Running-head
INDIVIDUAL DIFFERENCES IN L2 PRONUNCIATION

Title
Detangling Experiential, Cognitive, and Sociopsychological Individual Differences in Second Language Speech Learning: Cross-Sectional and Longitudinal Investigations

DOI: https://doi.org/10.1017/S1366728922000700

Authors
Yui Suzukida
Juntendo University / University College London
y.suzukida.fz@juntendo.ac.jp
yui.suzukida.18@ucl.ac.jp

Kazuya Saito
University College London
k.saito@ucl.ac.uk

Acknowledgments:
This research was supported by the Language Learning Dissertation Grant. We would like to thank Andrea Revesz, Joan Carles Mora, and Viktoria Magne for their constructive feedback. We also thank Talia Issacs, Pavel Trofimovich, Hui Sun, and Xiaojun Lu for their assistance and feedback in the process of data collection and analysis.

Corresponding Author Info:
Yui Suzukida

Juntendo University
School of Medicine
Department of General Education,
1-1 Hiraga-gakuendai, Inzai, Chiba
Post code: 270-1695
y.suzukida.fz@juntendo.ac.jp

University College London
20 Bedford Way, London
WC1H 0AL
United Kingdom
yui.suzukida.18@ucl.ac.jp

Key words: L2 pronunciation, individual differences, aptitude, anxiety, motivation, English as a Foreign Language

Abstract

In this two-part study, we conducted both cross-sectional and longitudinal investigations on the relative weights of experiential, cognitive, and sociopsychological factors in adult L2 speech learning. In the cross-sectional phase (Study 1), speech was elicited from 73 Japanese speakers of English via a picture description task, and rated for accentedness and comprehensibility. These scores were linked to scores on a range of tests designed to measure aptitude, motivation, and anxiety. The results showed that comprehensibility was exclusively linked to experiential variables (e.g., the amount of L2 use outside classrooms), while accentedness was linked to phonemic coding ability and anxiety. In the longitudinal phase (Study 2), we tracked the same participants' L2 comprehensibility and accentedness development when they received four weeks of explicit pronunciation instruction. According to the results of pre- and post-tests, participants significantly improved the comprehensibility and accentedness of their speech regardless of cognitive and sociopsychological differences.

In the area of second language (L2) speech research, many scholars have sought to understand which factors contribute to the process and product of learners' successful phonological acquisition (Trofimovich, Kennedy, & Foote, 2015). A large number of studies have reported that L2 speech outcomes are strongly linked to the quantity and quality of a learner's L2 experience (i.e., more practice is better) (Flege, 2016 for overviews), to aptitude (Baker-Smemoe & Haslam, 2013) and to sociopsychological factors such as motivation (Liu & Huang, 2011). However, few of these studies have justified their selection of IDs using a theoretical model, or have exclusively focused on either the cognitive or sociopsychological aspects of IDs. The current study departed from this trend, and sought to unravel the complexities of classroom-based L2 pronunciation learning from a DST perspective. In the context of 73 college-level Japanese speakers of English, we conducted both cross-sectional and longitudinal investigations on the relative weights of experiential, cognitive, and sociopsychological factors in adult L2 speech learning. In the cross-sectional phase (Study 1), we examined the relationship between participants' experiential, cognitive and sociopsychological profiles and two different aspects of L2 oral proficiency, i.e., comprehensibility (i.e., how difficult it is to understand what the speaker is saying) and accentedness (i.e., how heavily a speaker's speech is affected by his/her native language; Derwing & Munro, 2013). In the longitudinal phase (Study 2), we tracked the same participants' L2 comprehensibility and accentedness *development,* when they received four weeks of explicit pronunciation instruction.

Background

Individual differences in SLA research

Over the past 50 years, much scholarly attention has been given to examining how the process and product of L2 learning is characterized by various contextual, experiential, cognitive and sociopsychological factors. Although existing studies tend to focus on either cognitive or psychological aspects, little attempts have been made to investigate IDs *holistically* by investigating both at the same time (Serafini, 2017). However, scholars have begun to call for a more integrative approach with which to explore how individual learners with varied profiles of experience, cognition, motivation, and emotion can develop different dimensions of language (e.g., Ortega, 2013). One such framework could be Dynamic Systems Theory (DST). DST is an approach, or a *meta-theory* (Larsen–Freeman, 2013), that consists of a set of principles for exploring the changes in complex systems. The theory holds that such changes are sensitive to initial states, are resource-dependent, non-linear, and exhibit emergent outcomes when systems stabilize at attractor states (e.g., de Bot, 2008). A particular system consists of multiple components, and the interaction between the components characterize the state of the system (de Bot, 2008). Identifying the operating rules of these components allows for robust interpretations to be made about system behavior. From a DST perspective, learner-external and learner-internal factors can be considered to be components that shape developmental changes in language systems (e.g., Hiver & Al-Hoorie, 2016).

Another integrative approach towards individual differences concerns cognitive psychologists' account of the human mind, i.e., *the trilogy of mind*. Under this view, human intellectual functioning consists of motivation, emotion, and cognition (e.g., Matthews & Zeidner, 2004). Researchers typically categorize learner-internal IDs into these three domains (i.e., cognition, motivation, and emotion), and stress that it is crucial to give them each equal attention (Waninge, 2015). Furthermore, in the context of L2 *pronunciation* research, Moyer (2014) has shown that L2 learners who can produce near-nativelike L2 pronunciation often show superior scores on *multiple* IDs (both cognitive and sociopsychological IDs), suggesting a synergistic effect in the context of L2 pronunciation learning.

Following these lines of thoughts, we propose that L2 pronunciation acquisition can be conceptualized as a multidimensional and complex phenomenon. To detangle its complex mechanisms, the current study took a first step towards exploring how both cognitive, sociopsychological IDs dynamically interact to shape two different dimensions of the L2

pronunciation learning process (comprehensibility vs. accentedness) from multiple angles (cross-sectional vs. longitudinal).

Roles of individual differences in second language pronunciation learning

To date, researchers have extensively examined a range of IDs hypothesized to predict success in L2 pronunciation learning. For example, many studies have explored the role of different cognitive abilities in attaining advanced L2 pronunciation perception and/or production performance. Variables investigated to date have included working memory (e.g., Hu et al., 2013), attention control (Darcy et al., 2015), musical aptitude (Li & Dekeyser, 2017), domain general auditory processing (Saito, Sun, & Tierney, 2020), foreign language aptitude (Saito & Hanzawa, 2016) and personality profiles (Hu & Reiterer, 2009). Other scholars have suggested that social and psychological factors impact learning. For instance, factors such as ethnic group affiliation (Gatbonton & Trofimovich, 2008), contextual attitude (Huensch & Thompson, 2017), language awareness (Kennedy & Trofimovich, 2010), motivation to learn an L2 (e.g., Nagle, 2018a), and degree of anxiety towards learning an L2 (Baran-Łucarz, 2016; Sardegna et al., 2014) have been found to affect pronunciation attainment and performance. In what follows, we provide a selective overview of past research evidence on IDs in relation to L2 pronunciation learning in the classroom setting.

Foreign language aptitude

Many scholars have attributed exceptionality in L2 pronunciation to some underlying *talent*, what researchers have called aptitude (e.g., Muñoz & Singleton, 2007). Foreign language learning aptitude refers to the set of specialized cognitive factors that are thought to play a role in language learning (Li, 2016). According to Carroll's (1962) influential model, aptitude consists of phonemic coding ability, grammatical sensitivity, inductive learning, and associative memory. To respond to the growing interest in both implicit and explicit learning aptitudes (Suzuki & DeKeyser, 2015), several post-MLAT (Carroll & Sapon, 1959) batteries have been developed, including the LLAMA (Meara, 2005), the CANAL-F test (Grigorenko et al., 2000), and Hi-LAB (Doughty et al., 2010). Among these, the LLAMA tests have been widely used in the field of SLA to measure both implicit (sound sequence recognition) and explicit (associative memory, phonemic coding and grammatical inferencing) learning aptitude (Granena, 2013). Cross-sectional and longitudinal studies of aptitude suggest that (a) different explicit learning aptitudes work on different aspects of L2 speech development, and (b) explicit and implicit aptitudes determine different stages of speech development (Baker-Smemoe & Haslam, 2013; Hu et al., 2013; Saito & Hanzawa, 2016). Saito and Hanzawa (2016) reported that Japanese L2 English learners' aptitude scores (a composite of four sub-tests measured via LLAMA) showed positive correlations with segmental, word stress, and speech rate ratings obtained from native raters. Baker-Smemoe and Haslam (2013) examined the relationship between L2 learners' pronunciation proficiency (operationalized as production accuracy, reduced accentedness, and fluency) and aptitude (as well as motivation and various strategies). They also found that sound discrimination ability (measured via the PLAB) was associated with reduced accentedness, and that higher comprehensibility was predicted by higher motivation and the use of various learning strategies. Similarly, Hu et al. (2013) found that higher phonemic coding ability predicts better L2 pronunciation performance. More recent work has suggested that (a) phonemic coding ability (measured by the LLAMA E, B) and rote memory contributed to quick improvements in accuracy and fluency; and (b) sound sequence recognition (measured via the LLAMA D) facilitated comprehensibility in the long run by enhancing their accurate production of segmentals (Saito et al., 2019). Such evidence indicates that sound sequence recognition may also tap into L2 learners' implicit learning aptitude.

Motivation

Motivation is believed to play a role in initiating and maintaining learners' efforts to learn an L2 (e.g., Gardner, 2007). Researchers have found that learners' motivation, and especially their concerns for native-like L2 pronunciation, are a key predictor of reduced foreign accent (Gonet,

2006; Moyer, 2014). For example, Gonet's (2006) classroom study of Polish English as a Foreign Language (EFL) learners found that motivation was the strongest contributor to L2 pronunciation acquisition.

Recently, Dörnyei's (2005) L2 Motivational Self System has been increasingly used to explore different motivational orientations, learning behaviors, and learning outcomes in the FL classroom setting (Dörnyei & Chan, 2013). The model consists of two components, or self-guides: the *Ought-to L2 self* (i.e., imposed self-image related to obligation and avoidance) and the *Ideal L2 self* (idealized self-image of an L2 user). Both components are considered to be closely associated with the extent to which learners are committed to studying, practicing, and using an L2 for an extensive period of time (e.g., Ushioda, 2016). Furthermore, higher levels of Ideal L2 self have been linked with positive L2 learning outcomes (e.g., Dörnyei & Chan, 2013). In L2 *pronunciation* research, however, only a handful of studies examined the link between possible selves and L2 speech performance (e.g., Nagle, 2018a; Saito et al., 2018). Saito et al. (2018) found a link between the two self-guides and L2 experience, but also found a positive correlation between higher Ideal L2 self and comprehensibility. Based on these findings, the authors suggested that Ideal L2 self may be a key factor for enhancing information processing, and helping them make the most of the available opportunities of receiving input and producing speech in L2. However, as the available research evidence is limited (e.g., Nagle, 2018a), further research in the EFL setting is required to confirm the robust influence of self-guides on L2 pronunciation learning.

Anxiety

Another factor worthy of attention in L2 pronunciation learning is anxiety. Since Horwitz and colleagues' development of the Foreign Language Classroom Anxiety Scale (FLCAS; Horwitz et al.,1986), learners' anxiety in the classroom has been explored as a predictor of L2 performance (e.g., for a meta-analysis see Teimouri et al., 2019). According to Baran-Łucarz (2016), L2 pronunciation learning engenders a specific form of anxiety due to the perceived discrepancy between a learner's current pronunciation and the level of pronunciation they expect/desire to reach. Moreover, learners' self-perception of their pronunciation skill or their willingness to accept target-like pronunciation and modify their own pronunciation is believed to result in some changes to their actual behaviors (Baran-Łucarz, 2016). Therefore, more recently, scholars have begun to conceptualize an anxiety unique to pronunciation learning, identifying it either as Measure of Pronunciation Anxiety in the FL Classroom) (Baran-Łucarz, 2016), or as part of the Learner Attitudes and Motivations for Pronunciation inventory (Sardegna et al., 2014).

Research in the field of cognitive psychology has suggested that anxiety influences the cognitive, psychological, and behavioral aspects of learning. For instance, high anxiety has been shown to decrease the efficiency of cognitive functioning during task execution, can lead to panic and shakiness, and can result in task avoidance (e.g., Vasa & Pine, 2004). Because anxiety can hinder one's attention control, it is believed to deteriorate language learners' ability to receive and process input, and to produce output (Piechuurska-Kuciel, 2008). These negative impacts have been extended to L2 pronunciation learning as well (Baran-Łucarz, 2013). While pronunciation specific anxiety has been explored in relation to learners' self-rated proficiency (e.g., Szyszka, 2011), only a few empirical studies have explored proficiency as rated by others (cf. Saito et al., 2018). For example, Saito et al. (2018) found that anxiety, measured via the FLCAS, was significantly correlated with comprehensibility. Their findings not only support the assertion that anxiety is an emotion that is shaped through the accumulations of one's learning experience over time (Dewaele & Dewaele, 2017), but also shed light on the possible impact of negative emotions on pronunciation learning. However, more studies are needed to fully understand the relationship between anxiety and L2 pronunciation, particularly those which seek to identify how *pronunciation specific* anxiety influences L2 pronunciation learning.

Motivation for current study

As reviewed above, previous research has explored various cognitive and sociopsychological IDs as potential predictors of L2 pronunciation learning success. However, there is little crosstalk between the two different groups of ID researchers. In other words, we have yet to know how both cognitive and sociopsychological factors interact to impact different dimensions of L2 acquisition. One exception to this is Serafini (2017), which adopted a DST framework and took a longitudinal approach towards exploring the dynamic relationships between cognitive and sociopsychological IDs and general L2 proficiency. The study focused on the links between working memory (executive function, and phonological working memory), anxiety, attitude, and motivation of American learners of Spanish in the U.S. The results suggested that roles of IDs differed significantly depending on the timing of data collection (onset vs. endpoint) and learners' proficiency levels.

In discussing the results, Serafini (2017) stressed the importance of adopting an integrative perspective in researching IDs in order to accurately represent them as a set of dynamic and complex factors that affect L2 development. To our knowledge, however, no studies have taken such an approach towards investigating the differential impact of cognitive and socio-psychological IDs on L2 *pronunciation* learning (e.g., Baran-Łucarz, 2017 for motivation and anxiety; Baker-Smemoe & Haslam, 2013 for aptitude and motivation). Therefore, the primary focus of the current study was to understand the complex contributions of cognitive, motivational, and emotional IDs towards two different dimensions of L2 speech acquisition (enhancing comprehensibility vs. reducing foreign accentedness). To capture the dynamic nature of the ID-acquisition link, we designed a two-part study wherein we looked at the role of experience, aptitude, motivation, and emotion in L2 speech learning from both cross-sectional and longitudinal perspectives. In the cross-sectional phase (Study 1), the relationship between students' initial IDs and L2 pronunciation profiles was examined at the start of data collection. In the longitudinal phase (Study 2), the same participants' IDs was linked to their speech development during L2 pronunciation training. Following DST researchers' views on learner IDs (i.e., Serafini, 2017), and in keeping with the notion of *the trilogy of mind*, we focused on foreign language aptitude, motivation, and anxiety as proxies for the cognitive, motivational, and emotional aspects of L2 learners, respectively. Lastly, pronunciation was evaluated multidimensionally in terms of the degree of accentedness and comprehensibility. The research questions were formulated as follows:

1. Study 1: How are the comprehensibility and accentedness aspects of L2 speech differentially associated with speakers' experience and cognitive, motivational, and emotional ID factors at the onset of the project?
2. Study 2: How is L2 learners' speech development mediated by their cognitive, motivational, and emotional ID profiles when they receive explicit pronunciation instruction?

The following predictions were made based on previous ID research. Studies on L2 experience and pronunciation learning have demonstrated that accuracy in producing segmental and suprasegmental features develops according to the amount of recent and meaning-oriented interaction (Saito & Hanzawa, 2016). Specifically, it has been found that participants who have *recently* participated in extensive extracurricular L2 learning experiences (e.g., informal interactions with native and fluent non-native speakers in the target language) and classroom-based L2 speaking activities exhibit better comprehensibility and accentedness. In other words, it seems as though high quality speech can be achieved by means of exposure to rich linguistic input and receiving formal instruction (e.g., Derwing & Munro, 2013 for the evidence within naturalistic settings; Muñoz, 2014 for classroom settings).

When it comes to aptitude, research has shown that participants with greater phonemic coding ability and sound sequence recognition may demonstrate better accentedness (more nativelike) scores. This is arguably because they help learners attend to specific segmental and prosodic details in the input they receive (Saito et al., 2019). Therefore, we predicted that the same

pattern may be found in the current study. By contrast, the relationship between aptitude and comprehensibility has been shown to be weak at best. There is ample evidence that many L2 learners can continue to improve their comprehensibility (but not nativelikeness) as long as they are willing to use and practice the target language on a daily basis (Derwing & Munro, 2013). The linguistic features that contribute to comprehensibility are not necessarily limited to the accuracy of phonological features (e.g., Suzuki & Kormos, 2019), and thus may be unrelated to any aspects of phonological aptitude (e.g., phonemic coding ability).

With respect to the link between L2 learning motivation and pronunciation, previous studies have found that certain types of motivation may help learners notice detailed features of input under implicit learning conditions (e.g., Ushioda, 2016). In fact, there is evidence that learners who are more internally motivated (i.e., highly-developed Ideal L2 self) are able to make the most out of the available input and thus see greater improvements in comprehensibility (e.g., Saito et al., 2018). However, longitudinal studies of learners in naturalistic contexts have shown that reducing foreign accentedness requires years of experience using the target language (e.g., Munro & Derwing, 2008). Thus, a strong sense of Ideal L2 self may not be directly linked to higher degree of accentedness. When it comes to Ought-to L2 self (i.e., the perceived obligation for learning), evidence suggests that it may not significantly predict L2 pronunciation acquisition (Saito et al., 2018). The construct of Ought-to L2 self has multiple layers, and sense of obligation can be served as either facilitator or hinderance of L2 use. However, the current study follows the findings of the past study and predicts that learning a target language because of obligation may not necessarily lead to increased L2 use and L2 exposure.

Lastly, those who report a high degree of pronunciation learning anxiety may not be able to successfully refine their perception of L2 segmental and prosodic features (Piechurska-Kuciel, 2008). This is because anxiety can act as a further barrier to gaining opportunities to receive L2 input, and ultimately impede speech production and learning (e.g., Vasa & Pine, 2004). Hence, we predict that the learners with higher degrees of anxiety may show higher accentedness and lower comprehensibility scores.

As for the second objective of the current study (Study 2), we set out to explore the relationship between IDs and pronunciation learning in the context of explicit pronunciation instruction. Given that instruction is believed to *equally* facilitate adult L2 learners' pronunciation proficiency regardless of differences in the cognitive and sociopsychological profiles among L2 learners (Pennington, 2021), our prediction is that participants will be able to significantly enhance their comprehensibility and reduce their accentedness over time. Furthermore, the IDs variables that will be found to affect the participants' pronunciation proficiency at the onset may also influence the outcome of the instruction.

Study 1: Cross-sectional investigation

Participants

A total of 73 Japanese learners of English with varied learning experiences and backgrounds were recruited in Japan and included in the main analyses. Those learners reported that they had no prior experience in living or studying in English-speaking countries. None of them received any intensive pronunciation training in private English conversation schools or via private tutoring from English teachers at regular schools. They were first-year undergraduate students from various majors (e.g., engineering, medicine, sociology, education, literature, and cultural studies) and their average age was 19.41 years at the time of the project (*Range* = 18–20).

Procedure

After obtaining the necessary permissions from the universities in Japan, participants were recruited via posters and mailing lists. Interested students contacted one of the researchers, at which point the researcher scheduled individual appointments with each of the possible participants to determine candidacy. Upon completing a set of consent forms, the participants performed a spontaneous speech task, and took the LLAMA test on the researcher's laptop (approximately 30

minutes). Finally, they filled out a questionnaire sheet containing a set of questions about their language-learning background, L2 pronunciation learning motivation, and L2 pronunciation learning anxiety. The entire session lasted approximately 60 minutes.

Measures of individual differences

Aptitude test

In order to measure the participants' foreign language learning aptitude, the LLAMA test was used (Meara, 2005). The test was not only chosen for its popularity in SLA research (e.g., Bylund et al., 2010; Forsberg & Sandgren, 2013), but most importantly due to its first-language independent nature (in comparison to other available tests that are mainly for English native speakers). The sub-tests chosen for the current study included sound sequence recognition (LLAMA D) – for implicit learning aptitude (Granena, 2013; Suzuki, 2021 for the validation), associative memory (LLAMA B), and phonemic coding ability (LLAMA E) – for explicit learning aptitude. Except for LLAMA D whose maximum score is 75 %, maximum scores of LLAMA B and E are 100 %. The entire test session for measuring the aptitude took approximately 30 minutes. Descriptive statistics of participants' aptitude scores are illustrated in Supporting Information I.

Questionnaire instruments

After taking the aptitude test, the participants were asked to fill out a set of Likert-scale questionnaires that was designed to capture their L2 experience, L2 pronunciation-specific anxiety, and L2 pronunciation-specific motivation, respectively. Following previous ID studies (e.g., Kissling, 2014; Saito et al., 2018), we prepared a tailored questionnaire based on Language Contact Profile (Freed et al., 2004) to measure the participants' L2 experience. The items were designed to capture (a) the participants' *past* L2 learning experience before the university (i.e., at elementary, junior high, and high schools), and (b) the participants' *current* L2 learning experience at the university. In addition to the two distinctions (i.e., past and recent), the two types of L2 learning experience were further divided into either their time studying of English inside the regular curricular classes or their time using English for the conversations with other users of English (i.e., native and non-native speakers of English) outside the classroom (cf. Kissling, 2014 for a similar decision). Based on the participants' answers, total hours of L2 experience was calculated to create four types of experiential variables– past English learning inside the formal classrooms, past English use outside the formal classrooms, recent English learning inside the formal classrooms, recent English use outside the formal classrooms.

In terms of anxiety, the current study did not employ the oft-used Foreign Language Classroom Anxiety scale by Horwitz due to our emphasis on a skill-specific investigation – L2 pronunciation. Instead, the questionnaire developed by Baran-Łucarz (2016) was adopted in order to measure the participants' L2 pronunciation specific anxiety (see Supporting Information I for the items and descriptive statistics).

Finally, to measure the participants' pronunciation-specific motivation and anxiety, the current study utilized the questionnaire items of used in Baran-Łucarz (2017) which ask learners' degree of ideal L2 self, ought-to L2 self and anxiety in terms of L2 pronunciation learning (e.g., "I imagine myself as someone who is able to speak English *with accented but comprehensible pronunciation*."). The details of L2 pronunciation-specific anxiety and L2 pronunciation-specific motivation are summarized in Supporting Information I. In order to help the participants understand the questionnaire items, all the questions were translated into Japanese by the researcher and double checked by two translators. Since the Cronbach's alpha values of each construct indicated a relatively high level of internal consistency (α = .92 for ideal L2 self, α = .92 for ought-to L2 self, and α = .83 for anxiety), averaged score for each construct was computed. Finally, the interrelationship among the IDs and L2 experience variables were examined (see Supporting Information II).

Pronunciation proficiency measures

Speaking task

In order to tap into learners' less-controlled pronunciation knowledge, a semi-spontaneous speech task was adapted from EIKEN English Test (EIKEN, 2016; also see Lambert et al., 2017). Following the testing procedure established by EIKEN, the task sheet included four sequential pictures with several linguistic aids and a sentence to start their description. In order to prevent topic effect, two different pictures were used (Story A and Story B) (for the details of the task sheet, see Supporting Information III). A first half of the participants described Story A, and the remaining worked on Story B. The first 30 seconds of the approximately 2-minute speeches were taken from each of the 73 speech samples and saved as WAV files for the speech rating.

L2 pronunciation proficiency rating

Whereas some studies have examined L2 pronunciation proficiency via trained raters' assessments in accordance with detailed descriptors (e.g., Isaacs et al., 2015), much research attention has been given to untrained raters' *intuitive* judgements of L2 pronunciation proficiency. As seen in a range of existing studies (e.g., Derwing & Munro, 2013; Nagle, 2018a), we operationalized such intuitive judgements through scaler judgements of overall comprehensibility and accentedness.

Four raters (2 females, 2 males) with linguistic and pedagogical backgrounds were recruited in London. According to the research on listener factors, listeners' judgments are likely to be affected by factors such as their familiarity with the accent (e.g., Winke et al., 2013) and their language teaching experience (e.g., Kennedy & Trofimovich, 2008). Following the previous studies that employed subjective speech rating (e.g., Nagle, 2018a; Suzuki & Kormos, 2019), we carefully controlled the familiarity with Japanese-accented English. Based on a 6 point-scale (1 = *not at all*, 6 = *very much*), all four raters reported a high-level of familiarity with Japanese-accented English (*M* = 5.5; *Range* = 5–6). Thus, it was assumed that the leniency to the speech samples were relatively similar among the four raters and they are sufficiently sensitive to the speakers' use of Japanese sound system in the speech samples owing to their high familiarity to Japanese-accented English. All of them held master's degrees in applied linguistics and reported extensive experience in teaching English (*M* = 7.8 years) and participation in speech analyses of this kind. None of them reported any hearing problems.

Procedure of the pronunciation rating

The rating session was conducted via individual meetings with one of the researchers in a quiet room at a university in London, UK. The researcher helped the raters familiarize themselves with the rating procedure as well as the evaluation criteria. With a printed booklet, the raters were asked to listen to speech samples through headphones connected to a laptop computer, and subsequently evaluate the samples by circling a number on a 9-point scale for accentedness (*1 = heavily accented, 9 = not accented at all*) and comprehensibility (*1 = difficult to understand, 9 = easy to understand*) on a rating sheet. To ensure accurate and smooth rating, one of the researchers first provided a short training session to each of the raters prior to the main session. The training session included a brief explanation of the definitions of comprehensibility and accentedness, and a practice rating with three speech samples that were not included in the main dataset (see Supporting information IV for the training script). In order to ensure that the raters sufficiently understood the two constructs, the researcher asked the raters to explain their reasoning. Based on the explanations given, the researcher provided them with feedback. Subsequently, the raters proceeded to the main session. To avoid fatigue, the raters took 15 minutes breaks after one third, and two thirds of the speech samples were evaluated. The entire session lasted approximately 65 minutes per rater.

After all of the rating sessions were completed, the inter-rater reliability for the comprehensibility and accentedness results were calculated. The Cronbach's alpha of the four raters' judgments of comprehensibility was $\alpha = .82$ and accentedness was $\alpha = .80$. Since the Cronbach alpha analyses demonstrate acceptable agreements based on Larson-Hall's (2010) benchmark ($\alpha > .70$), the

results of the four raters' judgments were averaged to represent each speaker's comprehensibility and accentedness scores.[1]

Results

Constructing mixed-effects models

Study 1 was set to examine how experiential, cognitive and sociopsychological IDs differentially influence L2 pronunciation of 73 Japanese learners of English. For this purpose, the current study used mixed-effects modeling in R (R Core Team, 2018) with *Ime4* package, and built models that predict the learners' comprehensibility and accentedness scores. Prior to the model construction, the assumptions (linearity, homoscedasticity, normal distribution) were tested by the residual analyses. The fixed effects in the modelling included sound sequence recognition, phonemic coding ability, associative memory, ideal L2 self, ought to L2 self, and anxiety (those variables were collected at a single point in time). In order to control for L2 experience effect on the participants' comprehensibility and accentedness, past and recent L2 experience were also included as the fixed effects. These experience-related variables include the number of hours for regular English classes (inside-classroom experience); and the number of hours for the conversations with native and non-native speakers of English outside the regular English classes (outside-classroom experience). Furthermore, to ensure the comparability of the fixed effects that were measured through the different scaling systems, they were converted to z-scores prior to the analyses. For the evaluation of the models, we employed the pairwise Likelihood Ration Test (Baayen, 2008) to see whether the compared model decreases the Akaike's Information Criterion (AIC; an estimator of the relative amount of information lost by a particular model) with the forward selection method. The variables that did not improve the model fit via model comparisons were discarded. The variance inflation factors (VIFs) of all the predictors were below 2.0.

Predictors of L2 pronunciation proficiency at the onset of the project

Accentedness and IDs

According to series of model comparisons based on AIC values (for the details of constructed models, see Supporting information V), the final model suggested that phonemic coding ability (β = .24), anxiety (β = -.25), and recent English learning outside the classroom (β = .51) showed a significant contribution to determining accentedness score (Table 1). The predictive powers of these variables were further confirmed by the inspection of their confidence intervals at 95% level: All the values of the estimated regression coefficients were positive. The fixed effects in the final model explain a substantial amount of variance in the accentedness score (marginal R^2 = .44).

<Insert Table 1 about here>

Comprehensibility and IDs

The model comparisons revealed that the model with the lowest AIC value includes recent English learning outside the classroom (β = .30) and recent English learning inside the classroom (β =.28) as the statistically significant predictors of higher comprehensibility (AIC = 210.25; for the model comparisons, see Supporting Information V). Furthermore, the inspections of the confidence intervals at 95% level confirmed the positive contributions of these variables to comprehensibility. Therefore, among ten variables, the fixed effects in the final model accounted for 20 % of the total variance (marginal R^2 = .20).

<Insert Table 2 about here>

Study 2: Longitudinal investigation

The findings of Study 1 revealed that the ID profiles of Japanese EFL students (with years of foreign language education) were differentially related to comprehensibility and accentedness scores. The participants demonstrated higher comprehensibility as long as they regularly practiced the target language both inside and outside of the classroom. However, those with more nativelike

pronunciation tended to access L2 English beyond the classroom setting, demonstrated greater phonetic aptitude, and had less anxiety. One obvious limitation in Study 1 is that the data was collected at a single time point. Since the ID-proficiency link is dynamic and everchanging in nature, Study 2 was designed to replicate the findings of Study 1 (i.e., more IDs effects for accentedness than comprehensibility) from a longitudinal approach. The goal of Study 2 was to assess the mediating roles of aptitude, anxiety and motivation in the development of L2 comprehensibility and accentedness, when participants received explicit pronunciation instruction for four weeks (50 minutes × 4 weeks). Since the existing research on L2 pronunciation instruction has demonstrated the effectiveness of explicit instruction on L2 segmental and suprasegmental proficiency (Saito & Plonsky, 2019 for a review), it was assumed that the treatment (i.e., pronunciation instruction) in the current study would positively impact the comprehensibility and accentedness of participants' L2 speech.

Participants

Out of 73 participants who took the tests at the onset of the project, 63 agreed to participate in Study 2. In order to ensure that pronunciation instruction help L2 learners make tangible improvement in accentedness and comprehensibility, participants were assigned to the experimental group who receive pronunciation instruction (n = 51), and to the control group who received grammar instruction (n = 12). The latter group did not receive any pronunciation instruction. The number of participants in the experimental group was considerably larger because the main objective of Study 2 lay in the role of IDs in L2 pronunciation learning gains. The purpose of the control group was to demonstrate test-retest effects given that similar materials were used for pre- and post-tests. Both experimental and control groups received 50-minute-long instruction every week for 4 weeks. The procedure was summarized in Figure 1.

<Insert Figure 1 about here>

Treatment: Experimental group

Explicit pronunciation instruction was provided to the participants in the experimental group. L2 pronunciation instructions used in past research can be broadly categorized into articulatory-based and auditory-based instructions with the former highlighting L2 learners' understanding of the manner and place of articulation of sounds in contrast to their L1, and the latter emphasizing L2 learners' perceptual development of sounds by introducing similarities and dissimilarities of L2 and their L1 counterparts (Saito & Plonsky, 2019). Since perception and production are assumed to complement each other to facilitate L2 speech learning (Nagle, 2018b), the training materials in the current study comprised both perception- and production-based practice activities (see Couper, 2003 for a similar approach; for detailed description of intervention, see Supporting Information VII).The sessions were led by a researcher who is a native speaker of Japanese with a master's degree in TESOL and highly proficient in English. The study used non-native teachers who have been shown to be capable of providing effective pronunciation instruction (Levis et al., 2016), and teachers/listeners of the same L1 are better equipped at noticing pronunciation errors that are derived from the L1 phonological system (e.g., Riney et al., 2000).

Treatment: Control group

The control group received grammar instruction with exercises (e.g., filling in the blanks, passage comprehension, error recognition) chosen from the textbook for The Test of English for International Communication (TOEIC) (Trew, 2007).

Pronunciation proficiency measures

The same picture description tasks in Study 1 were used for post-tests. To ensure that participants did not work on the same prompts, however, two different versions of pictures were counterbalanced for each participant (Story A → Story B; Story B → Story A). Following the same procedure in Study 1, the same expert raters (four linguistically trained native speakers) listened to all the speech samples in a randomized order (122 samples), and made intuitive judgements for comprehensibility and accentedness. Given that the raters demonstrated an adequate level of

agreement ($\alpha > .80$), their rating scores were averaged to derive one single comprehensibility and accentedness score for each participant at pre- and post-tests, respectively.

Results

Constructing mixed-effects models

Study 2 was set to investigate the potential moderating effect of learner IDs and experiential variables on the effectiveness of pronunciation instruction. Therefore, after the inspection of the residuals of the variables for meeting the statistical assumptions for mixed effects model, the interactions between instruction (i.e., *Time*), and learner IDs were examined by following procedure. First, interaction terms were prepared by combining instruction (*Time*) and one fixed effect (e.g., *Sound sequence recognition*). After preparing the interaction terms for all the fixed effects, the codes were run individually.

Effectiveness of pronunciation instruction

In order to make sure that the groups did not differ in terms of their ID profiles, L2 experience and L2 pronunciation, a series of statistical analyses were conducted. First, prior to the t-tests, Levene's test was conducted to test the hypothesis of equal population variances. Since all the variables did not show any statistical significance, the null hypothesis of equal population variances was not rejected. Therefore, a series of t-tests were conducted to examine the possible differences between the two groups. Due to the uneven number of participants in each group (51 vs. 12), Welch's t-test was used. According to the results, the experimental and control groups were not statistically different in terms of the pre-test scores of comprehensibility and accentedness as well as the ID profiles. After the intervention, the post-test scores of the two groups were compared using paired-samples t-test. The results indicated that only experimental group showed statistically significant improvements in comprehensibility and accentedness ($t = 6.468$, $p > .001$ for comprehensibility; $t = 8.436$ $p > .001$ for accentedness). Concerning the effect size of the treatment, Cohen's d was calculated (Cohen's $d = 0.7$ for Comprehensibility, and Cohen's $d = 1.3$ for Nativelikeness). According to Plonsky & Oswald's (2014) field-specific benchmark of the effect size, these results can be considered as medium to large effect size. Therefore, the results suggest that pronunciation instruction was equally facilitative of L2 comprehensibility and accentedness.

The roles of aptitude, motivation and anxiety in the effectiveness of instruction

According to the result of mixed effects modelling, the estimated beta values of the ID variables elicited from the experimental group (who received the pronunciation instruction) did not show statistically significant interaction effect (i.e., $p > .220$). The estimated beta values, standard errors, t-values of the model that includes the interactions are summarized in Table 3 for accentedness and Table 4 for comprehensibility. The results suggest that (a) the unique contribution of IDs to comprehensibility and accentedness over time; and (b) that explicit instruction can help learners enhance the comprehensibility and nativelikeness aspects of L2 pronunciation proficiency regardless of IDs profiles.

<Insert Table 3 about here>

<Insert Table 4 about here>

Discussion

Focusing on the EFL context, the current study sought to examine the complex and dynamic mechanisms underlying adult L2 speech learning. To this end, we conducted cross-sectional and longitudinal investigations of how Japanese EFL students with different experiential, cognitive and sociopsychological IDs attained two different constructs of L2 pronunciation proficiency (comprehensibility and accentedness) after years of EFL education, and following pronunciation instruction. Two overall conclusions were derived. First, we argue that L2 speech learning is a highly complex phenomenon that needs to be scrutinized not only along learner dimensions (experiential, cognitive, and sociopsychological IDs), but also along linguistic dimensions (comprehensibility vs. accentedness). Second, we argue that provision of instruction can be equally effective regardless of differences in L2 learners' cognitive and sociopsychological profiles.

R1: Roles of IDs in L2 pronunciation learning

Overall, the results confirm that experiential factors, and different aspects of IDs, play important roles in determining how, and to what degree, learners can develop their L2 speech. A positive relationship was found between time spent in the regular English classes at the university and comprehensibility. In light of evidence from previous L2 pronunciation studies that accentedness is mainly linked to segmental and suprasegmental accuracy (i.e., phonological accuracy), and that comprehensibility is associated with wider range of linguistic features such as temporal, lexical, grammatical, and phonological accuracy (e.g., Trofimovich & Isaacs, 2012), it seems as though the participants' regular English classes may have helped them improve the temporal and lexicogrammatical aspects of their speech. It is noteworthy, however, that recent L2 use outside the regular English classes at the university (i.e., using L2 for communication with native and non-native speakers of English) was strongly associated with both comprehensibility and accentedness. Echoing findings from previous studies which have examined the influence of L2 experience (e.g., Baker-Smemoe & Haslam, 2013; Saito & Hanzawa, 2016), this confirms the importance of *extensive* exposure to, and use of, the target language in pronunciation learning (e.g., Flege, 2016). Since this variable was associated with both accentedness and comprehensibility, it can be concluded that input and output beyond one's regular L2 experience can help further strengthen and refine one's accumulated knowledge of pronunciation *and* lexicogrammar. The positive links between the two types of L2 experience (classroom English learning experience vs. extracurricular conversations with native/non-native speakers) and the two dimensions of L2 pronunciation offers additional evidence for the experience-driven account of successful L2 speech learning (e.g., Muñoz, 2014). This account holds that, in the EFL classroom setting, English learning experience can lead to improvements in comprehensibility via improvements in the accuracy of various pronunciation features. However, learners who make extra efforts to increase the amount of L2 use/exposure outside the classrooms (e.g., communications with international friends) may be able to reduce their degree of L1 phonological transfer and consequently reduce their accentedness.

Asymmetric patterns were found regarding the influence of cognitive and psychosocial factors: phonemic coding and lower anxiety were associated with L2 accentedness, but no factors were related to L2 comprehensibility. This could partially be explained by the differences in the constructs of accentedness and comprehensibility. Specifically, L2 pronunciation studies have revealed that accentedness is mainly linked to segmental and suprasegmental accuracy (i.e., phonological accuracy) whereas comprehensibility is associated with wider range of linguistic features such as temporal, lexical, grammatical, and phonological accuracy (e.g., Trofimovich & Isaacs, 2012). Based on these results, it can be concluded that L2 learners who have higher phonemic coding ability and lower anxiety may have been able to successfully reduce the use of their L1 sound system (i.e., Japanese) in L2 speech, resulting in improved segmental and suprasegmental accuracy. Because both phonemic coding and anxiety are believed to be involved in information processing (e.g., Baran-Łucarz, 2013; Skehan, 2016), it can also be concluded that higher phonemic coding ability and/or lower anxiety could help learners notice cross-linguistics differences, retain analyzed auditory information, and integrate it into their L2 systems.

In line with past research on explicit learning aptitude and L2 pronunciation (e.g., Baker-Smemoe & Haslam, 2013; Saito, et al., 2019 for a cross-sectional evidence), the results of the current study support the idea that phonemic coding ability helps learners improve the segmental and suprasegmental aspects of their speech (i.e., accentedness). However, unlike other cross-sectional studies which have found an association between associative memory, superior grammatical complexity, and speed fluency (e.g., Saito et al., 2019), higher associative memory was not found to be a predictor of comprehensibility or accentedness in the current study.

These results could be explained, on the one hand, by the notion that the participants' use of grammar and/or temporal features may not have been fully reflected in the raters' judgements. However, an alternative explanation can be provided as well. Previous aptitude research has shown

that associative memory can help learners retain a vast amount of lexical knowledge, relate new information to existing knowledge, and control the delivery of such knowledge efficiently so that it mainly involves in the *later stages* of L2 acquisition – i.e., the proceduralization and automatization of acquired knowledge (e.g., Skehan, 2016). Based on this, it is reasonable to assume that the participants in the current study may have yet reached the later stages of acquisition, and/or may not have had sufficient declarative knowledge to benefit from their superior associative memory. The same account could also explain the insignificant relationship found between sound sequence recognition and L2 pronunciation. Sound sequence recognition is believed to help L2 learners attend to L2 phonological and word sequences in an incidental and implicit fashion. It is thus considered to be essential in the later stages of L2 acquisition, i.e., for the further refinement of L2 sound processing ability and the attainment of nativelike L2 pronunciation (e.g., Granena, 2013 for naturalistic setting; Saito et al., 2019 for FL setting). Thus, participants with higher sound sequence recognition may have been in the earlier stages of L2 pronunciation acquisition, where the explicit processing and analysis of L2 sounds are more instrumental to success.

Next, a negative relationship was found between anxiety and reduced accentedness. Such a result concurs with previous studies showing that anxiety can affect L2 pronunciation acquisition (e.g., Saito et al., 2018; Szyszka, 2011). In the case of the current study, however, the participants' comprehensibility was not associated with their level of pronunciation specific anxiety. These contrasting results may suggest that, irrespective of anxiety, participants may be able to attend to phonological features with a degree of sufficient of accuracy in a way that makes their speech comprehensible. However, because anxiety is known to interfere with attention control (e.g., Piechurska-Kuciel, 2008), high-anxiety participants may have not been able to allocate sufficient attention to the differentiation of L1 and L2 sounds when speaking.

With respect to motivation, neither Ideal or Ought-to L2 self were linked to comprehensibility or accentedness. This provides counter evidence to past studies which have found a strong association between Ideal L2 self and comprehensibility (e.g., Saito et al., 2018). At the same time, a small but positive link was found between Ideal L2 self and recent L2 learning outside of the classroom (r = .226, p = .054, see Supporting Information III). Although this link did not reach the threshold of statistical significance, it may nevertheless suggest that participants with internalized motivation may have actively sought out opportunities to practice English *outside* of the classroom (e.g., Saito et al., 2018; Ushioda, 2016). Unlike past studies, which have used questionnaires for general English learning in general, the current study tailored the statements to elicit responses *specific to* pronunciation (e.g., Baran-Łucarz, 2017). Thus, further research is needed in order to confirm the relationship between pronunciation specific motivation and L2 pronunciation acquisition.

R2: Roles of instruction in learner individual differences

The second aim of the study (Study 2) was to examine the extent to which the relationship between IDs and proficiency varied over time following explicit pronunciation instruction. The results showed that there were no significant interactions between any ID variables and instructional gains. This runs counter to prior evidence showing that aptitude moderates the effectiveness of, for example, L2 grammar instruction (e.g., Yalçin & Spada, 2016). The findings rather suggest that instruction is facilitative of L2 pronunciation development *regardless of* learners' ID variables. Different from L2 grammar instruction, wherein learners need to process abstract and complex concepts of language, L2 pronunciation learning mainly comprises a perceptual-motor phenomenon. In this regard, the results support the view that the explicit explanation of L2 pronunciation features (i.e., articulatory-based and auditory-based instruction) may be *equally* beneficial for learners with various aptitude, motivation, and anxiety profiles (e.g., Couper, 2003)

Conclusion

The current study addressed the complex relationships between learner IDs and L2 pronunciation learning in the EFL classroom setting. Grounded in the view that pronunciation

proficiency is a multi-dimensional construct with interrelated components (Saito & Plonsky, 2019), we employed two holistic measurements of L2 proficiency (i.e., comprehensibility and accentedness) to illustrate their interconnectivity and interaction with an array of learner IDs. The results speak to the complex role of IDs in shaping the course of L2 pronunciation acquisition. First and foremost, the findings suggest that the extensive use of a target language greatly promotes the development of L2 comprehensibility and accentedness. In the context of the current study (i.e., English-as-a-Foreign-Language), such experience-related factors include the amount of language-focused practice *inside* classrooms and conversational interactions with users of English *outside* classrooms. When it comes to linguistic nativelikeness (accentedness), however, further improvement can be observed only among certain individuals with greater phonemic coding ability and lower levels of anxiety towards L2 pronunciation learning. The absence of any links between IDs and instructional gains suggests that pronunciation-focused instruction is effective for L2 learners regardless of their ID profiles.

As for theoretical contribution, the current study is the first attempt to extend the integrative framework of SLA to L2 pronunciation in EFL classroom contexts. Echoing the fundamental tenant of DST and the trilogy of mind, the study provides a comprehensive picture of the complex relationship between use, learner individual differences, and language development. On a broad level, our findings indicate that whereas both socio-psychological individual differences are tied to use (e.g., greater motivation leads to more practice inside and outside classrooms), cognitive aptitude servers as a factor of advanced L2 acquisition. In the field of L2 pronunciation, however, we add that one domain-specific crucial source of individual variation concerns the dimensions of proficiency, i.e., comprehensibility vs. accentedness. As for comprehensibility, which many scholars consider as an index of a functional user of L2 English (e.g., Derwing & Munro, 2013), there is a great possibility that more L2 practice leads learners to be comprehensible. As for accentedness, which has been claimed to represent an *ideal* (but not necessarily *realistic*) goal of L2 speech learning, foreign accent reduction can be an extremely difficult task especially among post-pubertal learners, and limited to certain individuals with high-level cognitive aptitude (Linck et al., 2013).

To close, several limitations of the study need to be acknowledged. First of all, the participants' L2 experience profiles were surveyed using a questionnaire (i.e., Language Contact Profile). Although the use of self-report data is common in SLA (cf. Derwing & Munro, 2013), it may not accurately reflect participants' actual language exposure. Therefore, the findings related to L2 experience in this study need to be treated as tentative. More accurate measurements of the quantity and quality of L2 experience should be obtained in future studies by, for example, asking participants to track their L2 interactions using their mobile phones (Surtees, 2013) or using electronic language logs (Ranta & Meckelborg, 2013).

Secondly, we would like to emphasize that the findings in the current study need to be replicated and verified. Following previous L2 pronunciation studies (e.g., Saito et al., 2019), we used the LLAMA test to gauge participants' foreign language learning aptitude. However, because several scholars have recently cast doubt on the reliability of this battery (e.g., Bokander & Bylund, 2020), the results need to be treated with some caution. In addition, as we illustrated in the literature review, there are a wealth of influential aptitude tests such as the CANAL-F test (Grigorenko et al., 2000) and Hi-LAB (Doughty et al., 2010) that can be employed as research tools. In order to confirm the relationship between L2 pronunciation and aptitude, it is thus important to replicate the study with more reliable aptitude measures.

Thirdly, the current study used single speaking task (i.e., picture description task) to evaluate participants' L2 pronunciation performance. However, it has been recognized that speaking style and the type of L2 knowledge used in L2 speech (i.e., controlled vs. spontaneous knowledge) varies depending on the nature of tasks and condition of its administration (e.g., controlled vs. semi-structure vs. fully free tasks). Because of this, it is crucial for future studies to assess speakers'

performance using multiple speaking tasks (see Saito & Plonsky, 2019 for more discussion of task type in relation to L2 declarative knowledge).

Fourthly, we acknowledge that pronunciation skills, the main focus of this study, comprise only one aspect of general L2 proficiency. More studies are needed to assess whether, to what degree, and how aptitude, motivation, and emotion mediate L2 pronunciation improvement, and how this ultimately impacts the process and product of *general* L2 learning. For example, to unpack the relationships between IDs and general proficiency, it would be interesting to examine the generalizability of our findings to reading, listening, writing, grammar, and vocabulary learning. Future studies should also develop, validate, and refine theoretically sound methods to tap into the highly complex nature of L2 general proficiency.

Lastly, the purpose of the current study was to capture the complex relationship between different IDs in relation to L2 pronunciation proficiency. However, we acknowledge that we only covered a small number of key IDs. In order to fully apply the principle of DST and provide a fuller picture of the relationship between IDs and L2 pronunciation development, future studies should include as many factors as possible, including working memory (e.g., Hu et al., 2012), musical aptitude (Li & Dekeyser, 2017), and personality (e.g., Hu & Reiterer, 2009).

Future direction

In this current project, we aimed to track the relationship between IDs and L2 learning over time (e.g., Serafini, 2017) via both cross-sectional and longitudinal investigations. Although participants' ID profiles were examined only once at the beginning of the project, we would like to emphasize that ID factors (especially related to sociopsychological dimensions of L2 learners) can be considered as a dynamic (rather than stable) phenomenon. There is ample evidence demonstrating the fluctuations among L2 learners' motivation (e.g., Pawlak, 2012; Waninge et al., 2014) and state anxiety (Gregersen, 2020). There has been an ongoing debate on the malleability of language aptitude among aptitude researchers (e.g., Kormos, 2013; Singleton, 2017; Wen, Skehan, & Biedroń, 2017). To this end, we call for future research which will examine the ever-changing nature of various L2 learners' IDs and its impact on L2 speech learning at different time points over an extensive period of L2 immersion and classroom instruction.

Notes

1. The two global constructs – accentedness and comprehensibility – showed a positive medium-to-strong correlation ($r = .65$, $p < .001$). As shown and discussed in many previous studies (e.g., Derwing & Munro, 2015), the two constructs could be considered as a somewhat overlapping but essentially different phenomenon, suggesting that some L2 speech can be strongly accented but highly comprehensible.

References

Baayen, RH (2008) *Analyzing Linguistic Data. A Practical Introduction to Statistics Using R*. Cambridge: Cambridge University Press.

Baker-Smemoe W and 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 (2016) Conceptualizing and Measuring the Construct of Pronunciation Anxiety: Results of a Pilot Study. In Pawlak, M (ed), *Classroom-Oriented Research.* Cham: Springer, pp. 39–56.

Baran-Łucarz M (2017) FL Pronunciation Anxiety and Motivation: Results of a Mixed-Method Study. In *At the crossroads: Challenges of foreign language*. Cham: Springer, pp. 107–133.

Bokander, L and Bylund, E (2020) Probing the internal validity of the LLAMA language aptitude tests. *Language Learning* 70, 11–47.

Bylund E, Abrahamsson, N and Hyltenstam, K (2010) The role of language aptitude in first language attrition: The case of prepubescent attriters. *Applied Linguistics* 31, 443–464.

Carroll, JB (1962) The prediction of success in intensive foreign language training. In R. Glaser (ed), *Training research and education*. Pittsburgh: University of Pittsburgh Press, pp. 87–136.

Carroll, JB and Sapon, S(1959) *Modern language aptitude test*. New York: Psychological Corporation.

Couper, G (2003) The value of an explicit pronunciation syllabus in ESOL teaching. *Prospect* 18, 53–70.

Darcy, I, Park, H, and Yang, CL (2015) Individual differences in L2 acquisition of English phonology: The relation between cognitive abilities and phonological processing. *Learning and Individual Differences* 40, 63–72.

de Bot, K (2008) Introduction: Second language development as a dynamic process. *The Modern Language Journal* 92, 166–178.

Derwing, TM and Munro, MJ (2013) The development of L2 oral language skills in two L1 groups: A 7-year study. *Language Learning* 63, 163–185.

Dewaele, JM and 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.

Doughty, CJ, Campbell, SG, Mislevy, MA, Bunting, MF, Bowles, AR and Koeth, JT (2010) Predicting near-native ability: The factor structure and reliability of Hi-LAB. In *Selected proceedings of the 2008 Second Language Research Forum*. Somerville, MA: Cascadilla Proceedings Project, pp. 10-31.

Dörnyei, Z (2005) *The psychology of the language learner: Individual differences in second language acquisition*. Mahwah, NJ: Lawrence Erlbaum Associates.

Dörnyei, Z and Chan, L (2013) Motivation and vision: An analysis of future L2 self images, sensory styles, and imagery capacity across two target langauges. *Language Learning* 63, 437–462.

Erlam, R (2005) Language aptitude and its relationship to instructional effectiveness in second language acquisition. *Language Teaching Research* 9, 147–171.

EIKEN Foundation of Japan (2016) *EIKEN Pre-1 level: Complete questions collection.* Tokyo: Oubunsha.

Flege, J (2016, June) The role of phonetic category formation in second language speech acquisition. Plenary address delivered at New Sounds, Aarhus, Denmark.

Forsberg, LF and Sandgren, M (2013) High-level proficiency in late L2acquisition–Relationships between collocational production, language aptitude and personality. In Grañena, G and Long M (eds), *Sensitive periods, language aptitude, and ultimate L2 attainment.* Amsterdam: John Benjamins, pp. 231–256.

Freed, BF, Dewey, DP, Segalowitz, NS and Halter, RH (2004) The language contact profile. *Studies in Second Language Acquisition* 26, 349–356.

Gatbonton, E and Trofimovich, P (2008) The ethnic group affiliation and L2 proficiency link: Empirical evidence. *Language Awareness* 17, 229–248.

Gardner, RC (2007) Motivation and second language acquisition. *Porta Linguarum* 8, 9–20.

Gonet, W (2006) Success in the acquisition of English phonetics by Poles. A pilot study. In Sobkowiak,W and Waniek-Klimczak, E (eds) *Phonetics in foreign language teaching 6*. Konin: Wydawnictwo PWSZ, pp. 70–88.

Granena, G (2013) Cognitive aptitudes for second language learning and the LLAMA Language Aptitude Test. In Granena G and Long M (eds) *Sensitive periods, language aptitude, and ultimate L2 attainment.* Amsterdam: John Benjamins Publishing, pp.105–129.

Grigorenko, E, Sternberg, R and Ehrman, M (2000) A theory-based approach to the measurement of foreign language learning ability: The Canal-F theory and test. *The Modern Language Journal 84*, 390–405.

Hiver, P and Al-Hoorie, AH (2016) A dynamic ensemble for second language research: Putting complexity theory into practice. *The Modern Language Journal* 100, 741–756.

Horwitz, EK, Horwitz, M and Cope, J (1986) Foreign language classroom anxiety. *The Modern Language Journal* 70, 125–132.

Hu, X, Ackermann, H, Martin, JA, Erb, M, Winkler, S and Reiterer, SM (2013) Language aptitude for pronunciation in advanced second language (L2) Learners: Behavioural predictors and neural substrates. *Brain and Language* 127, 366–376.

Hu, X and Reiterer, S (2009) Personality and pronunciation talent in second language acquisition. In Dogil, GS and Reiterer, M (eds) *Language talent and brain activity*. Berlin: Mouton de Gruyter, pp. 97–130.

Huensch, A., & Thompson, A. S. (2017). Contextualizing attitudes toward pronunciation: Foreign language learners in the United States. *Foreign Language Annals 50*, 410–432.

Isaacs, T, Trofimovich, P, Yu, G, and Chereau, BM (2015) Examining the linguistic aspects of speech that most efficiently discriminate between upper levels of the revised IELTS pronunciation scale. *IELTS research reports online series* 4.

Kennedy, S and Trofimovich, P (2008) Intelligibility, comprehensibility, and accentedness of L2 speech: The role of listener experience and semantic context. *Canadian Modern Language Review* 64, 459–489.

Kennedy, S and Trofimovich, P (2010) Language awareness and second language pronunciation: A classroom study. *Language Awareness* 19, 171–185.

Kissling, EM (2014) What predicts the effectiveness of foreign-language pronunciation instruction? Investigating the role of perception and other individual differences. *Canadian Modern Language Review 70*, 532–558.

Kormos, J. (2013). New conceptualizations of language aptitude in second language attainment. In G. Granena& M. Long(Eds.), Sensitive periods, language aptitude and ultimate language attainment (pp. 113 – 152). Amsterdam : Benjamins.

Lambert, C, Kormos J and Minn, D (2017) Task repetition and second language speech processing. *Studies in Second Language Acquisition* 39, 167–196.

Larsen–Freeman, D (2013) Complexity theory: A new way to think. *Revista Brasileira de Linguistica Aplicada* 13, 369–373.

Larson-Hall, J (2010) *A guide to doing statistics in second language research using SPSS*. New York: Routledge.

Levis, JM, Sonsaat, S, Link, S and Barriuso, TA (2016) Native and nonnative teachers of L2 pronunciation: Effects on learner performance. *TESOL Quarterly* 50, 894–931.

Li, S (2016) The construct validity of language aptitude: A meta-analysis. *Studies in Second Language Acquisition* 38, 801–842.

Li, M and DeKeyser, R (2017) Perception practice, production practice, and musical ability in L2 Mandarin tone-word learning. *Studies in Second Language Learning 39*, 593–620.

Linck, JA, Hughes, MM, Campbell, SG, Silbert, NH, Tare, M, Jackson, SR, Smith, BK, Bunting, MB and Doughty, CJ (2013) Hi-LAB: A new measure of aptitude for high-level language proficiency. *Language Learning* 63, 530–566.

Liu, M and Huang W (2011) An exploration of foreign language anxiety and English learning motivation. *Education Research International*, 1–8

Matthews, G and Zeidner, M (2004) Traits, states, and the trilogy of mind: An adaptive perspective on intellectual functioning. In Dai, DY and Sternberg, RJ (eds) *Motivation, emotion, and cognition: Integrative perspectives on intellectual functioning and development.* Mahwah, NJ: Erlbaum, pp. 143–174.

Meara, PM (2005) *LLAMA language aptitude tests: The manual*. Swansea: Lognostics.

Moyer, A (2014) Exceptional outcomes in L2 phonology: The critical factors of learner engagement and self-regulation. *Applied Linguistics* 35, 418–440.

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 and Singleton D (2007) Foreign accent in advanced learners: Two successful profiles, *EUROSLA Yearbook* 7, 171–90.

Munro, MJ and Derwing, TM (2008) Segmental acquisition in adult ESL learners: A longitudinal study of vowel production. *Language Learning* 58, 479–502.

Nagle, C (2018a) Motivation, Comprehensibility, and Accentedness in L2 Spanish: Investigating Motivation as a Time-Varying Predictor of Pronunciation Development. *The Modern Language Journal* 102, 199–217.

Nagle, C (2018b) Perception, production, and perception-production: Research findings and implications for language pedagogy. *Contact Magazine* 44, 5–12.

Ortega, L (2013) *Understanding second language acquisition*. New York: Routledge.

Pawlak, M. (2012). The dynamic nature of motivation in language learning: A classroom perspective. *Studies in Second Language Learning and Teaching 2*, 249–278.

Pennington, MC (2021) Teaching Pronunciation: The State of the Art 2021. *RELC Journal* 52, 3–21.

Plonsky, L., & Oswald, F. L. (2014). How big is "big"? Interpreting effect sizes in L2 research. *Language Learning 64*, 878–912.

Piechurska-Kuciel, E (2008) *Language Anxiety in Secondary Grammar School Students*. Opole: Wydawnictwo Uniwersytetu Opolskiego.

R Core Team (2018) *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/. R

Ranta, L and Meckelborg A (2013) How much exposure to English do international graduate students really get? Measuring language use in a naturalistic setting. *Canadian Modern Language Review* 69, 1–33.

Riney, TJ, Takada, M and Ota, M (2000) Segmentals and global foreign accent: The Japanese flap in EFL. *TESOL Quarterly* 34, 711–737.

Saito, K, Dewaele, JM, Abe, M and In'nami, Y (2018) Motivation, Emotion, Learning Experience, and Second Language Comprehensibility Development in Classroom Settings: A Cross-Sectional and Longitudinal Study. *Language Learning* 68, 709–743.

Saito, K and 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 and Plonsky, L (2019) Effects of second language pronunciation teaching revisited: A proposed measurement framework and meta-analysis. *Language Learning* 69, 652–708.

Saito, K, Sun, H and Tierney, A (2020) Domain-general auditory processing determines success in second language pronunciation learning in adulthood: A longitudinal study. *Applied Psycholinguistics* 45, 1083–1112.

Saito, K, Suzukida, Y and Sun, H (2019) Aptitude, Experience, and Second Language Pronunciation Proficiency Development in Classroom Settings: A Longitudinal Study. *Studies in Second Language Acquisition* 41, 201–225.

Sardegna, VG, Lee, J and Kusey, C (2014) Development and validation of the learner attitudes and motivations for pronunciation (LAMP) inventory. *System* 47, 162–175.

Serafini, EJ (2017) Exploring the Dynamic Long-Term Interaction Between Cognitive and Psychosocial Resources in Adult Second Language Development at Varying Proficiency. *The Modern Language Journal* 101, 369–390.

Singleton, D. (2017). Language aptitude: Desirable trait or acquirable attribute?. *Studies in Second Language Learning and Teaching* 7, 89–103.

Skehan, P. (2016). Foreign language aptitude, acquisitional sequences, and psycholinguistic processes. In Granena G, Jackson D and Yilmaz Y (eds), *Cognitive individual differences in L2 processing and acquisition*. Amsterdam: John Benjamins, pp. 15–38.

Surtees, V (2013) *An alternative approach to speech act research in the study abroad context* (Doctoral dissertation, Concordia University).

Suzuki, Y (2021) Probing the construct validity of LLAMA_D as a measure of implicit learning aptitude: incidental instructions, confidence ratings, and reaction time. *Studies in Second Language Acquisition*, First View.1–14.

Suzuki, Y and DeKeyser, R (2015) Comparing elicited imitation and word monitoring as measures of implicit knowledge. *Language Learning 65*, 860–895.

Szyszka, M (2011) Foreign language anxiety and self-perceived pronunciation competence. *Studies in Second Language Learning and Teaching 1*, 283–300.

Suzuki, S and Kormos, J (2019) The Effects of Readaloud Assistance on Second Language Oral Fluency in Text Summary Speech. In Rose, RL and Eklund, R (eds) Proceedings of DiSS 2019, The 9th Workshop on Disfluency in Spontaneous Speech, pp.31–34.

Teimouri, Y, Goetze, J and Plonsky, L (2019) Second language anxiety and achievement: A meta-analysis. *Studies in Second Language Acquisition* 41, 363–387.

Trew, G (2007) *Tactics for TOEIC*. New York: Oxford.

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

Trofimovich, P, Kennedy, S and Foote, JA (2015) Variables Affecting L2 Pronunciation Development. In Reed M and Levis JM (eds) *The handbook of English pronunciation.* West Essex, UK: Wiley & Sons, pp. 353–373.

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

Vasa, RA and Pine DS (2004) Neurobiology in anxiety disorders in children and adolescents. In Morris, TR and March, JS (eds) *Anxiety disorders in children and adolescents* New York: Guilford Press, pp. 3–26.

Waninge, F. (2015). Motivation, emotion, and cognition: Attractor states in the classroom. In Dörnyei, Z, MacIntyre, PD and Henry, A (eds) *Motivational dynamics in language learning* New York: Multilingual Matters, pp. 195–213.

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

Wen, Z. E., Biedroń, A., & Skehan, P. (2017). Foreign language aptitude theory: Yesterday, today and tomorrow. *Language Teaching 50*, 1–31.

Winke, P, Gass, S and Myford, C (2013) Raters' L2 background as a potential source of bias in rating oral performance. *Language Testing* 30, 231–252.

Yalçın, Ş and Spada, N (2016) Language aptitude and grammatical difficulty: An EFL classroom-based study. *Studies in Second Language Acquisition* 38, 239–263.

Table 1

Summary of the Final Model of Reduced Accentedness

Predictors	Estimate	*SE*	*t*-value	*p*	95% *CI*	
					Lower	Upper
(Intercept)	**5.01**	**.09**	**57.48**	**<.001***	**4.76**	**5.25**
Phonemic coding ability	**.24**	**.09**	**2.71**	**.008***	**.06**	**.42**
Anxiety	**-.25**	**.09**	**-2.83**	**.006***	**-.43**	**-.08**
Recent English learning outside the classroom	**.51**	**.09**	**5.7**	**<.001***	**.33**	**.69**
Random effect (intercepts)	Variance	*SD*				
Task	<.001	<.001				
Information criterion	Estimate					
LogLikelihood	-82.118					
DIC	164.24					
AIC	176.24					
BIC	189.98					
R^2	Estimate					
Marginal	.44					
Conditional	.45					

Note. DIC = Deviance Information Criterion; AIC = Akaike Information Criterion; BIC = Bayesian Information Criterion

Table 2

Summary of the Final Model of Improved Comprehensibility

Predictors	Estimate	*SE*	*t*-value	*p*	95% *CI*	
					Lower	Upper
(Intercept)	**5.44**	**.11**	**48.75**	**<.001***	**5.18**	**5.69**
Recent English learning inside the classroom	**.28**	**.12**	**2.357**	**.02***	**.04**	**.51**
Recent English learning outside the classroom	**.3**	**.12**	**2.57**	**.01***	**.07**	**.54**
Random effect (intercepts)	Variance	*SD*				
Task	<.001	<.001				
Information criterion	Estimate					
LogLikelihood	-100.12					
DIC	200.25					
AIC	210.25					
BIC	221.70					
R^2	Estimate					
Marginal	.20					
Conditional	.21					

Note. DIC = Deviance Information Criterion; AIC = Akaike Information Criterion; BIC = Bayesian Information Criterion

Figure 1
Summary of Study 2

Table 3

Inspections of Interactions between Instruction and IDs (Accentedness)

Variable	Estimate	*SE*	t-value	*p*	95% *CI*	
					Lower	Upper
(Intercept)	**6.11**	**1.26**	**4.86**	**<.001***	**4.87**	**10.01**
Time	.455	.656	.694	.492	-2.50	2.27
Sound sequence recognition	-.006	.010	-.612	.543	-.035	.005
Associative memory	<.001	.007	.104	.917	-.004	.026
Phonemic coding ability	**.017**	**.005**	**3.162**	**.003***	-.001	.022
Ideal L2 self	.079	.120	.660	.512	-.278	.213
Ought-to L2 self	.131	.135	.966	.339	-.200	.353
Anxiety	**-.803**	**.250**	**-3.22**	**.002***	**-1.42**	**-.400**
Time:Sound sequence recognition	.005	.005	.922	.361	-.014	.023
Time:Associative memory	-.001	.004	-.292	.772	-.022	.005
Time:Phonemic coding ability	-.001	.003	-.504	.617	-.008	.012
Time:Ideal L2 self	.078	.062	1.24	.220	-.230	.225
Time:Ought-to L2 self	-.024	.0706	-.337	.738	-.233	.281
Time:Anxiety	.046	.130	.355	.724	-.174	.774

Random effect (intercepts)	Variance	*SD*
Subject	.569	.755

Information criterion	Estimate
LogLikelihood	-126.56
DIC	253.12
AIC	285.12
BIC	327.12

R^2	Estimate
Marginal	.30
Conditional	.70

Table 4

Inspections of Interactions between Instruction and IDs (Comprehensibility)

Variable	Estimate	*SE*	t-value	*p*	95% *CI*	
					Lower	Upper
(Intercept)	**7.44**	**1.39**	**5.34**	**<.001***	**3.79**	**8.44**
Time	-.115	1.27	-.089	.930	-.762	1.67
Sound sequence recognition	-.015	.011	-1.39	.168	-.024	.012
Associative memory	.011	.008	1.40	.166	-.013	.014
Phonemic coding ability	.011	.006	1.83	.072	.007	.027
Ideal L2 self	-.033	.133	-.246	.806	-.143	.301
Ought-to L2 self	.076	.150	.509	.613	-.120	.381
Anxiety	**-.911**	**.277**	**-3.29**	**.002***	**-1.27**	**-.341**
Time:Sound sequence recognition	.004	.010	.456	.650	-.005	.014
Time:Associative memory	-.009	.007	-1.17	.250	-.008	.006
Time:Phonemic coding ability	.002	.006	.372	.712	-.007	.004
Time:Ideal L2 self	-.002	.123	-.020	.984	-.038	.194
Time:Ought-to L2 self	.024	.139	.175	.862	-.155	.107
Time:Anxiety	.300	.256	1.17	.247	-.196	.288
Random effect (intercepts)	Variance	*SD*				
Subject	.696	.834				
Information criterion	Estimate					
LogLikelihood	-91.23					
DIC	182.5					
AIC	214.5					
BIC	256.5					
R^2	Estimate					
Marginal	.36					
Conditional	.91					

Supporting Information I

Items and Descriptive Statistics of Participants' IDs profiles

Table 1 *Descriptive statistics of individual differences among 73 Japanese students at the onset of the study*

	M	*SD*	*Range* *Min–Max*
Language Learning Aptitude			
Sound sequence recognition (0–75 %)	25.316	14.86	0–60
Associative memory (0–100 %)	58.97	19.45	20–95
Phonemic coding ability (0–100 %)	70.95	23.84	20–100
L2 pronunciation specific motivation and anxiety			
Ideal L2 self	3.29	.97	3.81–4.11
Ought-to L2 self	3.6	.57	2.57–1.92
Anxiety	3.29	.97	3.81–4.11
Past L2 experience[a]			
Past English learning inside the classroom	1522.24	419.9	834.24–2502.72
Past English learning outside the classroom	486.64	563.47	0–2763.42
Recent L2 experience[b]			
Recent English learning inside the classroom	225.53	121.66	0–469.26
Recent English learning outside the classroom	182.08	176.47	0–729.96

Note. a. Past L2 experience was calculated based on the total weeks they engaged in learning English during elementary, junior high, and high school. Inside the classroom indicates that all the English lessons they received as regular English classes in the schools, while outside the classroom refers to communication with native and non-native speakers in English that was taken place outside elementary, junior high, and high schools.
b. Recent L2 experience was calculated based on the total weeks they engaged in learning English and communication with native and non-native speakers in English since they entered the university.

Table 2 *Items used in Motivation and Anxiety Questionnaire and Its Descriptive Statistics*

	M	SD	Range Min–Max

1. Questionnaire items of pronunciation specific anxiety			
Fear of negative evaluation related to pronunciation			
I (would) feel uneasy pronouncing English sounds and/or words with a Japanese accent.	4.1	1.42	1–6
I would rather others do not hear me making pronunciation mistakes.	3.63	1.5	1–6
I fear others might find my pronunciation of English strange or funny.	3.4	1.41	1–6
I am worried what others might think of me when they hear my English pronunciation.	3.59	1.48	1–6
I get nervous and feel shy when making a pronunciation mistake.	3.51	1.36	1–6
I feel stressed knowing that others are listening to me.	2.81	1.42	1–6
I feel more embarrassed making a pronunciation mistake that any other type of mistake when I speak in English.	2.6	1.29	1–6
Pronunciation self-efficacy and self-assessment			
I find it more difficult to improve pronunciation than grammar or vocabulary.	3.24	1.48	1–6
I remember the pronunciation of new words easily.	3.7	1.21	2–6
My pronunciation is at a lower level than that of people around me.	3.78	1.34	1–6
I am satisfied with my present level of English pronunciation.	4.48	1.56	1–6
I have a talent to pick up the pronunciation of English.	3.97	1.29	1–6
My pronunciation of English is far from acceptable.	3.35	1.18	1–6
Pronunciation self-image			
I look funny pronunciation ‘th’ sound.	2.57	1.29	1–6
I like singing and/or speaking to myself in English.	3.23	1.62	1–6
Sometimes I like to imitate English actors/singers.	3.67	1.48	1–6
I do not like listening to myself reading English aloud.	3.03	1.31	1–6
I think I sound unnatural speaking English.	3.75	1.09	1–6
I look natural speaking English.	3.89	1.4	1–6
Belief related to the anxiety of pronunciation of English			
The comprehensibility of a speaker depends on his/her level of proficiency.	4.49	1.08	1–6
Some words in English sound funny and /or awkward.	3.29	1.24	1–5
The pronunciation of English is difficult for Japanese.	4.46	1.12	1–6
The level of pronunciation affects the ability to understand spoken language	4.24	1.32	1–6
2. Questionnaire items of pronunciation specific motivation			
***Ideal L2-self* related to pronunciation**			
I can imagine myself living abroad and having a discussion in English with accented but comprehensible pronunciation.	3.9	1.27	1–6
I can imagine a situation where I am speaking with foreigners in English with accented but comprehensible pronunciation.	4.11	1.18	1–6

I imagine myself as someone who is able to speak English with accented but comprehensible pronunciation.	4.08	1.21	1–6
Whenever I think of my future career, I imagine myself using English with accented but comprehensible pronunciation.	3.99	1.13	2–6
I can imagine myself living abroad and having a discussion in English with nativelike pronunciation.	3.83	1.2	1–6
I can imagine a situation where I am speaking with foreigners in English with nativelike pronunciation.	3.81	1.2	1–6
I imagine myself as someone who is able to speak English with nativelike pronunciation.	3.86	1.2	1–6
Whenever I think of my future career, I imagine myself using English with nativelike pronunciation.	3.81	1.28	1–6
***Ought-to L2 self* related to pronunciation**			
I study English pronunciation to speak English with accented but comprehensible pronunciation because close friends of mine think it is important.	3.51	1.45	1–6
I have to study English pronunciation to speak English with accented but comprehensible pronunciation, because if I do not study it, I think my parents will be disappointed in me.	2.19	1.29	1–5
Speaking English with accented but comprehensible pronunciation is necessary because people surrounding me expect me to do so.	2.75	1.52	1–6
My parents believe that I must be able to speak English with accented but comprehensible pronunciation to be an educated person.	2.41	1.4	1–5
I study English pronunciation to speak English with near native-like pronunciation because close friends of mine think it is important to speak English with near native-like pronunciation.	2.94	1.56	1–6
I have to study English pronunciation to speak English with near native-like pronunciation, because if I do not study it, I think my parents will be disappointed in me.	2.21	1.45	1–6
Speaking English with near native-like pronunciation is necessary because people surrounding me expect me to do so.	2.84	1.62	1–6
My parents believe that I must be able to speak English with near native-like pronunciation to be an educated person.	2.41	1.49	1–6

Supporting Information II

Results of Correlation among IDs Variables

Table 1 *Results of Correlation Analysis among Aptitude Variables*

	2		3	
	r	*p*	*r*	*p*
1. Sound sequence recognition	.052	.662	.131	.271
2. Rote and associative memory			.188	.110
3. Phonemic coding ability				

Note. Statistical significance at a $p < .025$ (Bonferroni corrected)

Table 2 *Results of Correlation Analysis among Motivation and Anxiety*

	2		3	
	r	*p*	*r*	*p*
1. Ideal L2 self	.116	.159	-.238	.043
2. Ought to L2 self			.147	.215
3. Anxiety				

Note. Statistical significance at a $p < .025$ (Bonferroni corrected)

Table 3 *Results of Correlation Analysis among L2 Experience Variables*

Experience variables	2		3		4	
	r	*p*	*r*	*p*	*r*	*p*
Past L2 experience						
1. Past English learning inside the classroom	-.039	.744	.010	.932	-.062	.604
2. Past English learning outside the classroom			-.07	.557	-.1	.399
Recent L2 experience						
3. Recent English learning inside the classroom					.31	.08
4. Recent English learning outside the classroom						

Note. * p <.017 (Bonferroni corrected)

Table 5 *Results of Correlation Analysis between Aptitude variables and Motivation and Anxiety Factors*

Motivation and anxiety factors	Aptitude factors					
	Sound sequence recognition		Associative memory		Phonemic coding ability	
	r	*p*	*r*	*p*	*r*	*p*
Ideal L2 self	.076	.521	-.216	.066	-.231	.05
Ought to L2 self			-.110	.354	.004	.974
Anxiety					.040	.739

Note. * p <.017 (Bonferroni corrected),

Table 6 *Results of Correlation Analysis between L2 Experience Variables and Motivation and Anxiety Factors*

Experience factors	Motivation and anxiety variables					
	Ideal L2 self		Ought to L2 self		Anxiety	
	r	*p*	*r*	*p*	*r*	*p*
Past L2 experience						
1. Past English learning inside the classroom	.147	.215	.001	.933	.016	.893
2. Past English learning outside the classroom	.087	.466	-.124	.297	-.222	.059
Recent L2 experience						
3. Recent English learning inside the classroom	.140	.236	-.017	.886	-.135	.254

4. Recent English learning outside the classroom	.226	.054	.192	.103	-.055	.646

Note. * $p < .017$ (Bonferroni corrected)

Table 7 *Results of Correlation Analysis between L2 Experience Variables and Aptitude Variables*

Experience factors	Cognitive factors					
	Sound sequence recognition		Associative memory		Phonemic Coding ability	
	r	*p*	*r*	*p*	*r*	*p*
Past L2 experience						
Past English learning inside the classroom	-.011	.928	.085	.476	.052	.662
Past English learning outside the classroom	-.993	<.001*	.074	.534	.126	.29
Recent L2 experience						
Recent English learning inside the classroom	.017	.889	-.117	.326	.112	.346
Recent English learning outside the classroom	-.069	.562	-.056	.636	.119	.316

Note. * $p < .017$ (Bonferroni corrected)

Supporting Information III

Speech Tasks

Task A

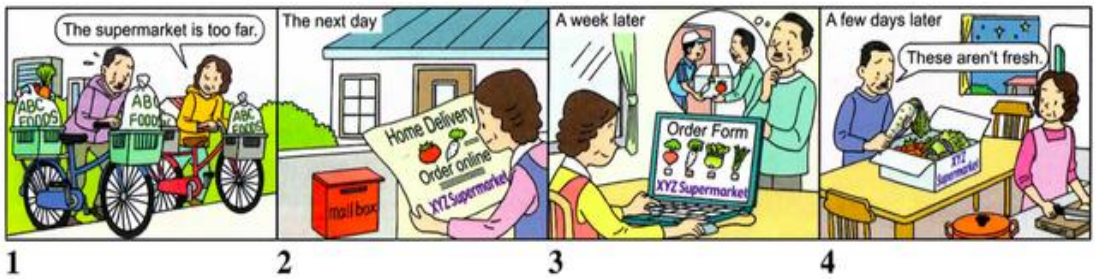
You have **one minute** to prepare.
This is a story about a girl who wanted a smartphone.
You have **two minutes** to narrate the story.

Your story should begin with the following sentence:
One day, a girl was at home with her parents.

Task B

You have **one minute** to prepare.
This is a story about an elderly couple who lived far away from the nearest supermarket.
You have **two minutes** to narrate the story.

Your story should begin with the following sentence:
One day, an elderly couple was coming home from the supermarket.



Supporting Information IV

Training Scripts and a Sample of Rating Scales in the Booklet

A. Training scripts (adopted from Trofimovich & Isaacs, 2012)

- Comprehensibility: The term comprehensibility refers to how difficult it is to understand what the speaker is saying. If you can understand what the speaker is describing (a story) easily regardless of his or her accent, the speech is regarded highly comprehensible. However, if you need effort to understand the speech or barely catch what is being said, then his or her speech has low comprehensibility.

- Nativelikeness: The term refers to how heavily a speaker's speech is affected by his/her native language. If you hear any features that are not in the native variety, then the speech has high foreign accentedness.

B. A Sample texts from the rating booklet

- Comprehensibility

Difficult to understand understand	1	2	3	4	5	6	7	8	9	Easy to understand

- Nativelikeness

Heavily accented at all	1	2	3	4	5	6	7	8	9	Not accented at all

Supporting Information V

Model Comparisons

Table 1 *Summary of Model Fits of Accentedness in Comparison to a Model with No Fixed Effect*

Variable	AIC	χ^2	*p*
Task (intercept)	213.67	n.a.	n.a.
In comparison to a model with Intercept + no Fixed Effect			
Sound sequence recognition	214.69	.98	.32
Associative memory	215.67	<.001	.99
Phonemic coding ability	206.86	8.81	.003*
Ideal L2 self	213.36	2.31	.13
Ought to L2 self	215.62	.045	.83
Anxiety	208.24	7.44	.006*
Past English learning inside the classroom	215.65	.019	.89
Past English learning outside the classroom	215.37	.299	.59
Recent English learning inside the classroom	205.72	9.95	.002*
Recent English learning outside the classroom	185.01	30.66	<.001*
In comparison to a model with Intercept + Recent English learning outside the classroom			
Recent English learning outside the classroom + Phonemic coding ability	181.86	5.15	.023*
Recent English learning outside the classroom + Anxiety	181.23	5.78	.016*
Recent English learning outside the classroom + Recent English learning inside the classroom	184.24	2.77	.1
In comparison to a model with Intercept + Recent English learning outside the classroom + Phonemic coding			
Recent English learning outside the classroom + Phonemic coding ability + Anxiety	176.24	7.62	.006*

Note. * $p < .05$

Table 2 *Summary of Model Fits of Comprehensibility in Comparison to a Model with No Fixed Effect*

Variable	AIC	χ^2	*p*
Task (intercept)	222.04	n.a.	n.a.
In comparison to a model with Intercept + no Fixed Effect			
Sound sequence recognition	224	.04	.84
Associative memory	223.69	.35	.55
Phonemic coding ability	220.14	3.9	.048*
Ideal L2 self	223.85	.19	.66
Ought to L2 self	223.16	.88	.35
Anxiety	221.06	2.98	.08
Past English learning inside the classroom	223.79	.25	.62
Past English learning outside the classroom	223.06	.99	.32
Recent English learning inside the classroom	214.59	9.45	.002*
Recent English learning outside the classroom	213.06	10.44	.001*
In comparison to a model with Intercept + Recent English learning outside the classroom			
Recent English learning outside the classroom + Recent English learning inside the classroom	210.25	5.35	.02*
Recent English learning outside the classroom + Phonemic coding	213.15	2.45	.12

Note. * $p < .05$

Supporting Information VI

Contents of Four Pronunciation Interventions

Based on the procedure and instruction used by Couper (2003), the four interventions involved the following three stages:

- First, the target items were introduced, and how to produce a particular feature was explicitly explained. Segmental features that were covered in the intervention were /b/, /v/, /z/, /ð/, /θ/, /r/, /l/, /s/, and /ʃ/ because they have been regarded as problematic for intelligibility among Japanese learners of English (Saito, 2011). In the case of the phonemes, graphical representations and explanations of the place and manner of articulation were also given (i.e., *articulatory-based instruction*).

- In the next stage, using multiple sound examples, the participants were asked to discriminate the target items (e.g., /r/ vs. /l/) with peers several times. Then, in order to compare their pronunciation with the models, they recorded themselves with their mobile phones (i.e., *auditory-based instruction*). Although feedback was provided between peers, the instructor was constantly monitoring the participants' performance and helped them produce the target forms if necessary.

- The third stage of the instruction involved meaning-oriented communication activities. Each session offered (a) a simple topic (e.g., "what is the last movie you watched?") to engage in a 4-3-2 activity (i.e., a type of fluency enhancement; De Jong & Perfetti, 2011; Tran & Saito, 2021), and (b) an argumentative topic to help the participants engage in a meaning-oriented communication (instead of a mundane drill or a simple greeting). Prior to the activity, the instructor reminded the participants the certain features they should attend to when listening to peers and producing speech by themselves. If necessary, recasts were used as a form of corrective feedback.