



The Role of Cognitive, Affective, and Experiential Individual Differences on L2 Pronunciation Production: Implications for the ESL Classroom

Individual differences (IDs) refer to the cognitive, affective, and experiential factors which determine variability in language learning. While ID research has had a strong appearance in second language acquisition, its presence in L2 pronunciation research has been sporadic. This study investigated a range of IDs, including cognitive (phonological short-term memory), affective (motivation and anxiety), and experiential (age, length of residence and English use) on pronunciation production. Speech data were collected from twenty ESL learners, and their patterns of segmental and suprasegmental performance were measured. Results of linear mixed effect modeling demonstrated an overall negative effect of age, an overall positive effect of motivation and L2 English use, and a dynamic and complex effect of phonological short-term memory, anxiety, and length of residence. Implications are discussed with respect to how IDs can be considered in pronunciation teaching and learning.

Keywords: individual differences, pronunciation, second language acquisition, segmentals, suprasegmentals

Introduction

Core to any discussion of second language learning and teaching is the great variability learners demonstrate when learning a second language (L2). This variability has been well documented in the large body of individual difference (ID) research (Dörnyei, 2005, 2006, 2009) within studies of second language acquisition (SLA), with special attention given to motivation and aptitude as the strongest predictors of second language success (Dörnyei & Skehan, 2003). However, while ID theory has been prominent in SLA research, the impact of IDs on L2 pronunciation has been much less documented. With age being ostensibly one of the most important factors in L2 pronunciation acquisition, many adult learners defy the effects of age and demonstrate exceptional L2 pronunciation, likely traceable back to a host of cognitive, socio-affective, and experiential IDs (Moyer, 2014; Nagle, 2022; Saito et al., 2018). In an English as a Second Language (ESL) context where English is used daily outside of the classroom, IDs can greatly determine the extent to which one acquires fundamental pronunciation skills for day-to-day communication with English users. Specifically, within an ESL speaking/pronunciation classroom, educators may notice great variability in pronunciation performance, despite the fact that all learners receive the same instruction. These differences may manifest themselves via segmental (i.e., consonant and vowel) and/or suprasegmental (i.e., temporal fluency, word stress, rhythm, sentence stress/prominence, intonation, pitch) use. It is with this backdrop in mind that the current study seeks to investigate how a host of cognitive, affective, and experiential IDs impact segmental and suprasegmental speech properties for L2 learners within an ESL context. Informed by the current findings, implications and recommendations are provided to ESL pronunciation educators as a way to contextualize and address learner IDs in the classroom.

Cognitive Individual Differences

Cognitive IDs are involved in L2 pronunciation acquisition given the nature of how L2 phonetic input is encoded and stored within one's mental representation (Nagle, 2022). Aptitude and the working memory system have been prioritized in this link between cognitive IDs and pronunciation performance (Moyer, 2014). Aptitude has often been measured via designated aptitude tests, such as the Pimsleur Language Aptitude Battery (PLAB) (see Baker Smemoe & Haslam, 2013) or the LLAMA (see Saito, 2017; Saito et al., 2019). Published first in 1959, the Modern Language Aptitude Test (MLAT) is also a widely drawn upon test of foreign language aptitude. For working memory, or the cognitive system which manages the simultaneous processing, manipulation, and storage of information, some researchers (e.g., Ahmadian, 2012; Isaacs, 2010; O'Brien et al., 2007; Purcell & Suter, 1980; Suter, 1976) have explored the link between aspects of phonological short-term memory (PSTM) and pronunciation variance. PSTM, also known as the phonological loop, is one of the most promising components of the working memory system due to the fact that it stores information encoded in the form of speech sounds. The phonological loop is a combined functionality of the phonological store (which holds speech-based data for up to a few seconds) and a rehearsal process (Baddeley, 2003). The phonological loop temporarily stores speech-based input as it works towards becoming a recognized, phonological system (Baddeley et al., 1998). Through the assistance of the phonological loop, target-like sounds and sequences are fed into the short-term memory store where they gradually lead to more accurate mental representations of the sound structure. However, the incorporation of this psycholinguistic construct into L2 ID pronunciation research is still in its early stages, and several advancements are to be made.

Affective Individual Differences

Motivation is likely one of the most discussed IDs in both SLA and pronunciation research; alongside aptitude, it has been one of the strongest predictors of L2 language learning success (Dörnyei & Skehan, 2003). Dörnyei's (2005, 2009) tri-partite L2 Motivational Self System has become a well-drawn upon framework of L2 motivation to capture the different selves of the learner. The ideal L2 self speaks to the personal goals, wishes, and aspirations a learner has with respect to their L2 in the future. The ought-to L2 self is responsible for the motivation the learner has due to outside pressures, obligations, and responsibilities from people, family, and society (Dörnyei, 2013). The L2 learning experience is accounted for by a learner's immediate learning environment and deals with current motivated behavior. Due to the ought-to L2 Self's inconsistent role in some studies, researchers (e.g., Papi et al., 2019; Teimouri, 2017) have made methodological revisions to capture its complexity. Papi et al. (2019) validated the bifurcation of both the ideal L2 self/*own* and *other* and the ought-to L2 self/*own* and *other*, while developing a neutral (no promotion/prevention regulatory focus) measure of motivated behavior. The bifurcated selves provide an additional distinction between personal and other-oriented futuristic goals and personal and other-oriented duties, obligations, or responsibilities, allowing more precise theoretical predictions about the behaviors of learners and ultimately their L2 use (Papi & Hiver, 2022).

Empirical investigations between motivation and L2 pronunciation are scattered and tend to draw on multiple motivational frameworks. Early work by Smit and Dalton (2000) found that pronunciation learning for English learners in Austria was affected by both intrinsic and extrinsic motivation. Another early work by Moyer (1999) measured L2 German learners' motivation towards their professional and personal current and long-term goals for the purposes of studying German. Moyer found professional motivation to account for up to 41% of the variance in listeners' ratings of the German learners' speech. After these earlier works, motivation research was lacking until more recent research by Saito and colleagues. In 2016, Saito and Hanzawa collected integrative and instrumental motivation data from Japanese learners of English and found that professional and integrative levels of motivation varied quite a bit from learner to learner, although they did not have an effect on listener speech ratings. Additionally, Saito et al. (2017) linked Japanese EFL students' motivation (instrumentality, integrativeness, and L2 learning metacognition) to improvements in listener-rated comprehensibility (but not accentedness) over the course of a semester. Finally, Saito et al. (2018) found that Japanese EFL students' improvement in listener-rated comprehensibility was impacted by their ideal L2 self, in addition to emotion and language contact. In general, motivation as a key to improved L2 pronunciation has been well-discussed (e.g., Baran-Łuczarska, 2012; Celce-Murcia et al., 2010, pp. 15-22; Hansen Edwards, 2018; Levis, 2005; Moyer, 2014; Saito et al., 2018; Smit & Dalton, 2000), yet empirical research to support this discussion has been lagging.

Language anxiety, another affective ID, is the most commonly studied emotion in SLA (Gregersen et al., 2014; MacIntyre, 2017); it "captures the worry and usually negative emotional reaction aroused when learning or using

an L2" (MacIntyre, 2007, p. 565). Comparisons have been made between the effect of anxiety on ESL learners in an ESL vs. an English as a Foreign Language (EFL) context. For example, Woodrow (2006) underscored the anxiety involved in the use of the target language in everyday ESL situations. Woodrow found a significant negative relationship between anxiety and L2 speaking (fluency, language use, and pronunciation). On the other hand, Sardegna et al. (2014) found that pronunciation anxiety and nervousness did not have a strong influence for Korean EFL learners, as they were not obliged to speak English regularly outside the classroom. While some studies have explored links between anxiety and aspects of pronunciation (e.g., Baran-Lucarz, 2012; Kessler, 2010; Saito et al., 2018), most have done so within EFL contexts.

Experiential Individual Differences

Age has been a factor which has notoriously received great attention in SLA, but this is especially true of L2 pronunciation given that "phonology relies on both speech-motor control and auditory-perceptual neural networks" (Moyer, 2014, p. 419). According to Scovel (1988), phonology is skill which depends largely on neuro-muscular development; that is, once someone's L1 sound system has been established, shifts in brain plasticity and neuro-muscular flexibility make it more challenging for both perceptive and productive skills. Flege's Speech Learning Model (1995) has also provided considerable attention to the effect of age on how phonetic categories of one's L1 become more defined.

In addition to age, another aspect of one's experience that should be considered, specifically within an ESL context, is length of residence (LOR). Some research has shown that LOR is related to accentedness, with longer-residing learners having less of an accent over time (see Moyer, 2011; Trofimovich & Baker, 2006). However, Trofimovich and Baker (2006) also found that non-native production of suprasegmentals was a feature of accented speech for all levels of LOR (3 month, 3 years, and 10 years), affirming the challenging nature of suprasegmental production. It should be noted that LOR is not a relevant construct in the EFL context and that there are of course highly advanced users of English who have never resided in an English-speaking country. Under such circumstances, other parallel constructs can be invoked, such as the length of time since first exposure to English, age of first exposure to the language, or quality of language use (Hansen Edwards, 2018).

Finally, language use, especially within an ESL context, varies from learner to learner and can affect both how learners acquire speech and how their speech is perceived by listeners (Hansen Edwards, 2018). English use can be both interactive and non-interactive. Interactive use, especially pertinent to an ESL context, involves interlocutors (e.g., friends, partners, roommates, strangers, etc.). Non-interactive use involves participating in media-related activities (e.g., television, movies, videos, gaming, and music). As an example, Derwing et al. (2008) found that language use was related to improvements in fluency and comprehensibility for speakers of various Slavic languages learning English. In the same study, media consumption (English talk radio and English television) was also linked to improvements in L2 fluency. What is notably lacking from the research, however, is the effect of L2 English use on a broader range of segmental and suprasegmental properties.

The Current Study

IDs, which are crucial to language learning in general, become even more critical to L2 pronunciation due to the age effect on L2 pronunciation acquisition and attainment (Moyer, 2014). However, empirical ID research has been surprisingly limited in L2 pronunciation research (Thomson & Derwing, 2015), especially when compared with its momentum in SLA research. The current study, exploratory in nature, is one of the first to evaluate the effect of a range of IDs on L2 pronunciation within a single design and within an ESL context. The current study is therefore guided by one major research question: To what extent do cognitive (PSTM), affective (motivation and anxiety) and experiential (age, LOR, and English use) individual differences impact second language English production of segmental and suprasegmental properties (i.e., segmental deviations, pausing, speech rate, sentence stress, varied intonation, level intonation, and pitch range)?

Methods

Participants

A total of twenty L2 English learners were recruited for this study. Ten learners were recruited from an intensive English program (IEP) at a major American university, and ten learners were recruited from the surrounding community in which the IEP was located. Efforts were made to represent the range and diversity of adult English language users in the targeted ESL community (see Norris & Ortega, 2000). A total of 12 males and 8

females participated from diverse L1 backgrounds ($n = 7$ Arabic; $n = 4$ Korean; $n = 1$ Spanish; $n = 2$ Russian; $n = 2$ Chinese; $n = 3$ Turkish; and $n = 1$ French). The youngest learner was 18 years old, and the oldest was 71 years old ($M = 30.80$). The average LOR was 5 years.

Materials and Instruments

Recorded Listening Stimuli

A spontaneous speech task was used to elicit speech stimuli from the learners. The topic of the task was to discuss differences between the learners' home city and the American city in which they currently lived. The recordings were cut at word boundaries around 45 seconds.

Phonological Short-term Memory

A forward and backward non-word repetition task was designed (see Appendix A) to assess PSTM. A non-word repetition test is a true test of phonological short-term memory (Baddeley et al., 1998; Gathercole, 2006). This test contained increasingly longer strings of nonsense words, exhibiting the following phonemic pattern: consonant, vowel, consonant, vowel. All non-words had primary stress on the first syllable and ended either with the unstressed vowel (schwa, /ə/), or a diphthong. Efforts were made to ensure that the non-words did not exist in the learners' L1s through several rounds of verification with native speakers of the L1 groups represented in the sample.

A member of the research team met with the learners individually and played the recording, pausing the recording after each string in order to give the learners the chance to repeat the non-words (in the order they heard them in on the forward test and in reverse order on the backward test). The learners heard the recorded strings only once. The repetitions were audio-recorded in order to verify their accuracy. The test was terminated at the level at which a learner could no longer repeat the non-words accurately. An accuracy score was calculated dichotomously ("correct" = 1 or "incorrect" = 0) for each non-word within a given string (e.g., Gathercole & Baddeley, 1989; Gathercole et al., 1992). The total of correctly repeated non-words for the forward test was added to the total of correctly repeated words for the backward test, and the total accuracy score was obtained to measure a speaker's capacity of the phonological loop. Several reliability indices were established with a trained psycholinguist coder who scored 30% of the non-word repetition tests: $\kappa = .733$; percentage of agreement = 89.5%; and intraclass correlation coefficient = .846.

Motivation

The instrument used to capture learners' pronunciation-oriented motivation was adapted from Papi et al.'s (2019) bifurcated version of Dörnyei's (2005, 2009) L2 Motivational Self System, previously validated in an ESL context. Our study additionally adapted Papi et al.'s (2019) neutral measure (no promotion/prevention regulatory focus) of motivated behavior. Although similar wording was used as in Papi et al.'s (2019) study, our motivation instrument (see Appendix B) was oriented towards L2 pronunciation and was aligned with the current thinking about the importance of intelligibility and comprehensibility as opposed to accentedness (Levis, 2005). In the end, the 18-item motivation survey was made up of 5 sub-scales: ideal L2 self/*own* (4 items; $\alpha = .77$); ideal L2 self/*other* (4 items; $\alpha = .91$); ought-to L2 self/*own* (3 items; $\alpha = .80$); ought-to L2 self/*other* (2 items; $\alpha = .76$); and motivated behavior (5 items; $\alpha = .81$).

The ideal L2 self/*own* focused on a learner's personal goals, hopes, and aspirations for pronunciation use, and the ideal L2 self/*other* focused on the goals, hopes, and aspirations the learner's significant others (e.g., family, parents, etc.) had for the learner and their pronunciation use. The ought-to self/*own* referred to the duties, responsibilities, and obligations a learner had with respect to their pronunciation use, and the ought-to self/*other* referred to the duties, responsibilities, and obligations a learner's significant others had for the learner. Motivated behavior focused on the learner's current time and investment dedicated to pronunciation learning. Due to the significant and strong correlations among the motivation sub-scales ($.45 < R < .86$), the scores from the motivation sub-scales were summed into a composite score operationalizing the larger construct of "Motivation."

Anxiety

The anxiety instrument (see Appendix C), adapted from Woodrow (2006) and Sardegna et al. (2014), measured pronunciation-related negative affect and ESL anxiety-provoking situations. The anxiety survey contained 7 items ($\alpha = .86$) which evaluated the extent to which the learners felt anxious in specific situations in

which they were required to speak English. In response to each situation, the learners expressed how anxious they felt using the same 1-5 point scale and descriptors as Woodrow (2006): 1= not at all anxious; 2 = slightly anxious; 3 = moderately anxious; 4 = very anxious; 5 = extremely anxious.

Age, Length of Residence, and L2 English Use

Age and LOR were simple quantifications of the learners' current age as well as their length of residence within the United States. The L2 English use survey (see Appendix D) drew on Taguchi et al.'s (2016) social contact questionnaire and developmental processes. The interactive L2 English use portion consisted of 8 items which measured communicative use of the L2 with different types of interlocutors (see Derwing et al, 2008). The non-interactive portion consisted of 5 items which measured time spent engaging with media (TV, movies, videos, gaming, music). The survey was administered to the learners over three weeks, once per week. Mean scores for each item across the three weeks were calculated, and reliability testing was conducted on the total of these 3-week averages for both interactive ($\alpha = .82$) and non-interactive L2 English use ($\alpha = .80$), which were combined into the larger construct of L2 English use.

Data Collection

In the first week, the learners' background and L2 English use were the only data collected. In the second week of data collection, speech stimuli, motivation data, anxiety data, L2 English use, and the non-word repetition test were all collected. Finally, only L2 English use was collected in the third week.

Segmental and Suprasegmental Analyses

Segmental and suprasegmental analyses were conducted based on previous work (e.g., Kang, 2010; Kang & Moran, 2014; Kang et al., 2022; Kormos & Dénes, 2004; Pickering, 2018), using Brazil's (1997) framework for prosodic analyses. All segmental and suprasegmental operationalizations and calculations can be found in detail in Appendix E. The total occurrences of segmental deviations (instances when consonants and vowels deviated from target-like sounds) were combined as one representative measure of segmental performance. Other representative measures of suprasegmental performance were analyzed, including *pausing* (total number of pauses in the speech sample); *speech rate* (articulation rate measured as number of syllables produced in the speech sample divided by the total length of the speech sample, excluding pause time); *sentence stress* (pace measured by total number of prominent syllables divided by the total number of runs in the speech sample); *varied intonation* (total number of rising, fall-rising, falling, and rise-falling tones); *level intonation* (total number of level tones), and *pitch range* (the difference between the highest and lowest values of pitch as determined on prominent syllables).

Data Analyses

To respond to the research question concerning the effect of learner IDs on pronunciation production, a linear mixed-effect modeling (LMEM) procedure was run. The LMEM treated the learners as random effects (accounting for their variance) and IDs (PSTM, motivation, anxiety, age, LOR, and English use) as fixed effect covariates. The pronunciation properties (segmental deviations, pauses, speech rate, sentence stress, varied intonation, level intonation, and pitch range) were the dependent variables being predicted by the models.

Results

Seven separate LMEMs were fitted to investigate the relationship of the IDs with the following seven pronunciation properties: (1) segmental deviations, (2) pauses, (3) speech rate, (4) sentence stress, (5) varied intonation, (6) level intonation, and (7) pitch range. Appendix F displays the summary of the LMEM results. All IDs significantly predicted the learners' segmental deviations, however with varying directionalities. Increases in age, LOR, and anxiety predicted more instances of segmental *inaccuracy*, while increases in motivation, PSTM, and L2 use predicted more instances of segmental *accuracy*. Specifically, 70.2% of variance in the model was explained by the IDs (*marginal R*² = .702) with even more, 85.1%, explained when considering the random effects (*conditional R*² = .851), i.e., other sources of individual variability for each separate learner that we did not explicitly measure.

All IDs predicted the pausing patterns of the learners, again with varying directionalities (*marginal R*² = .319 or 31.9% variance; *conditional R*² = .659 or 65.9% variance). The only two variables to have an inverse relationship with pausing were age and L2 use (older learners with more language use paused less frequently). For the

remaining ID variables, the relationship was positive: learners with longer LOR, higher motivation, higher PSTM scores, and more anxiety paused more frequently.

Speech rate was significantly predicted by all IDs, with differing relationships and a strong effect (*marginal R²* = .641 or 64.1% variance; *conditional R²* = .820 or 82% variance). The only two IDs which had a positive relationship with speech rate were motivation and L2 use (i.e., more motivation and language use predicted a faster speech rate). The remainder of the IDs negatively predicted the dependent variable: older learners with longer LOR, higher scores on the PSTM test, and more anxiety all predicted a slower speech rate.

For sentence stress, all ID variables predicted this suprasegmental feature except motivation. Effect sizes were also strong: *marginal R²* = .658 or 65.8% variance; *conditional R²* = .829 or 82.9% variance. Older learners and those with more L2 use used more sentence stress. However, learners with a longer LOR, higher PSTM scores, and more anxiety used less sentence stress.

For varied intonation, older age and PSTM scores predicted *fewer* patterns of varied intonation. At the same time, older learners and PSTM scores predicted *more* level intonation patterns. These results are compared to LOR, motivation, anxiety, and L2 use which had a significant and *positive* relationship with diverse intonation. Again, these results stand in logical opposition to those for level intonation patterns which had a *negative* relationship: longer LOR, more motivation, and more anxiety predicted fewer level tones. L2 use was not a significant predictor of level intonation as it was for varied intonation. The IDs in the model of varied intonation accounted for slightly more variance (*marginal R²* = .476 or 47.6% variance; *conditional R²* = .738 or 73.8% variance) than for level intonation (*marginal R²* = .335 or 33.5% variance; *conditional R²* = .667 or 66.7% variance).

Finally, for the results dealing with pitch range, all ID variables predicted the dependent variable (*marginal R²* = .318 or 31.8% variance; *conditional R²* = .659 or 65.9% variance). Age and L2 use were the only two variables with a negative relationship (i.e., older learners and more language use predicted a more narrow pitch range). The remaining ID variables had a positive relationship with the dependent variable: a longer LOR, higher motivation, higher PSTM scores, and more anxiety all predicted a wider pitch range.

To provide a visual summary, the results of the linear mixed-effects modeling are synthesized below in Figure 1. The “+” sign in the figure depicts a positive relationship with the pronunciation property, while the “-” sign depicts a negative relationship. N/S refers to a non-statistically significant outcome. For example, an increase in age was related to more segmental deviations, fewer pauses, slower speech rate, more sentence stress, less varied intonation, and more level intonation.

Figure 1
Associations of IDs with Segmental and Suprasegmental Properties

| + segmental deviations | + segmental deviations | - segmental deviations | - segmental deviations | + segmental deviations | - segmental deviations |
|------------------------|------------------------|------------------------|------------------------|------------------------|------------------------|
| - pauses | + pauses | + pauses | + pauses | + pauses | - pauses |
| - speech rate | - speech rate | + speech rate | - speech rate | - speech rate | + speech rate |
| + sentence stress | - sentence stress | N/S | - sentence stress | - sentence stress | + sentence stress |
| - varied intonation | + varied intonation | + varied intonation | - varied intonation | + varied intonation | + varied intonation |
| + level intonation | - level intonation | - level intonation | + level intonation | - level intonation | N/S |
| - pitch range | + pitch range | + pitch range | + pitch range | + pitch range | - pitch range |

Increases in ID



Age

LOR

Motivation

PSTM

Anxiety

English Use

Discussion and Pedagogical Implications

This exploratory study investigated the effect of a host of cognitive, affective, and experiential individual differences on ESL learners’ segmental and suprasegmental production. Findings overwhelmingly demonstrate that the effect of IDs on pronunciation is undeniable. IDs predicted between 32% and 70% of the variance in the quantifiable pronunciation properties (70%, segmentals; 32%, pausing; 64%, speech rate; 66%, sentence stress; 48%, varied intonation; 34%, level intonation; and 32%, pitch), with the largest effect on segmental performance. While IDs accounted for large portions of variance, their relationships with the pronunciation properties were not always straightforward. Rather, there were complex and dynamic relationships between the learner IDs and pronunciation (Hansen Edwards, 2018; Moyer, 2014; Nagle, 2022). Some IDs generally tended to trend in one direction, such as with the effect of motivation, age, and English use, but others (PSTM, anxiety, and LOR) had differential effects on different aspects of pronunciation. Despite these complex relationships, there are clear implications for teaching. Table 1 below summarizes some research-informed pedagogical strategies which can be adopted in the ESL classroom to address different facets of the IDs covered in this study. These points are elaborated on in detail below.

Table 1
Pedagogical Strategies and L2 Pronunciation Individual Differences

| Individual Difference | Pedagogical Strategies |
|---------------------------------------|---|
| Motivation | <ul style="list-style-type: none"> • Recognize the complex nature of L2 motivation • Use ongoing methods to raise learners’ self-awareness of L2 motivation and track personal changes • Assist learners with setting realistic, personal goals • Recognize that a perceived “lack” of motivation may be due to other factors (e.g., higher levels of L2 anxiety or a range of other personal circumstances) • Train learners to take charge of their own learning by accessing resources for autonomous learning and self-evaluation |
| Age | <ul style="list-style-type: none"> • Prioritize communicative success and intelligibility with older learners in conjunction with their needs and goals • Recognize that older learners may still attain native-like performance; do not give up on accuracy • Adjust expectations of pronunciation accuracy relative to a learner’s aptitude, motivation, and goals |
| English Use | <ul style="list-style-type: none"> • Understand that L2 English use in the target community alone is not enough; English use in the classroom is still beneficial, and attention to corrective feedback provided by both teachers and peers is crucial • Incorporate both authentic and modified input into the classroom environment, based on the goal of each activity • Create scaffolded tasks for simulated authentic language use, drawing on Celce-Murcia et al.’s (2010) well-established, 5-phase communicative framework |
| Phonological Short-Term Memory (PSTM) | <ul style="list-style-type: none"> • Expect to see more immediate results of PSTM-involved activities (e.g., chunking, repetition, etc.) on sound and word structure • Draw on activities which prioritize repetition (e.g., elicited imitation tasks) for automaticity training • Do not expect to be able to help boost individual learners’ PSTM • Provide and activate background knowledge, as background knowledge is known to boost comprehension; however, be aware that students with more available short-term memory capacity are likely to benefit from this type of support disproportionately |

| | |
|---------------------------|---|
| | <ul style="list-style-type: none"> Engage learners in targeted fluency practice using simplified materials to improve fluency through boosting automaticity |
| Speaking Anxiety | <ul style="list-style-type: none"> Provide ample practice to engage in free speech Give learners options for back-up plans (e.g., filled pauses, word searching strategies (e.g., synonyms, circumlocution, etc.), lexical bundles, formulaic expressions) in moments of high anxiety Incorporate simulated authentic scenarios for L2 use in the classroom (e.g., making phone calls, interacting with NSs, receiving corrective feedback, etc.) to boost confidence Focus on the positive opportunities for use of the L2 to naturally turn debilitating anxiety into facilitative anxiety Highlight attainable goals (e.g., intelligible pronunciation) as opposed to native-like attainment to reduce pressure Do not forget to provide ample encouragement and positive feedback |
| Length of Residence (LOR) | <ul style="list-style-type: none"> Understand that LOR alone does not ensure target-like pronunciation Do not expect longest-residing learners to have the most target-like pronunciation Encourage short-, mid-, and long-term residents to interact more with authentic and modified interactive and non-interactive sources of language outside the classroom which seems to have a much more far-reaching impact on their pronunciation production |

Our findings demonstrate that motivation had an overwhelmingly positive effect on segmental and suprasegmental production, in line with expectations about this ID (Celce-Murcia et al., 2010; Hansen Edwards, 2018; Moyer, 2014). It should be noted that our instrument did not have a focus on obtaining a native-like accent in English; it therefore illustrated the strong potential of motivation on intelligible and comprehensible pronunciation. Learners within an ESL context are faced with the ever-present need to use English clearly. While this study used a composite score of the sub-constructs of motivation, it clearly demonstrates that the role of reported motivation and the clear images of the possible selves are what drives a learner to a greater investment in producing overall target-like speech (Moyer, 2014). With more reported motivation, learners had fewer segmental deviations, a faster speech rate, more varied intonation, less level intonation, and a wider pitch range. According to Brazil's (1997) framework, rising and falling intonation contribute to the context of interaction between the listener and the speaker, while level intonation is a neutral tone which exists outside of this context. A wider pitch range is more dynamic. These characteristics of suprasegmental performance (which are all features of intelligible, listener-inclusive speech, see Kang et al., 2022), in conjunction with more segmental accuracy and a faster speech rate, were influenced by more motivated behavior, perhaps demonstrating the extent to which ESL learners are motivated for communicative success with a variety of English-speaking interlocutors within an ESL context.

Therefore, first and foremost, targeting multi-faceted motivation in the classroom is likely to have the most far-reaching effect on pronunciation. Teachers should recognize that motivation comes from several sources (how learners imagine their pronunciation in the future, how others imagine their pronunciation, personal consequences for not obtaining a specific level of pronunciation, other-imposed consequences, etc.). Pronunciation teachers can use ongoing methods to raise self-awareness of L2 motivation and track learners' personal changes, knowing that learners may be influenced by their ideal L2 selves (*own/other*) and their ought-to L2 selves (*own/other*). These vivid visions of their L2 selves, accompanied with strong motivated behavior, allow learners to quickly reduce the gap between their current and futuristic selves. Learners who are more motivated are ultimately more invested. This investment can provide the impetus for change for a broad range of segmental and suprasegmental properties. Teachers can assist pronunciation learners with setting realistic, personal goals (Horwitz, 2017). At the same time, due to the dynamism of L2 motivation, goal-setting should be an iterative process throughout a given time frame, and teachers may wish to re-evaluate motivation on several occasions. It is also important to remember that students who appear to be unmotivated, may not actually be unmotivated, but

they may have higher levels of language anxiety (Horwitz, 2017) or a range of other personal circumstances. Training learners to understand and take charge of their own learning, including an ability to access resources for autonomous learning and self-evaluation may further increase motivation.

While motivation tended to have an overall positive effect, age was the opposite. While they tended to pause less frequently, older learners had more segmental deviations, a slower speech rate, more sentence stress, less varied intonation, more level intonation, and a narrower pitch range. The effects of age on phonetics and phonology have been well-documented in the research (see Flege, 1995; Moyer, 2014), especially since phonology is a skill which depends on neuro-muscular flexibility, motor-control, and auditory-perceptual neural networks (Scovel, 1988). Pronunciation teachers working with older language learners should prioritize successful communication, intelligibility, and comprehensibility (Levis, 2005), always in conjunction with learners' needs and goals. It is crucial to note, moreover, that not a single domain of second language learning has been identified to date in which at least some late starters are unable to attain native-like levels of performance (Abrahamsson & Hyltenstam, 2009). Therefore, teachers should not automatically give up on the goal of improving older learners' pronunciation of individual sounds; some learners will be perfectly capable of performing well in this domain. However, since aptitude plays an important role (Abrahamsson & Hyltenstam, 2008), teachers should adjust their expectations when it comes to mastering segmentals in accordance with individual learners' abilities.

L2 English use mostly had an overall positive effect on pronunciation properties (more segmental accuracy, faster speech rate, less pausing, and more varied intonation). These findings add additional benefits of L2 English use to Derwing et al.'s (2008) findings, which demonstrated that conversation and media consumption had a part in improved fluency for speakers of Slavic languages learning English. Moyer (2014) notes that L2 use is more likely to have an effect on the L2 when it is used actively, consistently, and meaningfully, especially when it is used in the home. In this study, L2 English use was also a predictor of some speech tendencies typical of NNSs (e.g., more sentence stress and narrower pitch range) indicating that L2 use in the target community alone is not enough, and pronunciation instruction can still be quite beneficial here. More specifically, attention to corrective feedback provided by both teachers and peers within the classroom environment is crucial for raising metalinguistic awareness. Further supporting high-quality language use within the classroom, large amounts of authentic input (especially through media) can be especially influential on fluency and automaticity due to exposure to connected speech patterns, idiomatic expressions, formulaic language, and large amounts of variation in speech patterns, known to have a positive effect on pronunciation (Mahdi & Mohsen, 2024). This input should be modified based on the goal of each activity. Additionally, teachers should incorporate both authentic interactive and non-interactive language use and practice into the classroom with scaffolded tasks allowing learners to focus on pronunciation features¹. Celce-Murcia et al.'s (2010) five-phase communicative framework for pronunciation instruction provides a scaffolded approach to pronunciation learning building up from controlled to more communicative tasks (i.e., description and analysis, listening discrimination, controlled practice, guided practice, communicative practice) and can be used to target a range of segmental and suprasegmental properties.

The effect of other IDs on pronunciation was not as straightforward as motivation and age, presenting complex relationships with pronunciation. On the one hand, the expectation that PSTM would lead to more segmental accuracy did emerge. The phonological loop was anticipated to assist learners in acquiring more permanent, stable representations of segmental patterns associated with the target language. The phonological loop temporarily stores phonological forms on each occasion when a learner is incorporating new sound-based information into their mental lexicon (Gathercole, 2007). In essence, the more efficient a learner's phonological short-term memory, the more potential they might have for retaining target-like phonological patterns in their L2. Another finding that emerged was that the PSTM capacity also seemed to assist learners with reflecting nativelike patterns of sentence stress and pitch range. That is, native English speakers use sentence stress sparingly per tone unit, and they also have a wider pitch range when compared to NNSs (Kang, 2010; Pickering, 2018). However, unlike previous work (Ahmadian, 2012; O'Brien et al., 2007; Saito et al., 2019), PSTM was related to less fluent speech which was also less dynamic, illustrated by more pausing, a slower speech rate, and less varied and more level intonation. From an exploratory standpoint, there appears to be a divide between the effects of PSTM on higher level processes (e.g., content and discourse organization) vs lower-level ones (word and sound processing) (Segalowitz & Hultstijn, 2005, p. 371). Learners with higher amounts of available PSTM may also be more likely to

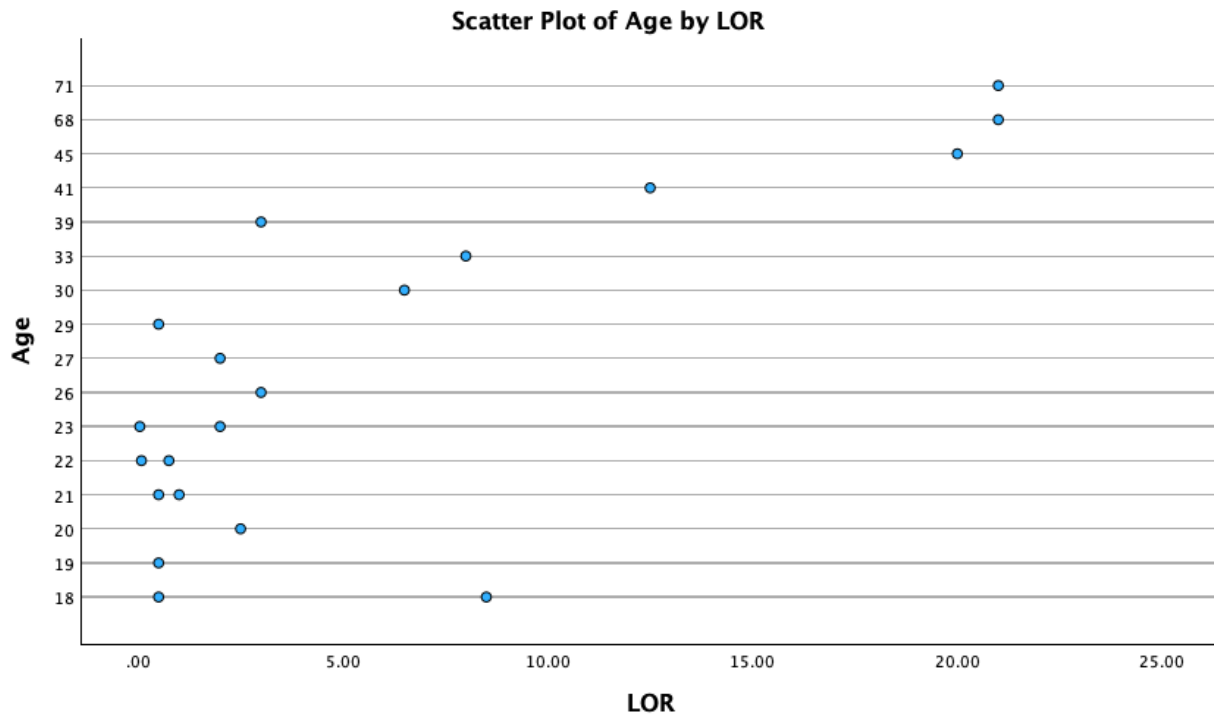
¹ Authentic listening activities are good for intensive listening. For extensive listening focused on comprehension as well as for fluency practice, learners need simplified input akin to graded readers in the visual modality (Nation, 2022).

engage in controlled L2 processing, which would be expected to have an effect on fluency (Segalowitz, 2010). Teachers are perhaps more likely to see immediate results of PSTM-involved activities (e.g., chunking, repetition, etc.) on sound and word structure. The more efficiently the system stores speech-based input, the more likely lower order processes are automatically retrievable. In fact, Segalowitz & Hulstijn (2005) highlight the role of repetition in automaticity training: repetition (in speech and text) reduces the cognitive load associated with information processing, freeing up resources to disperse more evenly between content and form. Using elicitation tasks with native and intelligible non-native models will provide this type of targeted practice in retaining and repeating speech patterns (see Kermad, 2025 for the facilitative effect of the elicitation task on pronunciation). There is unfortunately no evidence that targeted working memory training (PSTM is a component of working memory) has any effect on performance beyond what is practiced in the training and a small number of very similar tasks (Sala & Gobet, 2020; Soveri et al. 2017). Therefore, teachers cannot expect to be able to help boost individual learners' PSTM. Instead, teachers can provide and activate background knowledge, as background knowledge is known to boost comprehension; however, be aware that students with more available short-term memory capacity are likely to benefit from this type of support disproportionately (see Shin et al., 2019 on reading). Furthermore, targeted fluency practice, argued by Nation (2022) to be a crucial component of high-quality language teaching, can be used to improve fluency through boosting automaticity. Using simplified materials is recommended as it leaves more mental bandwidth for dealing with speech perception processes.

L2 speaking, in particular, is a highly complex skill due to its real-time balancing act of pronunciation, grammar, vocabulary, discourse, and pragmatics and can be more prone to situation-specific anxiety (Horwitz, 2017). In the present study, speaking anxiety specifically captured learners' anxiousness when making mistakes, repeating speech, receiving correction, and having to say difficult sounds, words, or phrases. These are realistic scenarios which arise regularly for learners in an ESL context, especially when interacting with native speakers of the L2 (see Woodrow, 2006). Unsurprisingly, anxiety was related to more segmental *inaccuracy* and *less* target-like temporal fluency; however, anxiety seemed to have a facilitative effect on prosody as learners used more target-like sentence stress, intonation, and pitch. This result may suggest that debilitating anxiety first manifests itself on sound and fluency. Gregersen et al. (2014) illustrated that anxiety can fluctuate up and down for both high- and low-anxiety learners. To address the debilitating effects of anxiety on segmental accuracy and fluency, some of Gregersen et al.'s (2014) suggestions can be used. Teachers can provide ample practice and strategies with free speech tasks for encouraging students to have a backup plan (e.g., using synonyms, circumlocution, etc.) in cases where they forget how to pronounce a word. Teachers can also help learners plan "escape routes," (Gregersen et al., 2014, p. 98) (e.g., filled pauses, word searching strategies, lexical bundles, formulaic expressions, etc.) to improve fluency. Incorporating simulated authentic scenarios for L2 use in the classroom (e.g., making phone calls, interacting with NSs, receiving corrective feedback, etc.) will likely provide more confidence for learners to perform in these situations outside of the classroom. Finally, focusing on the positive opportunities for use of the L2 might naturally help turn debilitating anxiety into facilitative anxiety. For example, teachers can help learners deal with anxiety by emphasizing the positives in each learner's competence and by reminding learners that language learning is a demanding process. Teachers may also wish to point out to learners that nativelike performance need not be a goal for second language pronunciation. In fact, intelligible (clear) and comprehensible (easy to understand) speech are much more attainable and should be the focus in the pronunciation classroom (Levis, 2005). Further, the average English speaker is much more likely to be appreciative of a learner's proficiency level, whatever it may be, than critical of it.

LOR, like anxiety, demonstrated a contrastive effect when comparing segmentals and temporal fluency with other suprasegmental features. LOR appeared to have a stronger effect on intonation and pitch than segmentals and fluency. LOR was related to more segmental inaccuracy, more pausing, and a slower speech rate. At the same time, longer LOR predicted the use of more target-like prosody, including more varied intonation patterns which were ultimately less level, along with a wider pitch range. Post-hoc analyses were performed to develop these findings, and a significant positive correlation was found between age and LOR (Pearson Correlation = .88, $p < .001$). The scatterplot below (see Figure 2) visually displays the relationship: the large majority of learners who had been living in the United States for less than a year were under 30 years old. As learners' LOR increased, their age also increased: those who had been living in the U.S. between 5-15 years were in their 30s and 40s (with the exception of one 18-year-old), and those who had been living in the U.S. over 20 years were 45 years and older.

Figure 2
Relationship between age and LOR



These post-hoc analyses demonstrate that there is a possible confound between LOR and age. In other words, because those who had resided in the U.S. longest were also the oldest learners, their segmental and fluency performance, could have also been reflected by age. This finding additionally supports that LOR, on its own, offers a limited contribution of a learner’s experience and is best considered alongside other variables which offer insights into learners’ language use (Hansen Edwards, 2018). Our findings demonstrated that L2 English use had a much more positive impact on one’s pronunciation than a simple quantification of how long one had lived in the country. This is quite meaningful for teachers working with short-, mid-, and long-term residents who can encourage learners to interact more with authentic, modified, interactive and non-interactive sources of language outside the classroom which seems to have a much more far-reaching impact on their pronunciation production. LOR alone does not ensure target-like pronunciation, and teachers should not expect their longest-residing learners to have the most target-like pronunciation.

In summary, considering the exploratory nature of this study, the above findings should be considered with respect to the study’s limitations. First, this study is a first to explore this broad a range of IDs on a comparably broad range of quantifiable pronunciation properties; therefore, results and patterns are preliminary and should be considered as such. It would be good practice in future research to explore the task-mediated effect of individual differences on pronunciation performance, as suggested in Kermad (2025), as the current findings are limited to a spontaneous speech task. Additionally, the effects of individual learner variability not captured by our measures (see the contribution of the random learner factor to our statistical models presented above) have two implications: (1) Incorporating other IDs would also be insightful in this line of work; for example, enjoyment and willingness to communicate can also be quite relevant to an ESL context as can aptitude; (2) The rich inter-learner variability inherent to SLA research often gets swept under the rug in the field’s tendency to emphasize the importance of quantitative findings and, in particular, measures of central tendency; closer qualitative scrutiny of individual learners, particularly outliers, would greatly enhance our understanding of language learning. Therefore, we recommend that quantitative analyses of measures of central tendency be complemented by careful examination of variability in the dataset as well as by qualitative follow-up examination and rich case studies of individual learners, particularly highly successful and highly unsuccessful ones, which is where we have the most to learn in SLA research. Finally, the current study made use of a cross-sectional design, and speech data were only

collected at a single time point. The field of SLA can benefit greatly by further research which traces L2 learners through time (e.g., Nagle, 2022; Saito et al., 2019), revealing the longer effect and directionality of IDs on pronunciation acquisition over time.

Conclusion

In closing, findings from this study corroborate the findings of prior research (e.g., Celce-Murcia et al., 2010; Moyer, 2014) suggesting that L2 pronunciation is a skill which is impacted by IDs from the whole learner. Therefore, educators cannot and should not expect uniform acquisition patterns in an ESL classroom, even for language learners of the same proficiency. On the contrary, the results of this study demonstrate that L2 pronunciation is a skill which is extraordinarily personal, complex, and dynamic. Awareness of the nature of this skill is the first step. Educators should actually *expect* differences from learner to learner within the classroom. Then, an educator's expectations within an ESL classroom can be continuously modified from class to class and learner to learner, knowing that some IDs are constantly in flux (e.g., motivation, anxiety, English use) while others (e.g., aptitude, age, LOR) are more fixed. For those IDs in flux, there are more strategies an educator can take to address them and instill change. On the other hand, for the IDs which are more fixed, awareness and context are key to framing these within the ESL classroom as well as for offering research informed strategies to address them in alternative ways.

Authors

Alyssa Kermad is an Associate Professor of Applied Linguistics and TESOL at California State Polytechnic University, Pomona. Her research areas span second language speaking and pronunciation, second language acquisition, speech assessment and methodology, and pragmatics and prosody.

Vedran Dronjic is Associate Professor of Applied Linguistics and TESL at Northern Arizona University. He holds a Ph.D. from the University of Toronto. His research relies on psycholinguistic and neurolinguistic tools to investigate language learning and processing, particularly as they relate to vocabulary, morphology, reading, and memory. He has a strong interest in popularizing research findings and translating them into recommendations for pedagogical practice.

Okim Kang is Professor of Applied Linguistics, Director of the Applied Linguistics Speech Lab, and Director of Global Communication Center at Northern Arizona University. Her research interests are speech production and perception, L2 pronunciation and intelligibility, L2 oral assessment and testing, automated scoring and speech recognition, World Englishes, and language attitudes.

Naoko Taguchi is Professor of Applied Linguistics at Northern Arizona University where she teaches courses in linguistics and TESOL. Her primary research interests include second language pragmatics, technology-enhanced language learning, and intercultural communication.

References

- Abrahamsson, N., & Hyltenstam, K. (2008). The robustness of aptitude effects in near-native second language acquisition. *Studies in Second Language Acquisition*, 30(4), 481-509.
- Abrahamsson, N., & Hyltenstam, K. (2009). Age of onset and nativelikeness in a second language: Listener perception versus linguistic scrutiny. *Language Learning*, 59(2), 249-306.
- Ahmadian, M. J. (2012). The relationship between working memory capacity and L2 Oral performance under task-based careful online planning condition. *TESOL Quarterly*, 46, 165-175.
- Baddeley, A. (2003). Working memory: Looking back and looking forward. *Nature Reviews Neuroscience*, 4, 829-839.
- Baddeley, A., Gathercole, S., & Papagno, C. (1998). The phonological loop as a language learning device. *Psychological Review*, 105, 158-173.
- Baker Smemoe, W., & Haslam, N. (2013). The effect of language learning aptitude, strategy use and learning context on L2 pronunciation learning. *Applied Linguistics* 34(4), 435-456.

- Baran-Łucarz, M. (2012). Individual learner differences and accuracy in foreign language pronunciation. In M. Pawlak (Ed.), *New perspectives on individual differences in language learning and teaching* (pp. 289–303). Springer.
- Brazil, D. (1997). *The communicative value of intonation in English*. Cambridge University Press.
- Celce-Murcia, M., Brinton, D. M., Goodwin, J. M., & Griner, B. (2010). *Teaching pronunciation: A course book and reference guide*. Cambridge University Press.
- Derwing, T. M., Munro, M. J., & Thomson, R. I. (2008). A longitudinal study of ESL learners' fluency and comprehensibility development. *Applied Linguistics*, 29(3), 359-380.
- Dörnyei, Z. (2005). *The psychology of the language learning: Individual differences in second language acquisition*. Lawrence Erlbaum Associates.
- Dörnyei, Z. (2006). Individual differences in second language acquisition. *AILA Review*, 19(1), 42-68.
- Dörnyei, Z. (2009). The L2 motivational self system. In Z. Dörnyei, & E. Ushioda (Eds.), *Motivation, language identity and the L2 self* (pp. 9-42). Multilingual Matters.
- Dörnyei, Z. (2013). Motivation in second language learning. In M. Celce-Murcia, D. M. Brinton, & M. A. Snow (Eds.), *Teaching English as a second or foreign language* (4th ed., pp. 518-531). National Geographic Learning/Cengage Learning.
- Dörnyei, Z., & Skehan, P. (2003). Individual differences in second language learning. In C. J. Doughty & M. H. Long (Eds.), *The handbook of second language acquisition* (pp. 589–630). Blackwell.
- Flege, J. E. (1995). Second language speech learning: Theory, findings, and problems. *Speech perception and linguistic experience: Issues in cross-language research*, 233-277.
- Gathercole, S. E. (2006). Nonword repetition and word learning: The nature of the relationship. *Applied Psycholinguistics*, 27, 513-543.
- Gathercole, S. E. (2007). Working memory and language. In M.G. Gaskell (Ed.), *The Oxford handbook of psycholinguistics* (pp.757-769), Oxford University Press.
- Gathercole, S. E., & Baddeley, A. D. (1989). Evaluation of the role of phonological STM in the development of vocabulary in children: A longitudinal study. *Journal of Memory and language*, 28, 200-213.
- Gathercole, S. E., Willis, C. S., Emslie, H., & Baddeley, A. D. (1992). Phonological memory and vocabulary development during the early school years: A longitudinal study. *Developmental Psychology*, 28, 887.
- Gregersen, T., MacIntyre, P. D., & Meza, M. D. (2014). The motion of emotion: Idiodynamic case studies of learners' foreign language anxiety. *The Modern Language Journal*, 98(2), 574-588.
- Hansen Edwards, J. G. (2018). Pronunciation and individual differences. In O. Kang, R. Thomson, & J. Murphy (Eds.), *The Routledge handbook of Contemporary English pronunciation* (pp. 385-393). Routledge.
- Horwitz, E. K. (2017). On the misreading of Horwitz, Horwitz, and Cope (1986) and the need to balance anxiety research and the experiences of anxious language learners. In C. Gkonou, M. Daubney & J.-M. Dewaele (Eds.), *New Insights into Language Anxiety: Theory, Research and Educational Implications* (pp. 31-50). Multilingual Matters.
- Isaacs, T. (2010). *Issues and arguments in the measurement of second language pronunciation*. (Unpublished doctoral dissertation). McGill University, Canada.
- Kang, O. (2010). Relative salience of suprasegmental features on judgments of L2 comprehensibility and accentedness. *System*, 38, 301-315.
- Kang, O., & Moran, M. (2014). Functional loads of pronunciation features in non-nativespeakers' oral assessment. *TESOL Quarterly*, 48, 176-187.
- Kang, O., Johnson, D., & Kermad, A. (2022). *Second Language Prosody and Computer Modeling*. Routledge.
- Kermad, A. (2025). Validity of task choice on global and specific pronunciation. *Research Methods in Applied Linguistics* (4)3. <https://doi.org/10.1016/j.rmal.2025.100277>
- Kessler, G. (2010). Fluency and anxiety in self-access speaking tasks: the influence of environment. *Computer Assisted Language Learning*, 23(4), 361-375.
- Kormos, J., & Dénes, M. (2004). Exploring measures and perceptions of fluency in the speech of second language learners. *System*, 32, 145-164.
- Levis, J. M. (2005). Changing contexts and shifting paradigms in pronunciation teaching. *TESOL Quarterly*, 39, 369-377.
- MacIntyre, P. D. (2007). Willingness to communicate in the second language: Understanding the decision to speak as a volitional process. *The Modern Language Journal*, 91(4), 564-576.

- MacIntyre, P. D. (2017). An overview of language anxiety research and trends in its development. In C. Gkonou, M. Daubney, & J.-M. Dewaele (Eds.), *New insights into language anxiety: Theory, research and educational implications* (pp. 11-30). Multilingual Matters.
- Mahdi, H. S., & Mohsen, M. A. (2024). Enhancing Pronunciation Learning through High Variability Phonetic Training: A Meta-Analysis. *Language Teaching Research Quarterly*, 40, 29-45.
- Moyer, A. (1999). Ultimate attainment in L2 phonology. *Studies in Second Language Acquisition*, 21, 81-108.
- Moyer, A. (2011). An investigation of experience in L2 phonology. Does quality matter more than quantity?. *Canadian Modern Language Review* (67), 191–216.
- Moyer, A. (2014). Exceptional outcomes in L2 phonology: The critical factors of learner engagement and self-regulation. *Applied Linguistics*, 35, 418-440.
- Nagle, C. (2022). Individual difference factors for second language pronunciation. In S. Li, P. Hiver, & M. Papi (Eds.), *The Routledge handbook of second language acquisition and individual differences* (pp. 269–281). Routledge.
- Nation, I. S. P. (2022). *Learning vocabulary in another language* (3rd ed.). Cambridge University Press.
- Norris, J. M., & Ortega, L. (2000). Effectiveness of L2 instruction: A research synthesis and quantitative meta-analysis. *Language Learning*, 50(3), 417-528.
- O'Brien, I., Segalowitz, N., Freed, B., & Collentine, J. (2007). Phonological memory predicts second language oral fluency gains in adults. *Studies in Second Language Acquisition*, 29, 557-581.
- Papi, M., & Hiver, P. (2022). Motivation. In S. Li, P. Hiver, & M. Papi (Eds.), *The Routledge handbook of second language acquisition and individual differences* (pp. 113–127). Routledge.
- Papi, M., Bondarenko, A, Mansouri, S., Feng, L., & Jiang, C. (2019). Rethinking L2 motivation research: The 2X2 model of L2 self-guides. *Studies in Second Language Acquisition*, 41(2), 337-361.
- Pickering, L. (2018). *Discourse intonation: A discourse-pragmatic approach to teaching the pronunciation of English*. University of Michigan Press.
- Purcell, E. T., & Suter, R. W. (1980). Predictors of pronunciation accuracy: A reexamination. *Language Learning*, 30(2), 271-287.
- Saito, K. (2017). Effects of sound, vocabulary and grammar learning aptitude on adult second language speech attainment in foreign language classrooms. *Language Learning*, 67, 665-693.
- Saito, K., & Hanzawa, K. (2016). Developing second language oral ability in foreign language classrooms: The role of the length and focus of instruction and individual differences. *Applied Psycholinguistics*, 37, 813-840.
- Saito, K., Dewaele, J.-M., Abe, M., & 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., Dewaele, J.-M., & Hanzawa, K. (2017). A longitudinal investigation of the relationship between motivation and late second language speech learning in classroom settings. *Language and Speech*, 60, 614-632.
- Saito, K., Suzukida, Y., & 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.
- Sala, G., & Gobet, F. (2020). Working memory training in typically developing children: A multilevel meta-analysis. *Psychonomic Bulletin & Review*, 27(3), 423-434.
- Sardegna, V. G., Lee, J., & Kusey, C. (2014). Development and validation of the learner attitudes and motivations for pronunciation (LAMP) inventory. *System*, 47, 162-175.
- Scovel, T. (1988). *A time to speak: A psycholinguistic inquiry into the critical period for human speech*. Newbury House.
- Segalowitz, N. (2010). *Cognitive bases of second language fluency*. Routledge.
- Segalowitz, N., & Hulstijn, J. (2005) Automaticity in bilingualism and second language learning. In J. F. Kroll & A. M. B. De Groot (Eds), *Handbook of bilingualism: Psycholinguistic approaches* (pp. 371-388). Oxford University Press.
- Shin, J., Dronjic, V., & Park, B. (2019). The interplay between working memory and background knowledge in L2 reading comprehension. *TESOL quarterly*, 53(2), 320-347
- Smit, U., & Dalton, C. (2000). Motivational patterns in advanced EFL pronunciation learners. *IRAL-International Review of Applied Linguistics in Language Teaching*, 38(3-4), 229-246.
- Soveri, A., Antfolk, J., Karlsson, L., Salo, B., & Laine, M. (2017). Working memory training revisited: A multi-level meta-analysis of n-back training studies. *Psychonomic Bulletin Review*, 24(4), 1077-1096.

- Suter, R. W. (1976). Predictors of pronunciation accuracy in second language learning. *Language Learning*, 26(2), 233-253.
- Taguchi, N., Xiao, F., & Li, S. (2016). Effects of intercultural competence and social contact on speech act production in a Chinese study abroad context. *The Modern Language Journal*, 100(4), 775-796.
- Teimouri, Y. (2017). L2 selves, emotions, and motivated behaviors. *Studies in Second Language Acquisition*, 39(4), 681-709.
- Thomson, R., & T., Derwing. 2015. The effectiveness of L2 pronunciation instruction: A narrative review, *Applied Linguistics*, 36, 326-344.
- Trofimovich, P., & Baker, W. (2006). Learning second language suprasegmentals: Effect of L2 experience on prosody and fluency characteristics of L2 speech. *Studies in Second Language Acquisition*, 28, 1-30.
- Woodrow, L. (2006). Anxiety and speaking English as a second language. *RELC journal*, 37(3), 308-328.

Appendix(ces)

Appendix A: Non-Word Repetition Test of PSTM

Forward Test

| | |
|-----------------|---|
| Level 2 | zæbə / fulou |
| | fɛvə / tɪzɛɪ |
| Level 3 | wasə / zɪfɛɪ / kɪbə |
| | mabɛɪ / dæzə / funaɪ |
| Level 4 | ɹɪfɛɪ / ruvə / ʃɛdɛɪ / wɪfə |
| | gəkə / tɪbɛɪ / ɹɪtə / lupou |
| Level 5 | ʃɪdɛɪ / lævə / huseɪ / gumə / daʊzɔɪ |
| | vɛgə / ɹaveɪ / wɪtə / tʃɪmɔɪ / manə |
| Level 6 | wɛɪzə / flɔɔɪ / sɪtʃə / fæbaɪ / gatə / ɹɪzɛɪ |
| | wɛfə / fɪkɔɪ / buɹə / ʃɪnɔɪ / hɛkə / ɹædaɪ |
| Level 7 | lɹə / hutou / mɛɹə / rɪkɔɪ / dɛlə / fɪbɛɪ / hæɹə |
| | laɪtɔɪ / vætə / tɪfɔɪ / sɪkə / nɛkɛɪ / jælə / dɪkɔɪ |
| Level 8 | vɛtə / wadou / rævə / rulou / ɹɔzə / zalɔɪ / vugə / tʃɪsɔɪ |
| | tɪsə / vɛkɔɪ / mævə / zɪkɔɪ / bɪvə / ɹaʊsɔɪ / vumə / mɛbuɪ |
| Level 9 | fɔɪ / wæɹə / mufoʊ / kɪnə / bɪnaɪ / tæfə / ɹouʃɔɪ / zatə / mɪɹɔɪ |
| | fouʃə / louɹɔɪ / ɹɔɪzə / jɪbɔɪ / gləkə / zɛbɛɪ / rakə / fɪkɔɪ / hɪɹə |
| Level 10 | wɛʃɔɪ / nɛfə / dʒubou / fɪsə / zɪtɔɪ / lɪkə / vælɔɪ / ɹouʃə / flɔɪ / tɪdə |
| | mæsɔɪ / sæmə / balɔɪ / kuza / hakɔɪ / rɪzə / louɹou / gɛvə / fɪzɔɪ / bæfə |

Diphthong symbol:

[aɪ]: I, bite, Stein, aisle

[aʊ]: bout, brown, doubt

[ɔɪ]: boy, rejoice, annoy

[oʊ]: oh, boat, beau, grow

[ɛɪ]: bait, reign, great

Backward Test

| | |
|-----------------|--|
| Level 2 | ριζει / κιθα |
| | τjιςκι / νεγα |
| Level 3 | zalaυ / buρα / laitou |
| | ruvα / falκι / wεfα |
| Level 4 | fiβει / vαtα / mufou / wαpα |
| | τjιmκι / hiρα / wεjou / kuza / |
| Level 5 | mαcaυ / γλka / loupci / γεvα / wadou |
| | pcιzα / ruloυ / vεtα / tifι / rλzα |
| Level 6 | mebou / sitjα / fikau / rakα / rλkau / gatα |
| | mιpci / vugα / balau / caemα / dikau / baεfα |
| Level 7 | fulou / pa3α / pouτjκι / foυbα / flbou / wifα / fizou |
| | weizα / fikou / gαkα / zitau / witα / dauzκι / zaβα |
| Level 8 | d3ubou / jαelα / pαdai / λpα / binau / tidα / tijει / mepα |
| | pavei / fevα / hutou / delα / funai / dαzα / zeβει / wasα |
| Level 9 | husei / laevα / zλkκι / hαpα / vekκι / manα / raucci / rαevα / jibκι |
| | pιtα / zλfei / cikα / mabeι / kιnα / lupou / tisα / jεdei / hekα |
| Level 10 | loupcou / gumα / pifei / tαεfα / tibeι / fisα / jλdei / mαevα / faβai / binα |
| | flnκι / vatα / jιnκι / vumα / hakou / nefα / nekei / pouτa / vaelau / like |

Diphthong symbol:

[ai]: I, bite, Stein, aisle

[au]: bout, brown, doubt

[oi]: boy, rejoice, annoy

[ou]: oh, boat, beau, grow

[ei]: bait, reign, great

Appendix B: Ideal Selves Motivation Survey

In the column “motivation,” fill in the circles according to how much you agree or disagree with the statement.

| | | | | | |
|-------------------|----------|-------------------|----------------|-------|----------------|
| 1 | 2 | 3 | 4 | 5 | 6 |
| Strongly Disagree | Disagree | Slightly Disagree | Slightly Agree | Agree | Strongly Agree |

| Statement | Motivation | | | | | |
|---|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|
| Ideal L2 Self/Own | | | | | | |
| 1. <i>I can imagine a day when I speak English that is very clear to listeners.</i> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| | 1 | 2 | 3 | 4 | 5 | 6 |
| 2. <i>I can imagine a day when I speak English that is very easy for listeners to understand.</i> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| | 1 | 2 | 3 | 4 | 5 | 6 |
| 3. <i>I can imagine a day when I speak English proficiently with international friends/colleagues.</i> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| | 1 | 2 | 3 | 4 | 5 | 6 |
| 4. <i>I can imagine a day when I have successful English communication with people from all around the world.</i> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| | 1 | 2 | 3 | 4 | 5 | 6 |
| Ideal L2 Self/Other | | | | | | |
| 5. <i>My family hopes that one day that I will speak English clearly.</i> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| | 1 | 2 | 3 | 4 | 5 | 6 |
| 6. <i>My family hopes that one day my English will be extremely easy to understand.</i> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| | 1 | 2 | 3 | 4 | 5 | 6 |
| 7. <i>It is my parents’ hope that one day I will speak English proficiently.</i> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| | 1 | 2 | 3 | 4 | 5 | 6 |
| 8. <i>The people who are important to me hope that one day I will master English pronunciation.</i> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| | 1 | 2 | 3 | 4 | 5 | 6 |
| Ought to Self/Own | | | | | | |
| 9. <i>If I don’t improve my English pronunciation, it will have a negative impact on my future.</i> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| | 1 | 2 | 3 | 4 | 5 | 6 |
| 10. <i>If I don’t improve my English pronunciation, I will fail in my future career.</i> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| | 1 | 2 | 3 | 4 | 5 | 6 |
| 11. <i>If I don’t improve my English pronunciation, I will fail in my social life.</i> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| | 1 | 2 | 3 | 4 | 5 | 6 |
| Ought-to Self/Other | | | | | | |
| 12. <i>If I don’t improve my English pronunciation, I will disappoint my parents/teachers.</i> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| | 1 | 2 | 3 | 4 | 5 | 6 |

| | | | | | | |
|--|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|
| <i>13. My family puts a lot of pressure on me to improve my English pronunciation.</i> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| | 1 | 2 | 3 | 4 | 5 | 6 |
| Motivational Intensity | | | | | | |
| <i>14. I work hard at studying English pronunciation.</i> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| | 1 | 2 | 3 | 4 | 5 | 6 |
| <i>15. I spend a lot of time studying English pronunciation.</i> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| | 1 | 2 | 3 | 4 | 5 | 6 |
| <i>16. I put a lot of effort in studying English pronunciation.</i> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| | 1 | 2 | 3 | 4 | 5 | 6 |
| <i>17. I constantly think about my English pronunciation activities.</i> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| | 1 | 2 | 3 | 4 | 5 | 6 |
| <i>18. Studying English pronunciation is very important to me these days.</i> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| | 1 | 2 | 3 | 4 | 5 | 6 |

Appendix C: Anxiety Survey

In the column “anxiety,” fill in the circles according to how anxious you feel when you speak English in the following situations.

| | | | | |
|--------------------|------------------|--------------------|--------------|-------------------|
| 1 | 2 | 3 | 4 | 5 |
| Not at all anxious | Slightly anxious | Moderately anxious | Very anxious | Extremely anxious |

| Situation | Anxiety | | | | |
|---|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|
| 1. Communicating with a native English speaker whom I do not know. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| | 1 | 2 | 3 | 4 | 5 |
| 2. Making mistakes when I speak English. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| | 1 | 2 | 3 | 4 | 5 |
| 3. Participating in a group discussion in English. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| | 1 | 2 | 3 | 4 | 5 |
| 4. Repeating speech that was not clear to listeners. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| | 1 | 2 | 3 | 4 | 5 |
| 5. Receiving correction on my English speech. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| | 1 | 2 | 3 | 4 | 5 |
| 6. Speaking on the phone in English. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| | 1 | 2 | 3 | 4 | 5 |
| 7. Speaking English sounds, words, or phrases which are difficult for me. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| | 1 | 2 | 3 | 4 | 5 |

Appendix D: Second Language English Use Survey

| | |
|--|----------------|
| <p>Interactive Contact: In the column “Hours per week,” please write the amount of time you spend using English in the following activities, according to your own experience in the United States. Please think of your normal week. If the activity does not apply to you, please write “0” hours per week.</p> <p>Please only write one number in the column.</p> | |
| | Hours per week |
| 1. I spend approximately _____ hours per week communicating with native speaker friends in English. | |
| 2. I spend approximately _____ hours per week communicating with non-native speaker friends in English. | |
| 3. I spend approximately _____ hours per week communicating with boyfriend /girlfriend / husband / wife in English. | |
| 4. I spend approximately _____ hours per week communicating with professionals (at banks, post offices, restaurants) in English. | |
| 5. I spend approximately _____ hours per week communicating with roommate(s) in English. | |
| 6. I spend approximately _____ hours per week communicating with people during online gaming (PlayStation, Xbox, etc.) in English. | |
| 7. I spend approximately _____ hours per week communicating with customers/patients/clients in English. | |
| 8. I spend approximately _____ hours per week communicating with student and/or work colleagues in English. | |

| | |
|---|----------------|
| <p>Non-Interactive Contact: In the column “Hours per week,” please write the amount of time you spend with English in the following activities, according to your own experience in the United States. Please think of your normal week. If the activity does not apply to you, please write “0” hours per week.</p> <p>Please only write one number in the column.</p> | |
| | Hours per week |
| 1. I spend approximately _____ hours per week watching TV in English. | |
| 2. I spend approximately _____ hours per week watching movies in English. | |
| 3. I spend approximately _____ hours per week watching videos (YouTube, DailyMotion, Facebook, etc.) in English. | |
| 4. I spend approximately _____ hours per week listening to music in English. | |
| 5. I spend approximately _____ hours per week listening to game content in English. | |

Appendix E:

Table 1
Summary of Segmental and Suprasegmental Analyses

| Feature | Description |
|---|---|
| 1. Segmentals | Auditory instances in which vowels or consonants deviated from what would be expected in Standard American English (e.g. consonant deletion/substitution/addition; vowel deletion/substitution/addition; or syllable addition/deletion). Deviations were normalized by dividing the number of segmental errors by the total number of syllables in the speech sample. |
| 2. Speech Rate Articulation rate | Number of syllables produced in the speech sample divided by the total length of the speech sampled (excluding pause time) and normalized to one minute. |
| 3. Pausing Total number of pauses | The total number of pauses in the speech sample. |
| 4. Sentence Stress Pace | Total number of prominent (stressed) syllables divided by the total number of runs in the speech sample. |
| 5. Varied Intonation | |
| Rising Tones | The total number of rising or fall-rising tones in the speech sample. |
| Falling Tones | The total number of falling or rise-falling tones in the speech sample. |
| 6. Level Intonation | |
| Level Tones | The total number of level tones in the speech sample. |
| 7. Pitch Range | The difference between the F ₀ minima and maxima appearing on prominent syllables per speech sample. |

Appendix F:

Table 2
Summary Estimates of Main Effects of Individual Differences on Pronunciation Properties

| Parameter | <i>df</i> | <i>Segmentals</i> | | <i>Pausing</i> | | <i>Rate</i> | | <i>Prominence</i> | | <i>Varied Intonation</i> | | <i>Level Intonation</i> | | <i>Pitch Range</i> | |
|------------|-----------|-------------------|----------|----------------|----------|-------------|----------|-------------------|----------|--------------------------|----------|-------------------------|----------|--------------------|----------|
| | | <i>b</i> | <i>p</i> | <i>b</i> | <i>p</i> | <i>b</i> | <i>p</i> | <i>b</i> | <i>p</i> | <i>b</i> | <i>p</i> | <i>b</i> | <i>p</i> | <i>b</i> | <i>v</i> |
| Intercept | 1193 | -1.28 | .000 | -2.92 | .000 | .78 | .000 | 1.22 | .000 | -.13 | .044 | .34 | .101 | -3.2 | .000 |
| Age | 1193 | .01 | .001 | -.01 | .008 | -.01 | .012 | .03 | .000 | -.02 | .000 | .05 | .000 | -.04 | .000 |
| Motivation | 1193 | -.01 | .000 | .02 | .000 | .01 | .000 | .00 | .309 | .004 | .000 | -.01 | .000 | .03 | .000 |
| PSTM | 1193 | -.02 | .000 | .05 | .000 | -.02 | .000 | -.10 | .000 | -.004 | .002 | .03 | .000 | .02 | .000 |
| Anxiety | 1193 | .12 | .000 | .04 | .000 | -.05 | .000 | -.03 | .000 | .01 | .000 | -.02 | .000 | .08 | .000 |
| L2 Use | 1193 | .01 | .000 | -.01 | .000 | .01 | .000 | .01 | .000 | .0001 | .004 | .00 | .775 | -.01 | .000 |
| LOR | 1193 | .02 | .018 | .10 | .000 | -.04 | .000 | -.15 | .000 | .071 | .000 | -.16 | .000 | .14 | .000 |