizing and updating aptitude. In P. Robinson (Ed.), *Individual differences and instructed language learning* (pp. 69–94). Amsterdam: Jhon Benjamins.
- Taguchi, N. (2008). Cognition, language contact, and the development of pragmatic

comprehension in a study-abroad context. *Language Learning*, 58, 33–71.

Taguchi, T., Magid, M., & Papi, M. (2009). The L2 motivational self system among Japanese,

Chinese and Iranian learners of English: A comparative study. In Z. Dörnyei & E.

Ushioda (Eds.), *Motivation, language identity and the L2 self* (pp. 66–97). Bristol:

Multilingual Matters.

Takase, A. (2007). Japanese high school students' motivation for extensive L2 reading.

*Reading in a Foreign Language*, 19, 1–18.

Tanaka, K., & Ellis, R. (2003). Study-abroad, language proficiency and learner beliefs about

language learning. *JALT Journal*, 25, 63–85.

Tanemura, S. (2014). *Affective components underlying EFL learners' extensive reading:*

*Examining the components' inner structures and predicting extensive reading behaviors*

(Doctoral dissertation). Retrieved from

<http://ir.nul.nagoya-u.ac.jp/jspui/handle/2237/21227>.

- Trentman, E. (2013). Arabic and English during study abroad in Cairo, Egypt: Issues of access and use. *The Modern Language Journal*, 97, 457–473.
- Tufféry, S. (2011). *Data mining and statistics for decision making*. New Jersey: Wiley.
- Ushioda, E. (1996). Developing a dynamic concept of L2 motivation. In T. Hickey & J. Williams (Eds.), *Language, Education and Society in a Changing World* (pp. 239–245). Dublin/Bristol: IRAAL/Multilingual Matters.
- Ushioda, E. (1998). Effective motivational thinking: A cognitive theoretical approach to the study of language learning motivation. In E. A. Soler & V. C. Espurz (Eds.), *Current Issues in English Language Methodology* (pp. 77–89). Castello de la Plana: Universitat Jaume.
- Wilkinson, L., & Task Force on Statistical Inference. (1999). Statistical methods in psychology journals: Guidelines and explanations. *American Psychologist*, 54, 594–604.
- Yamashita, J. (2013). Effects of extensive reading on reading attitudes in a foreign language.

- Reading in a Foreign Language*, 25, 248–263.
- Yashima, T. (2000). Orientations and motivation in foreign language learning: A study of Japanese college students. *JACET Bulletin*, 31, 121–133.
- Yashima, T. (2002). Willingness to communicate in a second language: The Japanese EFL context. *The Modern Language Journal*, 86, 55–66.
- Yashima, T. (2009). International posture and the ideal L2 self in the Japanese EFL context. In Z. Dörnyei & E. Ushioda (Eds.), *Motivation, Language Identity and the L2 Self* (pp. 144–163). Bristol: Multilingual Matters.
- Yashima, T., & Zenuk-Nishide, L. (2008). The impact of learning contexts on proficiency, attitudes, and L2 communication: Creating an imagined international community. *System*, 36, 566–585.
- Zimmerman, B. J. (2006). Development and adaptation of expertise: The role of selfregulatory process and belief. In A. Ericsson, N. Charens, P. Feltovich, & R.

Hoffman (Eds.), *Handbook of Expertise and Expert Performance* (p. 707—724).

Cambridge: Cambridge University Press.

## Appendix

### Appendix A: Participant information

ID	Pretest TOEFL iBT	Posttest TOEFL iBT	Gains in TL Oral Performance		Grammatical Knowledge Index		International Posture	
	Raw Score	Raw Score	Raw Score	z-score	Raw Score	z-score	Raw Score	z-score
1	17.00	19.00	2.00	-1.04	18.00	-1.43	78.00	-0.11
2	16.00	19.00	3.00	-0.43	19.00	-1.23	68.00	-1.14
3	14.00	20.00	6.00	1.42	26.00	0.20	83.00	0.41
4	18.00	21.00	3.00	-0.43	34.00	1.83	94.00	1.55
5	14.00	19.00	5.00	0.80	30.00	1.01	71.00	-0.83
6	17.00	21.00	4.00	0.19	28.00	0.61	75.00	-0.42
7	14.00	19.00	5.00	0.80	32.00	1.42	91.00	1.24
8	15.00	16.00	1.00	-1.66	22.00	-0.62	82.00	0.31
9	17.00	19.00	2.00	-1.04	30.00	1.01	81.00	0.20
10	16.00	21.00	5.00	0.80	21.00	-0.82	89.00	1.03
11	16.00	19.00	3.00	-0.43	16.00	-1.84	75.00	-0.42
12	16.00	22.00	6.00	1.42	32.00	1.42	91.00	1.24
13	17.00	21.00	4.00	0.19	29.00	0.81	83.00	0.41
14	17.00	21.00	4.00	0.19	27.00	0.40	87.00	0.83
15	15.00	17.00	2.00	-1.04	16.00	-1.84	70.00	-0.93

Participant information (continued)

ID	Pretest TOEFL iBT	Postest TOEFL iBT	Gains in TL Oral Performance		Grammatical Knowledge Index		International Posture	
	Raw Score	Raw Score	Raw Score	z-score	Raw Score	z-score	Raw Score	z-score
16	15.00	19.00	4.00	0.19	25.00	0.00	83.00	0.41
17	19.00	20.00	1.00	-1.66	28.00	0.61	91.00	1.24
18	13.00	20.00	7.00	2.03	26.00	0.20	94.00	1.55
19	15.00	17.00	2.00	-1.04	24.00	-0.21	76.00	-0.31
20	14.00	16.00	2.00	-1.04	19.00	-1.23	88.00	0.93
21	15.00	18.00	3.00	-0.43	24.00	-0.21	75.00	-0.42
22	21.00	24.00	3.00	-0.43	36.00	2.24	92.00	1.34
23	15.00	16.00	1.00	-1.66	25.00	0.00	68.00	-1.14
24	17.00	22.00	5.00	0.80	30.00	1.01	88.00	0.93
25	16.00	20.00	4.00	0.19	26.00	0.20	66.00	-1.35
26	15.00	17.00	2.00	-1.04	21.00	-0.82	76.00	-0.31
27	15.00	20.00	5.00	0.80	26.00	0.20	67.00	-1.25
28	13.00	19.00	6.00	1.42	26.00	0.20	80.00	0.10
29	14.00	16.00	2.00	-1.04	18.00	-1.43	75.00	-0.42
30	17.00	18.00	1.00	-1.66	27.00	0.40	71.00	-0.83
31	18.00	24.00	6.00	1.42	30.00	1.01	99.00	2.07

Participant information (continued)

ID	Pretest TOEFL iBT	Posttest TOEFL iBT	Gains in TL Oral Performance		Grammatical Knowledge Index		International Posture	
	Raw Score	Raw Score	Raw Score	z-score	Raw Score	z-score	Raw Score	z-score
32	18.00	23.00	5.00	0.80	25.00	0.00	85.00	0.62
33	17.00	21.00	4.00	0.19	23.00	-0.41	70.00	-0.93
34	16.00	19.00	3.00	-0.43	26.00	0.20	75.00	-0.42
35	14.00	18.00	4.00	0.19	21.00	-0.82	81.00	0.20
36	16.00	19.00	3.00	-0.43	18.00	-1.43	68.00	-1.14
37	14.00	18.00	4.00	0.19	28.00	0.61	68.00	-1.14
38	13.00	16.00	3.00	-0.43	24.00	-0.21	78.00	-0.11
39	14.00	20.00	6.00	1.42	28.00	0.61	75.00	-0.42
40	17.00	23.00	6.00	1.42	27.00	0.40	76.00	-0.31
41	19.00	20.00	1.00	-1.66	12.00	-2.65	50.00	-3.01
42	16.00	19.00	3.00	-0.43	28.00	0.61	74.00	-0.52
43	14.00	18.00	4.00	0.19	24.00	-0.21	81.00	0.20
44	17.00	21.00	4.00	0.19	25.00	0.00	78.00	-0.11
45	16.00	22.00	6.00	1.42	28.00	0.61	96.00	1.76
46	18.00	23.00	5.00	0.80	23.00	-0.41	73.00	-0.62

Participant information (continued)

ID	Out-of-Class TL contact (Reading)		Out-of-Class TL contact (Listening)		Out-of-Class TL contact (Writing)		Out-of-Class TL contact (Speaking)	
	Raw Score	z-score	Raw Score	z-score	Raw Score	z-score	Raw Score	z-score
1	13.70	-0.53	18.00	0.59	1.00	-0.82	6.00	-0.98
2	15.80	-0.09	11.00	-0.48	2.00	-0.31	7.00	-0.86
3	11.00	-1.11	22.00	1.20	3.00	0.20	19.00	0.65
4	10.00	-1.32	27.00	1.96	4.00	0.71	28.00	1.77
5	12.50	-0.79	27.00	1.96	3.00	0.20	19.00	0.65
6	15.00	-0.26	7.00	-1.09	1.00	-0.82	15.10	0.16
7	15.00	-0.26	19.00	0.74	4.00	0.71	15.60	0.22
8	17.50	0.28	8.00	-0.94	1.00	-0.82	3.00	-1.36
9	8.00	-1.75	22.70	1.30	2.00	-0.31	7.00	-0.86
10	21.00	1.02	7.00	-1.09	1.00	-0.82	18.00	0.52
11	13.00	-0.68	18.00	0.59	4.00	0.71	15.00	0.14
12	21.00	1.02	21.00	1.04	4.00	0.71	21.00	0.90
13	19.00	0.59	10.00	-0.64	4.00	0.71	22.00	1.02
14	19.50	0.70	13.00	-0.18	2.00	-0.31	15.00	0.14
15	15.00	-0.26	4.00	-1.55	1.00	-0.82	3.00	-1.36

Participant information (continued)

	Out-of-Class TL contact (Reading)		Out-of-Class TL contact (Listening)		Out-of-Class TL contact (Writing)		Out-of-Class TL contact (Speaking)	
ID	Raw Score	z-score	Raw Score	z-score	Raw Score	z-score	Raw Score	z-score
16	11.00	-1.11	19.00	0.74	1.00	-0.82	21.00	0.90
17	28.50	2.62	18.40	0.65	3.00	0.20	5.00	-1.11
18	9.00	-1.53	13.00	-0.18	11.00	4.26	29.00	1.90
19	21.00	1.02	8.00	-0.94	6.00	1.72	5.00	-1.11
20	13.50	-0.58	6.00	-1.25	2.00	-0.31	7.00	-0.86
21	12.00	-0.90	16.00	0.28	6.00	1.72	8.00	-0.73
22	8.50	-1.64	21.00	1.04	7.50	2.48	35.00	2.65
23	12.50	-0.79	4.00	-1.55	1.00	-0.82	3.00	-1.36
24	12.00	-0.90	25.50	1.73	3.50	0.45	12.00	-0.23
25	20.00	0.81	13.00	-0.18	2.00	-0.31	16.00	0.27
26	12.00	-0.90	6.00	-1.25	1.00	-0.82	2.50	-1.42
27	17.00	0.17	12.00	-0.33	2.00	-0.31	13.00	-0.11
28	12.00	-0.90	24.00	1.50	1.00	-0.82	20.00	0.77
29	14.00	-0.47	6.00	-1.25	1.00	-0.82	4.50	-1.17
30	15.00	-0.26	7.00	-1.09	2.00	-0.31	3.50	-1.30
31	23.00	1.45	23.00	1.35	4.00	0.71	22.00	1.02

Participant information (continued)

ID	Out-of-Class TL contact (Reading)		Out-of-Class TL contact (Listening)		Out-of-Class TL contact (Writing)		Out-of-Class TL contact (Speaking)	
	Raw Score	z-score	Raw Score	z-score	Raw Score	z-score	Raw Score	z-score
32	20.00	0.81	21.00	1.04	1.00	-0.82	25.00	1.40
33	20.50	0.91	12.00	-0.33	4.00	0.71	21.00	0.90
34	19.00	0.59	13.00	-0.18	2.00	-0.31	10.00	-0.48
35	19.50	0.70	13.00	-0.18	2.00	-0.31	13.00	-0.11
36	17.00	0.17	11.00	-0.48	2.00	-0.31	8.00	-0.73
37	19.00	0.59	11.00	-0.48	2.00	-0.31	10.00	-0.48
38	19.00	0.59	15.00	0.13	2.00	-0.31	10.00	-0.48
39	23.00	1.45	11.00	-0.48	1.00	-0.82	23.00	1.15
40	19.00	0.59	7.00	-1.09	1.00	-0.82	21.00	0.90
41	6.00	-2.17	5.00	-1.40	1.00	-0.82	2.00	-1.49
42	15.50	-0.15	15.50	0.20	2.00	-0.31	10.00	-0.48
43	23.50	1.55	15.00	0.13	2.00	-0.31	11.00	-0.36
44	20.00	0.81	7.00	-1.09	1.00	-0.82	18.00	0.52
45	20.00	0.81	22.00	1.20	4.00	0.71	21.00	0.90
46	16.50	0.06	16.50	0.36	2.00	-0.31	14.00	0.02

Participant information (continued)

Out-of-Class TL contact (General)

ID	Raw Score	z-score	ID	Raw Score	z-score
1	38.70	-0.58	16	52.00	0.37
2	35.80	-0.78	17	55.00	0.58
3	55.00	0.58	18	62.00	1.08
4	69.00	1.58	19	40.00	-0.49
5	61.50	1.04	20	28.50	-1.30
6	38.10	-0.62	21	42.00	-0.34
7	53.60	0.48	22	72.00	1.79
8	29.50	-1.23	23	20.50	-1.87
9	39.70	-0.51	24	53.00	0.44
10	47.00	0.01	25	51.00	0.30
11	50.00	0.22	26	21.50	-1.80
12	67.00	1.43	27	44.00	-0.20
13	55.00	0.58	28	57.00	0.72
14	49.50	0.19	29	25.50	-1.52
15	23.00	-1.69	30	27.50	-1.37

Participant information (continued)

---

Out-of-Class  
TL contact (General)

ID	Raw Score	z-score	ID	Raw Score	z-score
31	72.00	1.79	46	49.00	0.15
32	67.00	1.43			
33	57.50	0.76			
34	44.00	-0.20			
35	47.50	0.05			
36	38.00	-0.63			
37	42.00	-0.34			
38	46.00	-0.06			
39	58.00	0.79			
40	48.00	0.08			
41	14.00	-2.33			
42	43.00	-0.27			
43	51.50	0.33			
44	46.00	-0.06			
45	67.00	1.43			

---

Appendix B: Descriptive statistics on two groups' TL contact

Low Readiness Group					High Readiness Group				
Item	<i>M</i>	<i>CV</i>	<i>Min</i>	<i>Max</i>	Item	<i>M</i>	<i>CV</i>	<i>Min</i>	<i>Max</i>
Speaking	10.70	0.59	2.00	23.00	Speaking	17.60	0.46	3.00	35.00
Reading	16.12	0.23	6.00	23.00	Reading	16.31	0.35	8.00	28.50
Listening	9.88	0.41	3.00	20.00	Listening	14.22	0.35	5.00	21.50
Writing	2.12	0.67	1.00	6.00	Writing	3.17	0.74	1.00	11.00
ESL Class	11.36	0.19	8.00	15.00	ESL Class	10.74	0.14	7.50	12.00
Liberal arts	2.78	0.49	1.50	6.00	Liberal arts	3.19	0.45	1.50	6.00
Homework	10.30	0.43	3.00	22.00	Homework	9.67	0.25	6.00	15.00

*Note.* LRG = Low Readiness Group ( $n = 25$ ), HRG = High Readiness Group ( $n = 21$ ), descriptive statistics were computed using the raw scores.

Appendix C: Original question items in the language contact profile

---

Q1. During the SA program, how many classes did you participate?

1. English as Second Language Class [ ] class(es)
2. Regular Subjects/Liberal Arts [ ] class(es)

Q2. Please describe contents, hours of instructions and levels of the classes that you have answered in Q1.

[1] English as Second Language Class

1. Contents of the class [ ]

Hours of instruction [ ] hours

Class level  Elementary or equivalent

Please check  Intermediate-Low or equivalent

the box which  Intermediate or equivalent

represents  Intermediate-High or equivalent

your class  Advanced-Low or equivalent

level  Advanced or equivalent

*Note*. In an actual questionnaire, 14 identical questions followed this question.

[2] Regular Subjects/Liberal Arts

1. Contents of the class [ ]

Hours of instruction [ ] hours

*Note*. In an actual questionnaire, 4 identical questions followed this question.

---

## Appendix D: The definition of terminology

### A

1. *Automatization*: Automatization progresses as situation-specific rules (i.e., procedural knowledge) are synthesized into more general rules. The synthesis results in “spontaneous, effortless, fast and errorless (DeKeyser, 2007a, p.3)” language performance (e.g., the correct and fluent use of multiple grammar rules in a language activity (Robinson, 2007; Skehan, 2002) (see also declarative knowledge and procedural knowledge).

### D

1. *Declarative knowledge*: conscious knowledge of grammar (i.e., the grammar rules that one can explicitly explain) (DeKeyser, 2007a). The operation of declarative knowledge consumes time and moderate to high cognitive resources (see also procedural knowledge and automatization).

## I

1. *Instrumental motivation*: self-related benefits that one expects to attain as a result of his/her language learning behavior (Hernández, 2010).

2. *Integrative motivation*: an interest in communicating with the target language group as well as positive attitudes toward TL native speakers and their culture (Hernández, 2010).

3. *International posture*: “a tendency to see oneself as connected to the international community, to have concerns for international affairs and a readiness to interact with people other than Japanese (Yashima & Zenuk-Nishide, 2008, p. 567)”.

## L

1. *Language contact/TL contact*: the variety of language activities (a) from which SA participants can learn knowledge and skills in their TL (i.e., written and spoken inputs, including feedback for one’s production), and (b) where SA participants can practice their newly learned knowledge and skills in the TL for more proficient oral performance.

2. *Learner's factors*: See *Readiness*.

## P

1. *Procedural knowledge*: behaviorally routinized grammar knowledge in a specific situation.

Behavioral routinization takes place as a result of initial and/or repeated situation-specific

practice (cf., DeKeyser, 2007a). Because procedural knowledge is routinized within learners'

language systems, it can be applied to specific language performance (e.g., the use of specific

grammar form) much faster, and with a lower error rate and cognitive load (see also

declarative knowledge and automatization).

## R

1. *Readiness/Learner's factors*: the psychological attributes relevant to the efficiency of TL

leaning and the pursuit or avoidance of TL contact in SA contexts.

## S

1. *Study-abroad/SA*: learning one's TL as a main subject in a host country where it is spoken

by inhabitants as their first language.