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 experiencea |  |  |  |
| 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 experienceb |  |  |  |
| 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)

Table2 *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)

Table3 *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)

Table5 *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),

Table6 *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)

Table7 *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 1 2 3 4 5 6 7 8 9 Easy to understand understand

* Nativelikeness

Heavily accented 1 2 3 4 5 6 7 8 9 Not accented at all 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 | 𝜒2 | *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 | 𝜒2 | *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.