To be published in *Language Learning* (Wiley Blackwell)

Motivation, Emotion, Learning Experience and Second Language Comprehensibility Development in Classroom Settings: A Cross-Sectional and Longitudinal Study

Kazuya Saito[1], Jean-Marc Dewaele, Mariko Abe, & Yo In'nami

Abstract

This study presents a cross-sectional and longitudinal analysis of how 108 high school students in English-as-a-Foreign-Language (EFL) classrooms enhanced the comprehensibility of their second language (L2) speech according to different motivation, emotion and experience profiles. Overall, the students' learning patterns were primarily associated with their emotional states (anxiety vs. enjoyment), and secondarily with their motivational dispositions (clear vision of ideal future selves). The students' anxiety (together with weaker Ideal L2 Self) negatively related to their performance at the beginning of the project which they had achieved after several years of EFL instruction. Their enjoyment (together with greater Ideal L2 Self) predicted the extent to which they practiced and developed their L2 speech within the time framework of the project—three months. The results suggest that more regular/frequent L2 use with positive emotions directly impacts acquisition, which may in turn lead to the lessening of negative emotions and better L2 proficiency in the long run.

Key words: Second language speech, foreign language learning, motivation, emotion

[1] Parts of this study were presented at the 2017 Japan Second Language Acquisition Research Forum conference. We are grateful to anonymous *Language Learning* reviewers and the journal editor, Pavel Trofimovich, for their helpful comments on earlier versions of the paper. We also acknowledge Hui Sun for her help for data collection, coding and analyses. This study was funded by a grant from the Grant-in-Aid for Scientific Research in Japan (16H03455) and the Birkbeck College Additional Research Fund.

There is a theoretical consensus in the field of second language acquisition (SLA) that adult second language (L2) learners are able to learn new sounds and improve oral proficiency as a function of increased input and practice in the target language (e.g., Flege, 2016). Yet, the final outcome of such late SLA is subject to a great deal of variability, especially in foreign language classroom settings, where the amount of L2 input and output is substantially limited (Muñoz, 2014). To investigate the source of such individual differences, scholars have examined the role of L2 learners' experience inside and outside classrooms in L2 oral proficiency development (Saito & Hanzawa, 2016). With respect to learner-internal factors, much research attention has been given to the social and psychological dimensions of individual differences. For example, there has been extensive research conceptualizing, surveying, validating and refining different constructs of motivation (e.g., Ideal vs. Ought-to L2 Self: Dörnyei, 2005) and emotion (e.g., Anxiety vs. Enjoyment: Dewaele & MacIntyre, 2014) specific and relevant to successful foreign language learners in various classroom contexts. It has been recently proposed that motivation and emotion are intertwined as any motivated actions entail certain types of negative and positive emotion (Teimouri, 2016).

There are a set of questions which have remained unanswered, however, and which may contribute to our understanding of the mechanisms underlying successful foreign language learning: Whether, to what degree and how L2 learners' motivational and emotional states, different L2 experience profiles, and actual L2 speech development patterns are essentially related to each other. In the context of 108 Japanese English-as-a-Foreign Language (EFL) students, the present study elucidated (a) how L2 motivation and emotion orientations influenced participants' practice of the target language; and (b) how both learner-external (experience) and -internal (motivation, emotion) factors interacted to ultimately impact overall comprehensibility of their L2 speech over one academic semester (i.e., three months). The present study combines a cross-sectional perspective (comparing students' initial motivation, emotion and proficiency profiles at the start of data collection) and a longitudinal perspective (linking their motivation and emotion to any change in proficiency and experience during the project).

Background

Motivation, Emotion, Behaviours and SLA

Motivation is one of the most extensively-researched topics in the field of SLA. Motivation has been found to lead to different language learning behaviours in various contexts (for a research synthesis, see Boo, Dörnyei, & Ryan, 2015). The L2 Motivational Self System (Dörnyei, 2005) has established itself as the main theoretical framework for analyzing the motivational dispositions of L2 learners, especially in foreign language classrooms. In conjunction with the possible selves theory and self-discrepancy theory in social psychology (Higgins, 1987; Markus & Nurius, 1986), this model states that L2 learners' clear vision of their future selves exerts a significant impact on their current behaviours (i.e., whether, how often and in what way they use/practice a L2) and, by extension, their achievement (i.e., the extent to which they can improve their proficiency). Such possible selves can be conceptualized in terms of two self-guides: the Ideal L2 self and the Ought-to L2 self.

The Ideal L2 self, which roughly corresponds to not only integrative but also instrumental motives with a promotional focus, refers to the self-image of an ideal L2 user that one wants to become (Dörnyei, 2005). If L2 learners find that their current proficiency levels are distant from such desirable future level, they aim to fill in the discrepancy by striving to use, practice and

improve their L2 ability. In contrast, Ought-to Self relates to what characteristics one believes they ought to have in order to meet certain expectations and avoid negative outcomes in the future. These expectations are not one's own, but rather are imported and imposed images of the future that the learner internalizes to some extent. In the present study, we focus on the social aspect of Ought-to Self, assuming that L2 learners with stronger social Ought-to L2 Self make greater efforts to study the target language so as to achieve what their social networks or communities (e.g., friends, family members) expect them to achieve (i.e., instrumental motivation with a prevention focus) (Dörnyei, Csizér, & Németh, 2006).

To further understand the mechanism underlying motivation effects in SLA, L2 scholars have recently begun to emphasize the importance of including L2 learners' emotional states in the L2 Motivational Self System, as L2 learners' perception of actual and future selves may trigger different emotional reactions (Dörnyei & Ushioda, 2009; MacIntyre & Gregersen, 2012; Papi & Teimouri, 2014; Teimouri, 2016). So far, much of the discussion has been concerned with one kind of negative emotion in classroom settings—anxiety (Gkonou, Daubney, & Dewaele, 2017). MacIntyre and Gardner (1989) argued that situation-specific language anxiety builds gradually because of repeated experience of anxiety leading the learner to associate the L2 language class with anxiety. Moreover, while anxiety can fluctuate during a class, it is typically linked to what L2 learners have experienced over a prolonged period of learning, such as test scores, attitude towards the target language, and standing relative to their peers in class (Horwitz, 2017). Anxiety has been linked to harsh error correction (Gregersen, 2003), incompatibility between teachers and students (Gregersen, MacIntyre, & Maze, 2014), and thus has been identified as having a debilitating effect on L2 learning and achievement (for a review, see Horwitz, 2017, MacIntyre, 2017). Crucially, such negative emotions are claimed to influence the Ought-to Self aspect of motivation and vice versa; prevention-focused L2 learners feel anxious when they perceive difficulty in achieving their obligations, duties, and responsibilities regarding their foreign language learning (Papi & Teimouri, 2014).

More recently, certain scholars have argued in favor of a more holistic view of emotions, including the role of positive emotions in foreign language classrooms (Dewaele & MacIntyre, 2014; MacIntyre & Gregersen, 2014). For example, one kind of positive emotion—enjoyment—is believed to help L2 learners better attend to, process and acquire a target language (Dewaele & MacIntyre, 2014; Dewaele & Alfawzan, to appear). Such a positive, acquisition-friendly atmosphere can occur in certain classrooms, where activities are adequately challenging (slightly beyond L2 learners' competence), creative and unpredictable, have clear benefits and purpose, and where efforts are made by teachers to facilitate task completion (while providing praise, encouragement and feedback in a humorous, constant fashion). Different from anxiety, enjoyment may relate to the Ideal Self aspect of motivation, because a sense of elation arises when L2 learners achieve their more internalized future self-guides—i.e., their hopes, aspirations and ideals. There is also some evidence that L2 learners with a more promotional focus tend to express more positive emotions towards their own learning experience, peers and teachers (Teimouri, 2016). Dewaele and Dewaele (2017) carried out a pseudo-longitudinal study on British pupils aged between 12 and 18. The results identified a slight increase in L2 enjoyment over time while L2 anxiety remained constant. However, the learner-internal and external factors predicting enjoyment and anxiety did change dramatically over time, suggesting dynamic change below the surface.

Over the past 20 years, much empirical research has been conducted to expound the value of self guides as motivational orientations. The Ideal L2 self has been found to have relatively

strong relationships with L2 learners' motivated behaviours (Kormos, Kiddle, & Csizér, 2011; Taguchi, Magid, & Papi, 2009); however, the associations between the Ought-to Self and L2 learners' behaviours remain unclear (e.g., Lamb, 2012; Csizér & Kormos, 2009). From a methodological perspective, these behaviours have been typically measured via L2 learners' intended learning efforts by rating relevant statements on questionnaires (e.g., "*I think that I am doing my best to learn English*"). Some researchers have cast doubt on whether such reported intentions reflect the quantity and quality of L2 learners' actual use (e.g., Ryan, 2008). Moskovsky, Assulaimani, Racheva and Harkins (2016, p. 643) pointed out that "ultimately, SLA is about achievement, that is, about attaining an adequate level of proficiency in the [target language]. Therein lies the real test for the theory—in the capacity of the self guides to predict L2 achievement" (see also Ushioda, 2016).

To date, several empirical studies have explored the relationship between EFL students' motivation and their L2 performance, although no clear link between motivation and achievement has been found to date (e.g., Lamb, 2012; Moskovsky et al., 2016; Papi & Teimouri, 2014). Although these findings have hinted that stronger self-guides may not necessarily be linked to more successful SLA, they need to be replicated with greater methodological rigour. First, L2 learners' proficiency in these previous studies has been typically measured via general proficiency tests, final grade or self-ratings at only one data collection point. Such cross-sectional designs allow researchers to explore the relationship between L2 learners' motivation, emotion and achievement. To our knowledge, however, no longitudinal studies have ever delved into how such learner-internal factors impact L2 learners' development over time. This deficit in the research corresponds to the general lack of longitudinal work in the field of SLA (Ortega & Byrnes, 2008). Furthermore, these previous studies did not control for L2 learners' supposedly diverse EFL experience—another key factor affecting L2 speech learning in foreign language settings (reviewed in detail below). Even if certain L2 students have stronger motivation and show more acquisition, it remains unclear the extent to which such correlations could be tied to the way they have practiced the target language in EFL classrooms. Thus, more research is needed to elucidate the *triangular* relationships between L2 learners' motivation/emotion orientations, experience profiles, achievement and development.

Experience, L2 Speech Learning and Individual Differences

Second language speech (the focus of the present study) is a multifaceted phenomenon which comprises a wide range of different linguistic skills. According to previous literature on naturalistic L2 speech learning, learners tend to show a great deal of improvement within a short period of immersion (e.g., first 3-4 months of stay) in terms of global (e.g., Derwing, Munro, & Thomson, 2008 for comprehensibility and accentedness), temporal (e.g., Segalowitz & Freed, 2004 for fluency), segmental (Saito & Munro, 2014 for approximants) and lexicogrammatical (Mora & Valls-Ferrer, 2012) aspects of L2 speech, as long as they use the L2 as a main language of communication in various social settings (for review on experience effects in naturalistic SLA, see Saito, 2015).

In contrast, foreign language classrooms have been referred to as "minimum input" environments (Larson-Hall, 2008, p. 36), as L2 learners in such contexts typically receive only a few hours of instruction per week without many opportunities to use the target language for the purpose of communication. In this regard, successful classroom L2 speech learning crucially depends on the extent to which certain learners actively seek and utilize every possible opportunity to practice the target language both inside and outside of classrooms. It is

unsurprising that even L2 learners within the same language classroom setting can have different L2 learning experiences, resulting in varying amounts and kinds of improvement.

Under minimum input conditions, few longitudinal studies have indeed evidenced L2 learners' statistically significant improvement in oral proficiency (e.g., Mora & Valls-Ferrer, 2012; Muñoz & Llanes, 2014; Segalowitz & Freed, 2004). In these studies, the oral performance of participants between data collection points has failed to reach statistical significance, arguably because the researchers considered participants as a categorical group (without taking into consideration individual differences), using statistical analyses such as *t*-tests and ANOVAs. Rather, the incidence and degree of foreign language learning success widely varies between individuals according to a range of affecting factors. By using variance-based analyses (e.g., correlations, regression) rather than means-based comparisons (e.g., *t*-tests, ANOVAs), previous studies have found that the multivariate nature of L2 oral proficiency development in EFL classrooms is influenced by age (Larson-Hall, 2008) and length (Muñoz, 2014) of learning, and the nature of L2 use inside (Saito & Hanzawa, 2017, for form- vs. meaning-oriented instruction) and outside (Muñoz, 2014, for extracurricular activities) classrooms. For more discussion on the selection of appropriate statistical analyses for examining multivariate data, see Plonsky & Oswald (2017).[2]

To disentangle the intricate connections between experience and classroom L2 speech learning, Saito and Hanzawa (2016) scrutinized the linguistic and learner profiles of college-level Japanese students with similar EFL backgrounds (six years of foreign language education without any experience overseas). The students' oral proficiency attainment was significantly associated with their EFL experience inside (e.g., pronunciation training) and outside (e.g., cramming school[3]) classrooms at the high school (but not junior high school) level. These results indicated that the pedagogical potential of foreign language learning can be maximized by how students optimize their most immediate L2 experience beyond the regular syllabus, bringing to light the importance of the quantity, quality and timing of experience—three key factors in the usage-based account of SLA (Ellis, 2006).

Motivation for the Present Study

A number of L2 speech researchers have examined what types of L2 learners ultimately attain high-level pronunciation performance after years of immersion in naturalistic settings. One well-researched, yet controversial factor for successful L2 speech learning is motivation. Although some studies have shown that L2 learners with highly advanced oral proficiency are likely to demonstrate a strong concern for nativelike pronunciation accuracy (e.g., Moyer, 1999), others have failed to find such significant predictive power of motivation for successful L2 pronunciation ability at all (e.g., Purcell & Suter, 1980). Rather, it has been found that the final

[2]Plonsky and Oswald (2017) emphasized that comparing mean differences via *t*-tests and ANOVAs may not be an appropriate statistical analysis method for capturing individual differences typical of L2 data. Statistically, such means-based analyses consider between-group variance as the main focus and within-group variance as error variance. For any multivariate data, Plonsky and Oswald recommend using variance-based analyses (correlations, regression), which incorporate not only categorical/group independent variables, but also continuous predictors (e.g., students' different levels of motivation/emotion) in the same analysis model.

[3] Cramming schools are private institutes that students generally attend to prepare for university entrance exams.

quality of naturalistic L2 pronunciation is strongly determined by the age, length and intensity of immersion in an L2 speaking environment (e.g., Flege, Munro, & MacKay, 1995).

To our knowledge, very few studies have examined how inexperienced L2 learners with varied levels of motivation can differentially improve their interlanguage phonological systems, especially when they engage in different kinds of foreign language classroom learning. One of these studies, Baker-Smemoe and Haslam (2013), longitudinally investigated a total of 31 Chinese EFL learners' L2 motivation (elicited via a subsection of the Pimsleur Language Learning Aptitude Battery) and L2 oral proficiency development, finding significant associations between the two. Saito, Dewaele and Hanzawa (2017) devised a tailored questionnaire for a target population of the study ($N = 40$ Japanese EFL college students). Their different levels of context-specific motivation—studying English for their future career development as a vague and long-term goal in a globalized society (i.e., international posture: see Yashima, Zenuk-Nishide, & Shimizu, 2004)—were significantly predictive of the longitudinal development of L2 oral proficiency over one academic semester. Finally, Nagle (2018) tracked the longitudinal development of 26 English learners of Spanish's oral proficiency (comprehensibility, accentedness) and motivational orientations (Ideal L2 Self, Ought-to L2 Self) over one academic year. Whereas the results of the quantitative analyses did not reveal significant links between motivation and L2 speech learning, the study provided some qualitative evidence that the participants uniquely allocated motivational resources vis-à-vis their individually different learning contexts and objectives.

On the whole, these precursor studies suggest some acquisitional value for motivation, especially in classroom L2 speech learning. However, it is noteworthy that the constructs of L2 motivation investigated therein did not build on more recent frameworks of L2 motivation and emotion research (Baker-Smemoe & Haslam, 2013; Saito et al., 2017). Whereas Nagle (2018) featured Dornyei's L2 Motivational Self System, the statistical power of the dataset could be considered relatively weak ($N = 26$ learners); therefore, their findings (the lack of significant motivation effects in particular) need to be interpreted with caution. Additionally, none of these studies took into account recent trends considering motivation and emotion as interrelated with the concept of sociopsychological individual differences (Teimouri, 2016). Further, as observed in the aforementioned L2 motivation research (e.g., Moskovsky et al., 2016), the interaction of motivation and acquisition was not probed in relation to the quantity, quality and timing of the participants' L2 experience.

Focusing on 108 first-year Japanese high school students with varied EFL backgrounds, the main objective of the present study was to scrutinize L2 motivation, emotion and experience as key factors for explaining variance in the process and outcome of L2 oral proficiency development (measured via comprehensibility). Specifically, we set out to answer two research questions: (a) to what extent do L2 learners' motivation and emotion profiles predict the way they practice the target language inside and outside classrooms; and (b) to what extent do L2 learners' emotion, motivation and experience differentially predict their language development. Departing from the previous studies (Baker-Smemoe & Haslam, 2013; Saito et al., 2017), and providing generalizable and interdisciplinary insights to the existing EFL, motivation and emotion research, the learner-external and internal factors were assessed via solid, theory-driven instruments for experience (Saito & Hanzawa, 2016), motivation (Dörnyei et al., 2006) and emotion (Dewaele & MacIntyre, 2014).

Method

The present study combines both cross-sectional and longitudinal analyses. Data from the motivation/emotion questionnaire was collected at the end of the second term (T1: December 2016), the participants' experience and speech performance was examined at both T1 and the end of the third term (T2: March 2017). Using the T1 data only, we first cross-sectionally examined how the students' motivation and emotion were associated with their past (preschool, elementary/junior high school) and current (high school) L2 use inside/outside classrooms (RQ1), and how motivation, emotion and experience affected their long-term achievement after several years of EFL experience at the start of the project (RQ2). Using the data from T1 and T2, we then explored longitudinally how the participants' T1 motivation/emotion was predictive of their current L2 use between T1 and T2 (RQ1); and how their motivation, emotion and their L2 use could affect their L2 speech development over time (RQ2). As such, we aimed to pinpoint the role of motivation, emotion and experience in L2 oral proficiency development within a specific time framework (one academic term: three months). The timeline of the study is summarized in Figure 1.

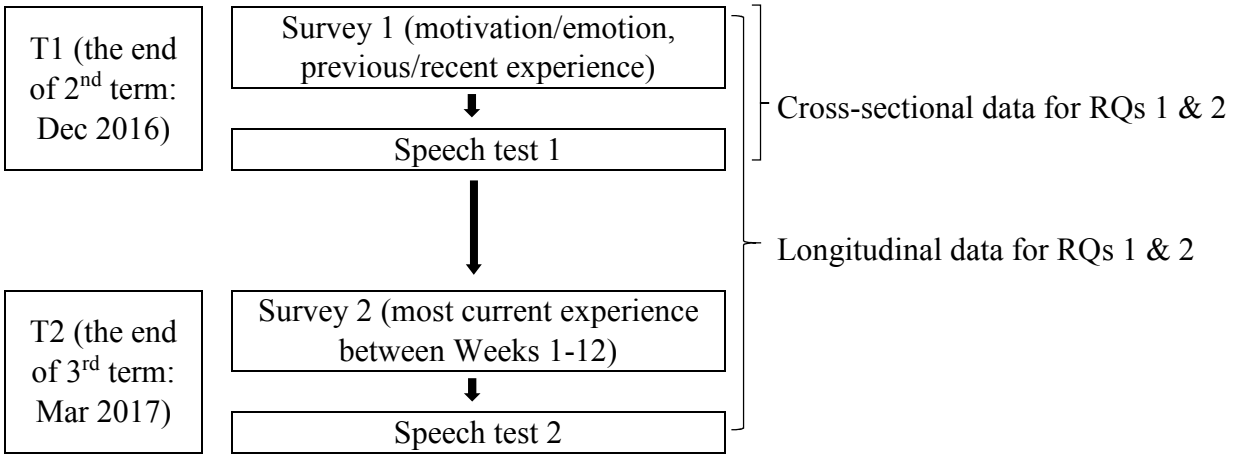

Figure 1. Summary of Research Time Framework

Participants

Japanese EFL Students. Although 122 students originally participated in the current project, 14 of them were eliminated from the main analyses because of low quality speech samples or because they did not complete the speaking tests at both T1 and T2 (for the details of the elicitation and assessment procedure, see the Measures of L2 Oral Proficiency section). In this study, all participants included in the final analyses (n = 108) were first-year students at the same prestigious high school located in Japan (age = 15-16 years, 44 males, 64 females). Although none of them had stayed abroad for more than one week (except for short family trips), their age of learning (i.e., the first exposure to L2 English) varied widely (M = 10.0 years; SD = 3.1; *Range* = 1-14 years). At this high school, these first-year students were divided into three class groups (regardless of their proficiency levels). Each group included 30-40 students and followed the same curriculum.

For the duration of the project (one academic term), all the students were required to enroll in seven 50-min EFL lessons per week. While these classes were taught by three different Japanese teachers (who equally demonstrated near-native L2 English proficiency), the students'

performance was assessed via the same test material at the end of the term, three months later. Different from other EFL studies, where participants substantially differed in terms of the number and type of classes they were taking (e.g., Muñoz, 2014 for college-level students), the EFL curriculum here was consistent across the first-year high school students in the present study: all of them took the same seven English lessons per week followed by the same final exams. The content, syllabi and observations of the EFL classes offered to the participating students have been described in detail in Supporting Information-A. In this section, therefore, only a brief description of the curriculum is provided.

According to the standard syllabus and our classroom observations, the EFL classes equally focused on writing, speaking, reading and listening throughout the term; the instruction was frequently delivered in English; and the students were always encouraged to interact with their teachers and peers in English. As we quantified later via the EFL Experience Questionnaire, the amount of L2 use (i.e., using L2 English to communicate with teachers and peers) was substantially different among the participants. While some reported that they actively and frequently used L2 English throughout classes, others reported that they chose to remain silent (see the Measures of Experience section).

In terms of the participants' general L2 English proficiency, their self-reports of the general proficiency test scores (i.e., The EIKEN Test in Practical English Proficiency) ranged from Grade Pre-2 to 2 at the beginning of the project. This indicated that they could be considered as A2 (basic) and B1 (independent) users according to the CEFR benchmarks (EIKEN Foundation of Japan, 2017).

Native speaking raters. To assess the overall comprehensibility of the Japanese students' oral proficiency, five native speakers of British English were recruited in London, UK. As described below, comprehensibility is typically measured using the scalar judgements of minimally-trained native raters. Even though such comprehensibility assessments are highly intuitive in nature, rater familiarity with foreign accented speech has been found to affect their judgements to some degree (e.g., Isaacs & Thomson, 2013). Thus, the decision was made to recruit only *experienced* raters (graduate students in an MA Applied Linguistics program at the time of the project) who reported extensive English teaching ($M_{\textit{years of teaching}}$ = 11.2 years, *Range* = 5-20 years) and L2 speech analysis experience (i.e., they had previously participated in similar L2 comprehensibility judgement sessions). Although none of them had studied Japanese nor visited Japan before the project, they reported high familiarity with Japanese-accented speech (*M* = 5.6 ranging from 5 to 6: *1 = not familiar, 6 = very familiar*).

Measures of L2 Oral Proficiency

Given that few adult L2 learners have been reported to attain nativelike L2 oral proficiency, and accent is a normal characteristic of post-pubertal SLA (Flege et al., 1995), it has been claimed that enhanced comprehensibility (rather than accent reduction) should be considered as a realistic goal for L2 speech teaching and learning (e.g., Derwing & Munro, 2015). To mirror what native speakers do in real-life situations when interacting with L2 users, many L2 speech researchers have used the intuitive judgment method for assessing the comprehensibility aspects of L2 speech (e.g., Trofimovich & Isaacs, 2012). Unlike high-stakes testing settings, where accredited raters make careful evaluations with reference to multiple proficiency descriptors, untrained raters make quick, intuitive assessments of comprehensibility (i.e., ease of understanding) on a 9-point scale (*1 = difficult to understand, 9 = easy to*

understand) right after listening to L2 speech samples. Although raters receive only minimal training/practice, they tend to show relatively high inter-rater reliability (Cronbach $\alpha > .90$), indicating the presence of native speakers' shared notion of comprehensibility (Saito, Trofimovich, & Isaacs, 2017). Extensive research on comprehensibility judgements seems to suggest that native raters selectively attend to those linguistic features which hinder successful communication and prompt understanding (for a list of communicatively-important linguistic features, see Derwing & Munro, 2015; Saito et al., 2017; Trofimovich & Isaacs, 2012).

In the present study, the participants' oral proficiency was measured by native speakers' intuitive judgement of L2 comprehensibility for the following reasons. First, comprehensibility serves as an adequate index to assess the extent to which the beginner-to-intermediate Japanese students reached the *minimum* phonological, lexical and grammatical requirements for speaking L2 English understandably regardless of foreign accentedness. Second, the results of comprehensibility judgements correspond to what matters for students in their future communicative settings—i.e., getting their meaning across successfully to native speakers. Third, although L2 learners' accentedness (linguistic nativelikeness) is resistant to change in even naturalistic settings, the comprehensibility aspect of L2 oral proficiency tends to improve under naturalistic *and* foreign language learning conditions, provided that learners use the target language (Derwing & Munro, 2015; Saito, 2015).

Materials. To collect speech samples from a relatively large number of participants (N = 122) within a short period of time (i.e., in a span of ten days), all the students participated in an automated English speaking test (i.e., the Telephone Standard Speaking Test) (ALC Press Inc, 2017) using their home phones or cellphones at their convenience at both T1 and T2. They were explicitly asked to take the exam in a quiet room to provide clear speech for proper ratings (noisy samples would negatively affect raters' assessments). The students' successful completion of the test was monitored by the testing agency and used as a part of their final grade assessment by their teachers. Outside of rare cases where the students failed to finish the test for technical reasons (e.g., unstable cellphone signals) and retook the test, they had only one opportunity to take the test. The test was 15 minutes long.

During the test, the students responded to 10 recorded questions. For each question, the students were given 45 seconds to elaborate and complete their answer without any preparation time. To avoid misunderstanding of the task instructions, all instructions were delivered in both English and Japanese. The structures of these 10 questions differed in terms of target grammar structures (present/past tense vs. comparatives) and task type (providing narratives, descriptions vs. reasoning). For each question, the testing system was programmed to ensure that the students engaged in different topics (randomly picked from the data bank of the test) at T1 and T2. The speaking test was considered to elicit extemporaneous speech, given that it was impossible for the students to do any preparation in advance.

Students' responses to the seventh question were chosen for analysis in the present study. This question measured their ability to describe and explain the events in the past. According to the test guideline (ALC Press Inc, 2017), Question 7 was considered relatively difficult (given that participants had to use the past tense consistently), and thus expected to elicit much individual variability among the participants. For each participant, a topic was randomly chosen from 17 different alternatives (e.g., favorite movie, family trip, shopping). Following the standards of previous L2 speech literature (e.g., Derwing & Munro, 2015), the first 30 seconds of each speech sample were cut and saved in a WAV file, then digitized and normalized for peak

intensity and perceived loudness. Furthermore, the researchers listened to all the speech samples and discussed whether the quality of each recording was adequate for the comprehensibility judgements. Through these discussions we decided to remove four students whose T1 or T2 samples sounded slightly distorted due to background noise, as well as two students who did not successfully complete the assigned task (exhibiting more than 20 seconds of silence) either at T1 or T2. Eight students were also removed because they did not take the speaking test at either T1 and T2. The final dataset for rating consisted of the speech of 108 participants at T1 and T2.[4]

Procedure. All the rating sessions individually took place in a quiet room at a university in London. First, the five raters received a brief explanation on the construct of comprehensibility (for training scripts, see Supporting Information-C) and familiarized themselves with the task format (i.e., a monologue task based on 17 different topics). Second, the raters practiced the procedure by rating the comprehensibility of three speech samples not included in the main dataset. During the judgement session, 216 speech samples (108 students × 2 testing sessions) were played in a randomized order via *Praat* software (Boersma & Weenink, 2017); and the raters listened to the full 30 seconds extracted for each learner. Upon listening to each sample only once, they assigned a comprehensibility score using a 9-point numerical scale. To avoid listener fatigue, the speech samples were divided into four different blocks (54 samples per block) with a ten-minute intermission between the blocks. In total, the entire session took approximately three hours per rater.

Inter-rater Reliability. The results of the Cronbach's alpha analysis demonstrated relatively high inter-rater agreement for the 9-point comprehensibility ratings among the five native raters ($\alpha = .90$). By pooling over the native raters' judgement scores, one mean rating score was given to each of the 108 students at T1 and T2, respectively.

Measures of Experience

As operationalized in previous EFL studies (Muñoz, 2014; Larson-Hall, 2008; Saito & Hanzawa, 2016), the language backgrounds of the student in our study were surveyed via a structured experience questionnaire—termed here as EFL Experience Questionnaire (see Supporting Information-B). At T1, the students filled in the first version of the questionnaire to report how they had practiced L2 English prior to as well as at the beginning of the project. The past experience variables included age of onset of L2 learning and length of learning (how many hours they had practiced L2 English inside and outside classrooms at preschool, elementary and junior high schools, respectively). As for recent experience, the students reported what percentage of time they spoke L2 English during class; and how many hours they practiced English outside of the classroom per week. In the present study, the latter extracurricular activities were divided into three subcategories: to prepare for classes, to study at cram schools and to engage in conversational activities with native speakers and English L2 users.

The second experience questionnaire was administered at T2 to probe the students' most current L2 use over the term (T1 → T2). Similar to the T1 questionnaire, the longitudinal experience variables included (a) the ratio of their L2 use inside classrooms and (b) the number

[4] To examine if topic type affected the raters' comprehensibility judgements, we ran a one-way ANOVA with their comprehensibility scores as dependent variables and topic type as independent variables at T1 and T2, respectively. The results showed that their performance did not significantly differ according to topics at T1, $F(16, 91) = 1.080$, $p = 386$, $\eta_p^2 = .160$, nor T2 $F(16, 91) = 1.552$, $p = .099$, $\eta_p^2 = .214$.

of hours they spent preparing for classes, at cram schools and on conversation activities with native speakers and English L2 users per week.

Measures of Motivation and Emotion

To survey the participants' motivation and emotion orientations, a composite questionnaire was developed in Japanese and administered at T1. In total, the composite questionnaire consisted of 58 items. For each item, participants rated the extent to which they agreed or disagreed by marking one of the 6 responses ranging from strongly agree to strongly disagree on a 6-point scale. For the items used for the final analyses, see Table 5.

The first part of the questionnaire consisted of 40 items adapted from the Japanese version of Taguchi's (2009) questionnaire, which was designed to measure multiple dimensions of motivation (e.g., integrativeness, instrumentality, family influence, attitudes to L2 community and culture) based on the L2 Motivational Self System theory (Dörnyei et al., 2006). Among them, four items corresponded to Ideal L2 Self and four items to Ought-to L2 Self. Therefore, they were used as motivation measures in the present study.

The second part of the questionnaire featured 18 statements adapted from the Foreign Language Enjoyment Questionnaire (Dewaele & MacIntyre, 2014) as well as Foreign Language Classroom Anxiety Scale (Horwitz et al., 1986). The first 10 statements were designed to tap into two dimensions of enjoyment specific to foreign language learning— private and social enjoyment in a teacher-controlled environment. The remaining eight statements reflected the physical symptoms of anxiety, nervousness and lack of confidence related to foreign language learning. Since two out of the eight statements were rephrased to indicate low anxiety, their scores were reverse-coded so that all the statements equally indicated high anxiety.

Results

L2 Oral Proficiency and Experience

The descriptive statistics of the participants' speaking performance and their past/current EFL experience are summarized in Tables 1 and 2. L2 comprehensibility development patterns appeared to be uniquely different between individuals; but such change took place mostly within a range of 3 to 6 (*1 = difficult to understand, 9 = easy to understand*). In terms of their previous EFL experience, the students substantially varied regarding how they practiced English inside and outside of their classrooms throughout preschool, elementary school and junior high school. Similarly, their recent (T1) and current (T2) EFL experience profiles inside and outside classrooms were subject to much individual variability. Whereas some students used English very often during class and made extra efforts to practice through extracurricular activities, such as preparation for classes, cram schools and conversations with native speakers and L2 users, others did not. Such EFL experience was quite comparable between the beginning (T1) and end (T2) of the term. Since our dataset is multivariate in nature and not suitable for mean-based analyses (ANOVAs, *t*-tests), following Plonsky and Oswald (2017), we used variance-based analyses (correlations, regression) to compare the participants' varied oral performance at T1 and their gain scores between T1 to T2 according to different experience, motivation and emotion profiles.

Table 1 *Descriptive Results of 108 Students' Oral Performance (Comprehensibility) at T1 and T2*

	T1				T2			
	M	*SD*	*95%CI*		*M*	*SD*	*95%CI*	
			Lower	Upper			Lower	Upper
Comprehensibility	4.55	1.54	4.25	4.84	4.43	1.52	4.14	4.72

Note. 1 = Difficult to understand, 9 = Easy to understand

Table 2 *Descriptive Statistics of Japanese Students' Past, Recent and Current EFL Experience*

A. Past EFL experience	*M*	*SD*	*Range*
• Total hours of L2 practice prior to the project	1349.9hr	977.6	300-7200
• Total hours of L2 practice at preschool	20.3hr	91.2	0-750
• Total hours of L2 practice inside EFL classroom at elementary school	175.0hr	271.8	0-1500
• Total hours of L2 practice outside EFL classroom (e.g., cram school) at elementary school	127.0hr	264.6	0-1800
• Total hours of L2 practice inside EFL classroom at junior high school	699.4hr	403.6	150-4500
• Total hours of L2 practice outside EFL classroom (e.g., cram school) at junior high school	368.0hr	435.2	0-3000
B. Recent EFL experience at T1 (end of Term 2)	***M***	***SD***	***Range***
• L2 use during class (%)	58.7%	16.8	23-95
• Total hours of all L2 practice outside EFL classrooms (per week)	8.0hr	4.4	1-21
• Total hours of preparation for class (per week)	6.7hr	3.4	0.5-19
• Total hours of cram school (per week)	0.7hr	2.1	0-10
• Total hours of conversation activities with native and non-native speakers (per week)	0.4hr	1.3	0-8
C. Current EFL experience at T2 (end of Term 3)	***M***	***SD***	***Range***
• L2 use during class (%)	55.2%	19.9	0-90
• Total hours of all L2 practice outside EFL classrooms (per week)	7.4hr	5.9	0-52
• Total hours of preparation for class (per week)	5.9hr	3.3	0-15.5
• Total hours of cram school (per week)	1.0hr	4.2	0-35
• Total hours of conversation activities with native and non-native speakers (per week)	0.5hr	1.4	0-7

Motivation and Emotion

Reliability. As shown in previous studies (Dewaele & MacIntyre, 2014; Taguchi et al., 2009), the Cronbach alpha analyses demonstrated high reliabilities for Ideal L2 Self, Ought-to L2 Self, Foreign Language Enjoyment and Anxiety (see Table 3).

Table 3 *Descriptive Statics and Reliability Coefficients of Motivation and Emotion Variables*

Variables	*M*	*SD*	*α*	95% *CI*	
				Lower	Upper
Ideal L2 Self	3.7	1.2	.82	3.5	3.9
Ought-to L2 Self	3.0	1.1	.85	2.7	3.2
Enjoyment	4.5	0.8	.83	4.4	4.7
Anxiety	3.5	0.9	.86	3.3	3.7

Validity. Next, we examined a number of underlying factors among eight items from the motivation questionnaire (4 for Ideal L2 Self, 4 for Ought-to L2 Self) and 18 items from the emotion questionnaire (10 for Enjoyment, 8 for Anxiety). In the present study, we followed Loewen and Gonulal's (2015) field-specific guidelines for analyzing factorability and determining a threshold for factor loadings. First, the students' motivation and emotion ratings were submitted to a factor analysis with Promax rotation and the minimum Kaiser criterion eigenvalue set to 1.0, respectively. As for the motivation questionnaire, the factorability of the entire dataset was confirmed via two tests: the Bartlett's test of sphericity ($\chi^2 = 431.78$, $p < .001$) and the Kaiser-Meyer-Olkin measure of sampling adequacy (.799). A decision was made to identify a "two-factor" solution which accounted for 63.3% of the total variance in the students' motivation ratings. Corresponding to the original conceptualization of the motivation questionnaire (Taguchi et al., 2009), all the items for Ideal L2 Self and Ought-to L2 Self were clustered into two different groups. Therefore, Factor 1 was labelled as Ideal L2 Self and Factor 2 as Ought-to L2 Self (see Table 4).

Table 4 *Summary of a Two-Factor Solution Based on a Factor Analysis of the Motivation Questionnaire*

Items	Factor 1 (Ideal L2 Self)	Factor 2 (Ought-to L2 Self)
A. Ideal L2 Self		
• I imagine myself as someone who is able to speak English.	**.929**	-.052
• I can imagine a situation where I am speaking English with foreigners.	**.924**	-.141
• I can imagine myself living abroad and having a discussion in English.	**.910**	-.035
• Whenever I think of my future career, I imagine myself using English.	**.673**	.368
B. Ought-to L2 Self		
• My parents believe that I must study English to be an educated person.	.104	**.750**
• Learning English is necessary because people surrounding me expect me to do so.	.069	**.749**
• I have to study English, because if I do not study it, I think my parents will be disappointed with me.	-.230	**.729**
• I study English because close friends of mine think it is important.	-.037	**.705**

Note. All loadings > .5 were highlighted in bold.

An additional factor analysis was conducted on the students' responses in the emotion questionnaire, with Promax rotation and a Kaiser criterion eigenvalue of 1.0. The two tests again confirmed the factorability of the dataset: the Bartlett's test of sphericity ($\chi^2 = 1051.82$, $p < .001$) and the Kaiser-Meyer-Olkin measure of sampling adequacy (.850). A "three-factor" solution was chosen, as it explained 60.8% of the total variance in the students' emotion ratings. According to this model, the 10 items for Enjoyment appeared to tap into two different aspects of the students' positive emotional states. Following Loewen and Gonulal's (2015) suggestion, we assigned each item to the factor that it loaded most highly on. Five items which loaded onto Factor 1 greatly represented the social aspects of positive emotions (i.e., how much students feel positive about their relationship with other peers); thus, Factor 1 was labeled as "Social Enjoyment." The other five items (loaded onto Factor 2) seemingly highlighted the students' inner perception of positive emotions (how much students feel positive about their own EFL learning in classrooms); thus, Factor 2 was labelled as "Private Enjoyment." Finally, Factor 3 was neatly loaded onto seven out of eight statements for Anxiety; Factor 3 was thus labelled as "Anxiety" (see Table 5). This three-factor structure reproduces the one uncovered by Dewaele and MacIntyre (2016).

Table 5 *Summary of a Three-Factor Solution Based on a Factor Analysis of the Emotion Questionnaire*

Items	Factor 1 (Social enjoyment)	Factor 2 (Private enjoyment)	Factor 3 (Anxiety)
A. Enjoyment			
• My class has a good atmosphere.	**.892**	-.086	.009
• My peers are nice.	**.804**	-.183	-.124
• We laugh a lot in class.	**.672**	.270	.102
• I always feel like there is a positive environment in my class.	**650**	.093	-.003
• English class is fun.	**.585**	.466	.027
• I'm a worthy member of my foreign language class.	-.212	**.836**	-.070
• I don't get bored in class.	.289	**.629**	.165
• I enjoy my foreign language class.	.454	**.557**	-.076
• It's cool to know English as a foreign language.	.134	**.533**	.346
• In class, I feel proud of my accomplishments.	.355	**.531**	-.028
B. Anxiety			
• I feel confident when I speak in my Foreign language class.	.205	**-.566**	.376
• I can feel my heart pounding when I'm going to be called on in my Foreign language class.	-.007	.091	**.832**
• I start to panic when I have to speak without preparation in my Foreign language class.	-.348	.211	**.761**
• I get nervous and confused when I am speaking in my Foreign language class	-.095	-.136	**.715**
• Even if I am well prepared for my Foreign language class, I feel anxious.	.031	.223	**.708**
• I am embarrassed to volunteer answers in my Foreign language class.	-.446	.035	**.626**
• I always feel that the other students in my class speak the Foreign language better than I do.	.494	-.300	**.617**
• I don't worry about making mistakes in my Foreign language class.	-.035	-.394	**.583**

All loadings > .5 were highlighted in bold.

Motivation vs. Emotion. To examine the relationship between the students' motivation and emotion, the resulting two motivation (Ideal L2 Self; Ought-to L2 Self) and three emotion (Social Enjoyment; Private Enjoyment; Anxiety) factor scores were then used for the following analyses. Pearson correlation analyses were performed to investigate the strength of the associations between the motivation and emotion factor scores (see Table 6). The participants' Ideal L2 Self was strongly correlated with the degree of their Private Enjoyment and inversely with the degree of Anxiety. Conversely, the participants' Ought-to L2 Self was unrelated to any aspect of their emotion at a $p < .025$ level (Bonferroni corrected). The results here suggested that when L2 students clearly visualize and internalize their future selves (i.e., Ideal L2 Self), they likely experience less negative emotions (Anxiety) and more positive emotions about their own EFL learning (i.e., Private Enjoyment).

Table 6 *Correlations between PCA Motivation and Emotion Dispositions*

	Social Enjoyment	Private Enjoyment	Anxiety
Ideal L2 Self	.21†	.48*	-.33*
Ought L2 Self	.10	.02	.07

Note. * indicates statistical significance at a $p < .025$ (Bonferroni corrected); † indicates marginal significance at $p < .051$ (Bonferroni corrected).

Cross-Sectional Investigations of Motivation, Emotion, Use and Achievement

Motivation, Emotion vs. Past/Recent Use. In this subsection, we first examined the extent to which the students' motivation and emotions could be related to their past L2 use prior to the project (preschool, elementary school and junior school) and their recent L2 use at the beginning of the project (high school). As shown in Table 7, the students' motivation factor scores demonstrated significant and marginally significant associations with their early EFL experience (Ideal/Ought-to L2 Self vs. the total hours of L2 practice at preschool and elementary school). Their positive emotion scores were significantly and marginally correlated with how much they recently engaged with the L2 inside and outside classrooms (Social/Private Enjoyment vs. L2 use during class; the total hours of preparation for class). The relationship between their negative emotion scores and previous/recent experience profiles did not reach statistical significance in any contexts.

Taken together, the results here indicated that motivation could be tied to previous experience; that positive emotion could be influenced by recent experience; and that negative emotion may be independent of particular EFL experience at specific time points (i.e., preschool vs. elementary vs. junior high schools).

Table 7 *Correlations between Previous and Recent Experience Variables and Motivation/Emotion Factors*

	Ideal L2 Self	Ought-to L2 Self	Social Enjoyment	Private Enjoyment	Anxiety
A. Past EFL experience					
• Total hours of L2 practice prior to the project	.13	.28*	-.01	.10	.01
• Total hours of L2 practice at preschool	.09	.23†	.06	.03	-.04
• Total hours of L2 practice inside EFL classroom at elementary school	.23†	.23†	.15	.16	-.12
• Total hours of L2 practice outside EFL classroom at elementary school (e.g., cram school)	.21	.20	.08	.10	-.09
• Total hours of L2 practice inside EFL classroom at junior high school	.02	.08	-.10	.01	.01
• Total hours of L2 practice outside EFL classroom (e.g., cram school) at junior high school	.01	.25*	-.07	.08	.09
B. Recent EFL experience at T1					
• L2 use during class (%)	.14	-.06	.22	.24†	-.14
• Total hours of all L2 practice outside EFL classrooms (per week)	.04	.17	.08	.09	-.01
• Total hours of preparation for class (per week)	-.01	.13	.26*	.10	-.01
• Total hours of cram school (per week)	.02	.12	-.18	-.01	.07
• Total hours of conversation activities with native and non-native speakers (per week)	.11	.02	-.12	.01	-.14

Note. * indicates statistical significance at a $p < .010$ (Bonferroni corrected); † indicates marginal significance at $p < .020$ (Bonferroni corrected).

Motivation, Emotion, Use vs. Achievement. Next, we conducted a series of Pearson correlation analyses to explore the students' long-term achievement (i.e., their comprehensibility scores at T1) relative to their motivation, emotion and experience profiles. The results summarized in Table 8 showed that the students' T1 performance was significantly and marginally correlated not only with their past (preschool) and recent (L2 use in classrooms) experience, but also with their motivation (Ideal L2 Self), positive (Private Enjoyment) and negative (Anxiety) emotion factors.

Table 8 *Correlations between Motivation, Emotion, Experience and Long-Term Achievement*

	Correlations with achievement	
	r	*p*
A. Motivation and emotion		
• Ideal L2 Self	.30	.002*
• Ought-to L2 Self	.01	.962
• Social Enjoyment	.09	.328
• Private Enjoyment	.24	.014*
• Anxiety	-.34	<.001*
B. Past EFL experience		
• Total hours of L2 practice prior to the project	.03	.753
• Total hours of L2 practice at preschool	.18	.058†
• Total hours of L2 practice inside EFL classroom at elementary school	.05	.562
• Total hours of L2 practice outside EFL classroom at elementary school (e.g., cram school)	.10	.286
• Total hours of L2 practice inside EFL classroom at junior high school	-.02	.819
• Total hours of L2 practice outside EFL classroom (e.g., cram school) at junior high school	-.01	.841
C. Recent EFL experience at T1		
• L2 use during class (%)	.22	.021*
• Total hours of all L2 practice outside EFL classrooms (per week)	.13	.167
• Total hours of preparation for class (per week)	.11	.224
• Total hours of cram school (per week)	.07	.447
• Total hours of conversation activities with native and non-native speakers (per week)	.02	.793

Note. * indicates statistical significance at a $p < .05$; † indicates marginal significance at $p < .10$.

To further examine the potentially different contributions of the motivation, emotion and experience factors to the participants' oral performance attainment, a multiple regression analysis was performed with their T1 comprehensibility scores as dependent variables and their learner intrinsic (motivation, emotion) and extrinsic (experience) factors as independent variables. To avoid multicollinearity problems, the decision was made to reduce the number of independent variables by using only significant and marginally significant predictors identified in the aforementioned correlation analysis—i.e., Ideal L2 Self, Private Enjoyment, total hours of L2 practice at preschool, and recent L2 use during class. According to the regression models (see Table 9), the students' T1 comprehensibility scores were mainly explained by Anxiety (10.6%), and to a lesser extent by Ideal L2 Self (4.8%) but without any clear evidence of multicollinearity (i.e., variance inflation factors [VIF] < 1.451).

Table 9 *Significant Results of Multiple Regression Analyses Using Motivation, Emotion and Past/Recent Experience as Predictors of L2 Comprehensibility at T1*

Predicted variable	Predictor variables	*Adjusted* R^2	R^2 *change*	*F*	*p*
Comprehensibility at T1	Anxiety	.106	.106	13.503	<.001
	Ideal L2 Self	.154	.048	10.598	<.001

Note. The variables entered into the regression equations included Ideal L2 Self; Private Enjoyment; Anxiety; total hours of practice at preschool; and recent L2 use during class.

In sum, the results here suggest that whereas the link between students' attained L2 oral proficiency and experience may be weak in EFL classrooms, their long-term achievement could be more clearly susceptible to the students' motivation and emotional differences; namely, certain L2 students with stronger Ideal L2 Self tend to have less anxiety towards L2 learning, which results in higher L2 oral proficiency after several years of EFL experience.

Longitudinal Investigation of Motivation, Emotion, Use and Development

Motivation, Emotion vs. Use. In this subsection, we take a longitudinal look at participants' motivation and emotion at T1 and the extent to which it could be predictive of their EFL experience inside and outside classrooms at T2. The Pearson correlation analysis (summarized in Table 10) reveals significant and marginally significant correlations between the participants' Private Enjoyment and the amount of L2 use inside (L2 use in class) and outside classrooms (extra L2 practice; conversations with other native and non-native speakers). Both the Ideal and Ought-to L2 Self aspects of motivation were equally related to, in particular, the students' L2 practice activities outside (but not inside) classrooms at T2. Similar to the cross-sectional analyses reported above, no significant correlations were identified between the students' negative emotion and experience profiles. To summarize, the results here hint that the degree of enjoyment, strongly tied to Ideal L2 Self motivation, could predict the quantity and quality of the most immediate and current EFL experience, whereas the predictive power of negative emotions for such experience seems to be weak.

Table 10 *Correlations between Most Current Experience Variables and Motivation/Emotion Factors*

	Ideal L2 Self	Ought-to L2 Self	Social Enjoyment	Private Enjoyment	Anxiety
Current EFL experience at T2					
• L2 use during class (%)	.18	-.08	.10	.21†	-.09
• Total hours of all L2 practice outside EFL classrooms (per week)	.22†	.25*	.03	.26*	-.02
• Total hours of preparation for class (per week)	.18	.09	.04	.17	-.01
• Total hours of cram school (per week)	.09	.24†	.01	.18	.02
• Total hours of conversation activities with native and non-native speakers (per week)	.26*	.21†	.03	.25*	-.13

Note. * indicates statistical significance at a $p < .010$ (Bonferroni corrected); † indicates marginal significance at $p < .020$ (Bonferroni corrected).

Motivation, Emotion, Use vs. Acquisition. Finally, we aimed to illustrate the extent to which the students' motivation/emotion (collected at T1) and experience (collected at T2) interacted to relate to their L2 comprehensibility development over the term (T1 → T2)—i.e., an index for short-term acquisition within one semester (three months). To this end, a set of partial correlation analyses were conducted with their gain scores over the term (T1→T2) as dependent variables, and their motivation/emotion and use profiles as independent variables. In order to isolate the influence of the participants' previous EFL experience (i.e., 3 to 14 years of EFL experience prior to the project), the students' comprehensibility scores at T1 were statically factored out as a covariate. As shown in Table 11, the students' longitudinal development of L2 comprehensibility was significantly associated with their motivation (Ideal L2 Self) and emotion (Private Enjoyment; Anxiety); and marginally correlated with their most current L2 use outside classrooms (cram schools, conversation activities).

To elucidate the relative weights of motivation, emotion and use in the participants' L2 comprehensibility development over one academic term, their gain scores were submitted to a multiple regression analysis with the motivation, emotion and experience factors as independent variables. Similar to the partial correlation analysis, the participants' comprehensibility scores at T1 were also selected as another independent variable as a way to control for the influence of the participants' previous EFL experience on their longitudinal development between T1 and T2. To avoid multicollinearity, only significant and marginally significant predictors identified in the partial correlation analyses—Ideal L2 Self, Private Enjoyment, Anxiety, total hours of cram school and conversation activities—as well as their T1 comprehensibility scores were entered into the model.

Table 11 *Partial Correlations between Motivation, Emotion, Experience and Development*

	Correlations with development	
	r	*p*
A. Motivation and emotion		
• Ideal L2 Self	.33	.001*
• Ought-to L2 Self	.01	.938
• Social Enjoyment	-02	.813
• Private Enjoyment	.33	.001*
• Anxiety	-.28	.006*
B. Current EFL experience at T2		
• L2 use during class (%)	.02	.885
• Total hours of all L2 practice outside EFL classrooms (per week)	.15	.124
• Total hours of preparation for class (per week)	-.03	.756
• Total hours of cram school (per week)	.19	.065†
• Total hours of conversation activities with native and non-native speakers (per week)	.19	.069†

Note. * indicates statistical significance at a $p < .05$; † indicates marginal significance at $p < .10$. Their initial proficiency (representative of their past EFL experience and proficiency) scores at T1 were partialled out.

The results (summarized in Table 12) showed that the model significantly explained 45.0% of variance in the participants' gain scores over the term (T1 → T2). More specifically, whereas their longitudinal SLA greatly built on their past EFL experience (T1 comprehensibility scores: 34.9%), participants who experienced stronger positive emotion in their current EFL learning experience (Private Enjoyment: 7.5%) and clearer future guides (Ideal L2 Self: 2.7%) actually showed more improvement in their oral proficiency even within a relatively short period of time (one academic term: three months), compared to peers with weaker positive emotion and motivation profiles. No evidence of strong multicollinearity was observed (VIF < 1.02).

Table 12
Significant Results of Multiple Regression Analyses Using Motivation, Emotion and Use as Predictors of L2 Comprehensibility at T1

Predicted variable	Predictor variables	*Adjusted* R^2	R^2 *change*	*F*	*p*
Gain scores in comprehensibility (T1 → T2)	Comprehensibility at T1	.349	.349	52.489	<.001
	Private Enjoyment	.423	.075	35.613	<.001
	Ideal L2 Self	.450	.027	26.225	< .001

Note. The variables entered into the regression equations included comprehensibility scores at T1; Ideal L2 Self; Private Enjoyment; Anxiety; total hours of extracurricular activities (cram school, conversations).

Discussion

The present study took an exploratory approach towards examining how first-year high school students in Japanese EFL classrooms enhanced their English L2 oral proficiency (analyzed via comprehensibility measures) according to different motivation, emotion and experience profiles. As predicted earlier, the questionnaires used in the study appeared to tap into five dimensions of the students' motivational and emotional orientations. They included the students' perception of their own desirable images (Ideal L2 Self) and others' imposed images (Ought-to L2 Self), their positive emotions resulting from the classroom atmosphere and their relations with teachers and peers (Social Enjoyment) and towards their individual and present EFL learning experience (Private Enjoyment), and the negative emotions linked to the learning and using of English as a foreign language in general (Anxiety). According to the results of the correlation analyses, Ideal L2 Self was, in particular, associated with Private Enjoyment and Anxiety, arguably because Ideal-self driven L2 learners likely have strong, well-defined self-guides, a promotional focus, and can easily orient towards future positive outcomes. As such, their greater motivational configurations elicit not only more enjoyment, but also less anxiety about EFL learning. In line with the recent motivation and emotion literature (Papi & Teimouri, 2014; Teimouri, 2016), the findings echo that having strongly internalized motivation (Ideal L2 Self) evokes various kinds of emotion (anxiety, enjoyment).

Table 13

Summary of Marginal and Significant Links between Motivation, Emotion, Experience and Acquisition

Motivation/emotion factors	Predicted experience factors[a]	Predicted acquisition factors
Ideal L2 Self	• Past use[a] • Current use[c]	• Achievement (T1 scores) • Development (T1/T2 gains)
Ought-to L2 Self	• Past use[a] • Current use[c]	*n.s.*
Social Enjoyment	• Recent use[b] • Current use[c]	*n.s.*
Private Enjoyment	• Recent use[b] • Current use[c]	• Development (T1/T2 gains)
Anxiety	*n.s.*	• Achievement (T1 scores)

Note. [a]Past use = elementary & junior high schools; [b]Recent use = high school; [c]Current use = T1-T2

In response to the first research question (i.e., the interaction patterns between L2 learners' motivation/emotion factors and their experience), our cross-sectional and longitudinal analyses demonstrated that both Ideal L2 Self and Ought-to L2 Self demonstrated small-to-medium correlations with various stages of the students' experience at preschool, junior high and high schools (r = .20-.30), indicating that motivation is closely aligned with L2 learners' behaviours to study, use and practice the target language throughout their EFL experience (Dörnyei, 2009; Ushioda, 2016). Interestingly, the degree of the students' positive feelings about their own EFL learning (Private Enjoyment) predicted the quantity/quality of their most recent experience (at both T1 and T2). However, their negative emotions (Anxiety) were not linked to any particular phase of their EFL experience.

Our findings on the differential effects of positive and negative emotions on experience throw a new light on previous cross-sectional research (Dewaele & MacIntyre, 2014, 2016; Dewaele, Witney, Saito & Dewaele, 2017). It suggested that positive and negative emotions may affect the L2 learning differently and over different timescales with enjoyment being more strongly predictive of how frequently L2 learners actually use the target language (see also Dewaele & Alfawzan, to appear). Negative emotions are known to build up gradually over time, they are more likely to become more permanent whereas positive emotions may be more fleeting and do not crystallise in either state-like or trait-like features (MacIntyre, 2017). Experiencing more enjoyment in the English class during their term may have been more emotionally salient for our participants and may have pushed anxiety to a more diffuse throbbing in the emotional background.

Turning to the second research question (i.e., the relative weights of motivation, emotion and experience in successful L2 speech learning), the students' comprehensibility at the beginning of the project (T1), which they had achieved after several years of EFL experience, was moderately correlated with their motivational (Ideal L2 Self) and emotional (Private Enjoyment, Anxiety) states, and weakly so with their recent experience during high school (L2 use inside classrooms). The results of multiple regression analyses suggest that the students' L2 oral performance was primarily explained by anxiety (10.6%) and secondarily by Ideal L2 Self (4.8%). The students' change in their comprehensibility between T1 and T2, which was indicative of the students' short-term acquisition within one academic semester, was similarly linked to motivation (Ideal L2 Self), emotion (Private Enjoyment, Anxiety) and current, extracurricular L2 use (total hours of cram school and conversation activities). According to the results of the multiple regression analyses, the students' L2 longitudinal development was explained by Private Enjoyment (7.5%) as well as by Ideal L2 Self (2.7%).

Overall, the results here replicated those of previous studies identifying the importance of L2 learners' most current and immediate experience with successful foreign language learning (Saito & Hanzawa, 2016; Ellis, 2006). However, it is probably more important to note that the study further revealed that this experience factor may be a necessary, but not sufficient condition for SLA; and that L2 learners' individual differences in motivation and emotion may greatly determine the extent to which they can ultimately make the most of and turn such experience into acquisition.

Supporting the theoretical claims of many L2 motivation scholars (Dörnyei, 2005, 2009; Ushioda, 2016), our study is the first attempt to provide empirical evidence that having strongly internalized future guides—Ideal L2 Self—could be a strong antecedent for successful SLA. One reason for this could be that motivation is believed to help L2 learners foster their awareness, noticing and understanding of input, especially when there is no input enhancement, as evidenced in the SLA (e.g., Takahashi, 2005) and social psychology (e.g., Newsome, 1986) literature. Ultimately, L2 learners' deeper engagement in language is thought to be the most important source of L2 speech learning (Flege, 2016).

Another reason could be that the emotions of such motivated, successful L2 learners appear to make unique contributions to the process and product of SLA in the short- and long-term. As shown in the present study, certain L2 learners with clearer vision of ideal future selves likely experience stronger positive emotions (Enjoyment) and weaker negative emotions (Anxiety) during their EFL learning. The L2 of these "positive" learners develops more quickly thanks to their current, enjoyable EFL experience. In contrast, students' anxiety was tied only to their long-term achievement but not to specific experience factors nor to short-term

development. The results here concur with the existing view that while L2 anxiety can fluctuate in the short term, it is relatively stable over longer periods time as a result of academic, cognitive and social causes throughout EFL experience (Dewaele & Dewaele, 2017; MacIntyre, 2017). The results here suggest that regular L2 use in positive emotional atmosphere directly impacts acquisition; and that accumulating such positive experience in various L2 learning contexts might decrease students' anxiety in the long run (Dewaele & MacIntyre, 2014), and significantly boosts their ultimate attainment in EFL classrooms.

Pedagogical Implications

As the present study highlighted the effects of motivation, emotion and experience profiles on L2 development, these findings have pedagogical implications. The results suggest that L2 learners' most immediate L2 learning experience may influence their SLA to some degree (significant experience effects were found only in the cross-sectional analyses). Yet, it is students with a predominant promotion focus (i.e., Ideal L2 Self) that most benefit from this experience, resulting in more tangible L2 speech development. In EFL classrooms, where many students typically have stronger Ought-to L2 Self and still consider learning English as an obligation (Li, 2014), teachers could incorporate motivation-enhancing activities into their lessons, specifically by adopting both promotion and prevention motivational strategies (Dörnyei & Kubanyiova, 2014).

As described earlier, our recent descriptive studies (e.g., Dewaele et al., 2017) have revealed that what teachers do in classrooms can strongly stimulate student enjoyment (but not necessarily reduce their anxiety). These effective practices include using the target language frequently; creating a friendly, pleasant and amusing atmosphere where students are constantly encouraged to use the target language without too much concern for making errors; and devising a range of interesting challenges involving risk-taking, autonomy and unpredictability beyond regular routine. Although researchers and teachers alike pay much attention to reducing L2 learners' anxiety, it is important to point out that anxiety is related to a combination of learner-internal and learner-external factors. In this sense, teachers should not be overly concerned about anxiety (which is not necessarily related to classroom practices), but should rather focus on boosting enjoyment (related to their L2 use and development) by making classroom environments adequately unpredictable, surprising and challenging for students.

Limitations

To close, several limitations need to be acknowledged with an eye towards further replication and elaboration of this topic. First, we need to acknowledge that the dataset drew on the performance of Japanese EFL students in the same high school, where much emphasis was put on the continuous improvement of English within the curriculum. According to our observations, teachers frequently used English as the medium of instruction, and students were constantly encouraged to use English (despite a wide range of variance in their L2 use in classrooms). In other words, these students were more likely to overcome the social anxiety that silences many Japanese L2 learners (King & Smith, 2017). Since our population can be said to represent "good language learners" (in a sense that they were recruited from the same prestigious high school with a great deal of emphasis on EFL education), the findings reported here need to be replicated with different groups of L2 learners with a wide range of proficiency levels (beginner, intermediate, advanced) and ages (young vs. adult learners) across different learning

contexts (foreign vs. second language learning), L1-L2 parings (linguistically similar vs. distant) and in a wide range of schools.

Second, the students' achievement and development was analyzed through L2 comprehensibility measures in the present study, as they were assumed to accurately reflect what they were practicing in EFL classrooms (how to make themselves successfully understood in oral communication). Alternatively, students' L2 speech learning could be examined in depth from various perspectives by adopting a set of linguistic measures tapping into multiple types of L2 knowledge which differ in linguistic domains (pronunciation, vocabulary, grammar: Saito et al., 2017) and processing modes (receptive vs. productive: Flege, 2016), with and without awareness (explicit vs. implicit: Moyer, 1999).

Third, the study found that all the motivation and emotion effects were exclusively tied to Ideal L2 Self, but not to Ought-to L2 Self. Whereas those with strong Ought-to L2 Self tended to practice the L2 in their past and current learning environments, such Ought-to driven experience did not seem to result in much L2 achievement/development. Nonetheless, such findings should be considered as tentative at best. Notably, certain L2 motivation scholars have begun to propose the re-examination of the theoretical framework of Ought-to L2 Self, as it does not cover the multifaceted nature of L2 learners' views of their duties, responsibilities and obligations related to SLA (e.g., Ought-to L2 Self own vs. others: Teimouri, 2016). Future studies may reconceptualise the theories and methods to examine the complex role of motivation and emotion in SLA.

Fourth, we assessed the participants' motivation and emotion related to L2 learning in general through existing questionnaires (Dewaele & MacIntyre, 2014; Dörnyei, 2005), and linked the scores to the global domain of their L2 oral proficiency (i.e., perceived comprehensibility, which is supposed to reflect the accurate and fluent use of various linguistic skills: Trofimovich & Isaacs, 2012). In the field of L2 speech research, certain scholars have recently elaborated constructs and instruments in order to tap into motivation specific to the "pronunciation" aspects of L2 oral proficiency (e.g., Sardegna, Lee, & Kusey, 2017 for *Learner Attitudes for Pronunciation Inventory*; Baran-Łucarz, 2014 for the *Pronunciation Anxiety Questionnaire*). Smit (2002) provided cross-sectional evidence that L2 learners who demonstrated stronger intrinsic (rather than extrinsic) motivation and less anxiety for L2 pronunciation learning obtained better final grades in L2 English pronunciation classes. To probe the relationship between motivation, emotion and L2 oral proficiency at a fine-grained level, future studies could investigate the extent to which such pronunciation-focused motivation and emotion influence L2 segmental, prosodic and temporal learning over time.

Finally, given that our study was one of the first attempts to provide quantitative evidence on the triangular relationship between L2 motivation/emotion, use and L2 speech learning, the findings need to be complemented by qualitative perspectives. L2 motivation researchers argue that L2 motivation is a complex, dynamic, emergent and adaptive system, as it constantly develops according to a wide range of contextual, cultural and psychological factors (e.g., Dörnyei, 2009). To this end, it is necessary for future studies to use mixed-method approaches to provide detailed description of how each individual learner's motivation and emotions change over time; and how such emotions and motivation growth interacts with the students' learning contexts (cf. Nagle, 2018; Waninge, Dörnyei, & de Bot, 2014).

References

ALC Press Inc. (2017). Telephone Standard Speaking Test. Retrieved from https://tsst.alc.co.jp/tsst/e_index.html April 2017.

Baker-Smemoe, W., & Haslam, N. (2013). The effect of language learning aptitude, strategy use and learning context on L2 pronunciation learning. *Applied Linguistics, 34*, 435-456.

Baran-Łucarz, M. (2014). The link between pronunciation anxiety and willingness to communicate in the foreign-language classroom: The Polish EFL context. *Canadian Modern Language Review*, *70*, 445-473.

Boersma, D., & Weenink, P. (2017). *Praat: Doing phonetics by computer version 6.0.29.* Retrieved from http://www.praat.org April 2017.

Boo, Z., Dörnyei. Z., & Ryan, S. (2015). L2 motivation research 2005-2014: Understanding a publication surge and a changing landscape. *System, 55*, 147-157.

Csizér, K., & Kormos, J. (2009). Learning experiences, selves and motivated learning behaviour: a comparative analysis of structural models for Hungarian secondary and university learners of English. In Z. Dörnyei, & E. Ushioda (Eds.), *Motivation, language identity and the L2 self*. (pp. 98-117). Bristol: Multilingual Mattters

Derwing, T. M. & Munro, M. J. (2015). *Pronunciation fundamentals: Evidence-based perspectives for L2 teaching and research*. Amsterdam: John Benjamins.

Derwing, T. M., Munro, M. J., & Thomson, R. I. (2008). A longitudinal study of ESL learners' fluency and comprehensibility development. *Applied Linguistics*, *29*, 359-380.

Dewaele, J.-M. & Alfawzan, M. (to appear). Anxiety vs Enjoyment: which has a greater influence on foreign language performance? Special issue *Emotions in SLA* J.-M. Dewaele (guest ed.), *Studies in Second Language Learning and Teaching*

Dewaele, J.-M. & Dewaele, L. (2017). The dynamic interactions in Foreign Language Classroom Anxiety and Foreign Language Enjoyment of pupils aged 12 to 18. A pseudo-longitudinal investigation. *Journal of the European Second Language Association, 1*, 12-22. doi: 10.22599/jesla.6

Dewaele, J.-M. & MacIntyre, P. D. (2014). The two faces of Janus? Anxiety and Enjoyment in the Foreign Language Classroom. *Studies in Second Language Learning and Teaching*, 4, 237-274.

Dewaele, J.-M. & MacIntyre, P. (2016). Foreign Language Enjoyment and Foreign Language Classroom Anxiety. The right and left feet of FL learning? In P. MacIntyre, T. Gregersen & S. Mercer (Eds.), *Positive Psychology in SLA*. (pp. 215-236). Bristol: Multilingual Matters.

Dewaele, J.-M., Witney, J., Saito, K. & Dewaele, L. (2017). Foreign language enjoyment and anxiety in the FL classroom: the effect of teacher and learner variables. *Language Teaching Research.* doi: 10.1177/1362168817692161

Dörnyei, Z. (2005). *The psychology of the language learner: Individual differences in second language acquisition*. London: Routledge.

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). *Motivation, language attitudes and globalisation: A Hungarian perspective*. Clevedon: Multilingual Matters.

Dörnyei, Z., & Kubanyiova, M. (2014). *Motivating learners, motivating teachers: Building vision in the language classroom*. Cambridge: Cambridge University Press.

Dörnyei, Z., & Ushioda, E. (Eds.). (2009). *Motivation, language identity and the L2 self*. Clevedon: Multilingual Matters.

Ellis, N. C. (2006). Language acquisition as rational contingency learning. *Applied Linguistics. 27*, 1-24.

Eiken Foundation of Japan. (2017). Comparison table. Retrieved from http://stepeiken.org/comparison-table

Flege, J.E. (2016). *The role of phonetic category formation in second language speech acquisition*. Eight International Conference on Second Language Speech, 10 June, 2016, Aarhus University, Denmark.

Flege, J., Munro, M, & MacKay, I. R. A. (1995). Factors affecting degree of perceived foreign accent in a second language. *Journal of the Acoustical Society of America, 97*, 3125-3134.

Gkonou, C., Daubney, M., & Dewaele, J. M. (2017). *New Insights into Language Anxiety: Theory, Research and Educational Implications*. Multilingual Matters.

Gregersen, T. S. (2003). To err is human: A reminder to teachers of language-anxious students. *Foreign Language Annals*, *36*, 25-32.

Gregersen, T., MacIntyre, P. D., & Meza, M. D. (2014). The motion of emotion: Idiodynamic case studies of learners' foreign language anxiety. *The Modern Language Journal*, *98*, 574-588.

Horwitz, E. K. (2017). On the misreading of Horwitz, Horwitz, and Cope (1986) and the need to balance anxiety research and the experiences of anxious language learners. In C. Gkonou, M. Daubney, & J.-M. Dewaele (Eds.), *New Insights into Language Anxiety: Theory, Research and Educational Implications* (pp. 31-47). Bristol: Multilingual Matters.

Horwitz, E., Horwitz, M., & Cope, J. (1986). Foreign language classroom anxiety. *Modern Language Journal*, *70*, 125-132.

Higgins, E. T. (1987). Self-discrepancy: A theory relating self and affect. *Psychological Review*, *94*, 319-340.

Isaacs, T., & Thomson, R. I. (2013). Rater experience, rating scale length, and judgments of L2 pronunciation: Revisiting research conventions. *Language Assessment Quarterly*, *10*, 135-159.

King, J., & Smith, L. (2017). Social anxiety and silence in Japan's tertiary foreign language classrooms. In C. Gkonou, M. Daubney, & J-M. Dewaele (Eds.), *New insights into language anxiety: Theory, research and educational implications* (pp. 91-109). Bristol: Multilingual Matters.

Kormos, J., Kiddle, T., & Csizér, K. (2011). Systems of goals, attitudes, and self-related beliefs in second-language-learning motivation. *Applied Linguistics*, *32*, 495-516.

Lamb, M. (2012). A self system perspective on young adolescents' motivation to learn English in urban and rural settings. *Language learning*, *62*, 997-1023.

Larson-Hall, J. (2008). Weighing the benefits of studying a foreign language at a younger starting age in a minimal input situation. *Second Language Research, 24*, 35-63.

Li, Q. (2014). Differences in the motivation of Chinese learners of English in a foreign and second language context. *System, 42,* 451-461.

Loewen, S., & Gonulal, T. (2015). Exploratory factor analysis and principal components analysis. In Plonsky, L. (Ed), *Advancing quantitative methods in second language research*. New York: Routledge.

MacIntyre, P. D. (2017). An overview of language anxiety research and trends in its development. In C. Gkonou, M. Daubney, & J-M. Dewaele (Eds.), *New insights into language anxiety: Theory, research and educational implications* (pp. 11-30). Bristol: Multilingual Matters.

MacIntyre, P. D., & Gardner, R. C. (1989). Anxiety and second-language learning: Toward a theoretical clarification. *Language Learning*, *39*, 251-275.

MacIntyre, P., & Gregersen, T. (2012). Emotions that facilitate language learning: The positive-broadening power of the imagination. *Studies in Second Language Learning and Teaching*, *2*, 193-213.

Markus, H., & Nurius, P. (1986). Possible selves. *American Psychologist*, *41*, 954-969.

Mora, J. C., & Valls-Ferrer, M. (2012). Oral fluency, accuracy, and complexity in formal instruction and study abroad learning contexts. *TESOL Quarterly*, *46*, 610-641.

Moskovsky, C., Assulaimani, T., Racheva, S., & Harkins, J. (2016). The L2 motivational self system and L2 achievement: A study of Saudi EFL learners. *The Modern Language Journal*, *100*, 641-654.

Moyer, A. (1999). Ultimate attainment in L2 phonology: The critical factors of age motivation and instruction. *Studies in Second Language Acquisition*, *21*, 81-108.

Muñoz, C. (2014). Contrasting effects of starting age and input on the oral performance of foreign language learners. *Applied Linguistics, 35*, 463-482.

Muñoz, C. & Llanes, À. (2014). Study abroad and changes in degree of foreign accent in children and adults. *The Modern Language Journal*, *98*, 432-449.

Nagle, C. (2018). Motivation, comprehensibility, and accentedness in L2 Spanish: Investigating motivation as a time-varying predictor of pronunciation development. *Modern Language Journal*. DOI: 10.1111/modl.12461

Newsome III, G. L. (1986). The effects of reader perspective and cognitive style on remembering important information from texts. *Journal of Reading Behavior*, *18*, 117-133.

Ortega, L., & Byrnes, H. (Eds.). (2008). *The longitudinal study of advanced L2 capacities*. New York: Routledge.

Papi, M., & Teimouri, Y. (2014). Language learner motivational types: A cluster analysis study. *Language Learning*, *64*, 493-525.

Plonsky, L., & Oswald, F. L. (2017). Multiple regression as a flexible alternative to ANOVA in L2 research. *Studies in Second Language Acquisition*, *39*, 579-592.

Ryan, S. (2008). *The ideal L2 selves of Japanese learners of English*. (Unpublished doctoral dissertation). University of Nottingham, Nottingham, UK.

Sardegna, V. G., Lee, J., & Kusey, C. (2017). Self-efficacy, attitudes, and choice of strategies for English pronunciation learning. *Language Learning*. doi:10.1111/lang.12263

Saito, K. (2015). Experience effects on the development of late second language learners' oral proficiency. *Language Learning, 65*, 563-595.

Saito, K., Dewaele, J.-M., & Hanzawa, K. (2017). A longitudinal investigation of the relationship between motivation and late second language speech learning in classroom settings. *Language and Speech*. doi: 10.1177/0023830916687793

Saito, K., & Hanzawa, K. (2016). Developing second language oral ability in foreign language classrooms: The role of the length and focus of instruction and individual differences. *Applied Psycholinguistics, 37*, 813-840.

Saito, K., & Hanzawa, K. (2017). The role of input in second language oral ability development in foreign language classrooms: A longitudinal study. *Language Teaching Research*. doi: 10.1177/1362168816679030

Saito, K., & Munro, M. (2014). The early phase of /r/ production development in adult Japanese learners of English. *Language and Speech, 57*, 451-469.

Saito, K., Trofimovich, P., & Isaacs, T. (2017). Using listener judgements to investigate linguistic influences on L2 comprehensibility and accentedness: A validation and generalization study. *Applied Linguistics, 38,* 439-462.

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. (2015). Foreign language aptitude and its relationship with grammar: A critical overview. *Applied Linguistics, 36*, 367-384.

Smit, U. (2002). The interaction of motivation and achievement in advanced EFL pronunciation learners. *IRAL, 40*, 89-116.

Spada, N., & Tomita, Y. (2010). Interactions between type of instruction and type of language feature: A meta-analysis. *Language Learning, 60,* 263-308.

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). Clevedon: Multilingual Matters.

Takahashi, S. (2005). Pragmalinguistic awareness: Is it related to motivation and proficiency? *Applied Linguistics*, *26*, 90-120.

Teimouri, Y. (2016). L2 Selves, emotions, and motivated behaviours. *Studies in Second Language Acquisition*. Advanced publication. doi: 10.1017/S0272263116000243

Trofimovich, P., & Isaacs, T. (2012). Disentangling accent from comprehensibility. *Bilingualism: Language and Cognition*, *15*, 905–916.

Ushioda, E. (2016). Language learning motivation through a small lens: A research agenda. *Language Teaching*, *49*, 564-577.

Waninge, F., Dörnyei, Z., & de Bot, K. (2014). Motivational dynamics in language learning: Change, stability, and context. *The Modern Language Journal*, *98*, 704-723.

Yashima, T., Zenuk-Nishide, L., & Shimizu, K. (2004). The influence of attitudes and affect on willingness to communicate and second language communication. *Language Learning*, *54*, 119-152.

Supporting Information-A: Details of L2 Use inside Classrooms

During the time of the project (December 2016 – March 2017), the participating students in the present study were enrolled in seven 50-minute EFL lessons per week—(a) five lessons entitled ***General English*** and (b) two lessons entitled ***English Production***. These classes were taught by the same three teachers throughout the academic year, incorporating a number of comprehension and production tasks to apply newly learned grammar points and useful expressions to real communication. We had full access to all the standard syllabus and lesson plans for General English and English Production, and conducted classroom observations for all the instructors on February 8th, 2017. In what follows, we provided a detailed description of the two different EFL classes that the participating students engaged during the time of the project.

General English (*n* = 5 lessons per week)
The content and foci of General English were mainly based on textbook materials where each lesson was designed to enhance students' accurate and fluent use of certain phonological, lexical and grammatical structures. For each lesson, the instructors began with an explanation of target linguistic features, followed by a range of comprehension-based (reading and listening passages) and production-based (e.g., translation, dictation, discussion, crossword puzzles) activities. To enhance the students' interests in using the target language for meaningful communication, the students were also encouraged to use English to discuss a range of familiar and interesting topics in geography, psychology and sociology—a concept comparable to Content and Language Integrated Learning. According to our classroom observations, the instructors constantly provided support to their students during the comprehension activities, and gave an adequate amount of corrective feedback during the production activities.

English Production (*n* = 2 lessons per week)
To follow up and reinforce what the students learned in General English (5 lessons per week), English Production (2 lessons per week) mainly offered more speaking and writing practice opportunities. The students engaged in oral communication in pairs and groups through a wide variety of communication tasks, such as debates, interviews, role plays and monologue. At the same time, students were encouraged to pay attention to the accurate and fluent use of the target language while using L2 English for meaningful purposes. In addition, the students also took part in structured and free writing tasks in order to learn the various crucial components of L2 English writing processes, such as word choice, linking words, grammar usage and overall organization. Throughout these opportunities, the students were encouraged to build a sense of confidence and achievement by carrying out various kinds of commands in L2 English.

Supporting Information-B: EFL Experience Questionnaire

English Version (Translated)

【English Use Inside School】

1) What is your most recent exam score in English class?

Name of class：_____ （_____/100 points）

Name of class：_____ （_____/100 points）

2）How much do you use English during class? （e.g., presentation, conversation, discussion）

Name of class：_____ I use English （_____）% of the time during class.

Name of class：_____ I use English （_____）% of the time during class

【English Use Outside School】

3）How many hours do you study at a cram school <u>outside</u> school?

Approximately （_____）hours per week

4）How many hours do you prepare for your class <u>outside</u> school?

Approximately （_____）hours per week for speaking-related activities (e.g., practicing presentation)

Approximately （_____）hours per week for non-speaking activities (e.g., memorizing vocabulary, reading aloud, studying for college entrance exams)

5）How many hours do you voluntarily spend speaking English?

Approximately （_____）hours per week with native speakers of English (from the USA, the UK, Australia etc.)

What kinds of activities do you generally do (e.g., practicing speaking and presentation, free conversations)?

Approximately （_____）hours per week with non-native speakers of English (e.g., advanced-level Korean, Chinese and Japanese learners of English)

What kinds of activities do you generally do (e.g., practicing speaking and presentation, free conversations)?

【Previous English Learning Experience】

6) When did you start learning English?

_____years old

Where? (at elementary/junior high school, conversation school etc.)

_____

7) Please tell us about your previous English learning experience

Preschool _____ hours per week in class; _____ hours per week outside class

Elementary school _____ hours per week in class; _____ hours per week outside class

Junior high school _____ hours per week in class; _____ hours per week outside class

8) Have you been abroad more than 10 days (other than short family trips)?

yes / no

If yes, where did you visit and what did you do?

When? _____ years old Where? _____ (city, country) How long? _____ days/months/years
Why (e.g., study abroad, homestay)? _____

When? _____ years old Where? _____ (city, country) How long? _____ days/months/years
Why (e.g., study abroad, homestay)? _____

When? _____ years old Where? _____ (city, country) How long? _____ days/months/years
Why (e.g., study abroad, homestay)? _____

Japanese Version (Used for the Present Study)

【学校内における英語学習について調査】

1)「最新」のテスト結果を教えてください。

クラス名：_____（_____点）

クラス名：_____（_____点）

2) 授業中にどのくらい英語を話していますか？（プレゼン・会話・ディスカッションなど）

クラス名：_____ 授業中（_____）％は、英語で話している。

クラス名：_____ 授業中（_____）％は、英語で話している。

【学校外における英語学習についての調査】

3) 授業以外に、どのくらい塾で勉強していますか？

一週間に（ ）時間くらい

4）授業以外に、どのくらい学校の授業の復習・予習・宿題をしていますか？

スピーキングに関する勉強時間（プレゼンの練習など）：一週間に（　　　　）時間くらい

スピーキング以外の勉強時間（単語ドリル・音読・入試勉強など）：一週間に（　　　　）時間くらい

5）授業以外で、自主的に英語話者と、どのくらい英語を話しますか？（英会話学校・クラブ活動も含む）

ネィティブスピーカー（アメリカ・イギリス・オーストラリア人など）と一週間に（　　　　）時間くらい話します。

具体的に何をしていますか（スピーチの練習、会話の練習など）：

ノンネィティブスピーカー（英語を流暢に話す日本・韓国・中国人など）と一週間に（　　　　）時間くらい話します。

具体的に何をしていますか（スピーチの練習、会話の練習など）：

【外国語学習の経験について】

6) 英語を勉強し始めたのはいつですか？

何歳ごろ　　　　　　　　　　　　　______________________歳ごろ

どの様に（学校で・英会話学校・塾など）　____

7) あなたの過去の英語学習の経験について教えてください

小学生のころ　　　　学校の授業は一週間に（　　　　）時間くらい；塾での勉強は一週間に（　　　　）時間くらい

中学生のころ　　　　学校の授業は一週間に（　　　　）時間くらい；塾での勉強は一週間に（　　　　）時間くらい

8) これまでにあなたは 10 日間以上の海外滞在・留学をした経験がありますか（10 日間未満の家族旅行などは除く）？（〇で囲んでください）

yes / no

yes の場合、具体的にはどこで何を勉強しましたか： いつ______歳のとき　どこで________（都市・国名）どの位滞在しましたか　______日・月・年　どの様な事をしましたか（留学・ホームステイ）　______________________________ いつ______歳のとき　どこで________（都市・国名）どの位滞在しましたか　______日・月・年　どの様な事をしましたか（留学・ホームステイ）　______________________________ いつ______歳のとき　どこで________（都市・国名）どの位滞在しましたか　______日・月・年　どの様な事をしましたか（留学・ホームステイ）　______________________________

Supporting Information-C: Training Scripts and Onscreen Labels of L2 Comprehensibility Judgements

A. Training scripts for comprehensibility judgement

Comprehensibility	This term refers to how much effort it takes to understand what someone is saying. If you can understand with ease, then a speaker is highly comprehensible. However, if you struggle and must listen very carefully, or in fact cannot understand what is being said at all, then a speaker has low comprehensibility.

B. Onscreen labels