



Can lower inhibitory control lead to better second language phonological contrast learning?*

Jieun Lee** · Hanyong Park

(University of Kansas · University of Wisconsin-Milwaukee)

Lee, Jieun and Hanyong Park. 2025. Can lower inhibitory control lead to better second language phonological contrast learning? *Linguistic Research* 42(Special Edition): 149-178. Inhibitory control, a domain-general cognitive function involved in selective attention, has been linked to individual differences in second language (L2) learning. However, previous research on its role in L2 phonological learning has yielded mixed findings. To clarify this relationship, the present study investigates how individual differences in inhibitory control relate to L2 speech learning, targeting intermediate Korean learners of English. Twenty-four participants completed five days of computer-based auditory training on two English vowel contrasts (/ɪ/–/i/ and /ɔ/–/u/). Inhibitory control was measured using the Stroop task. Participants completed pre- and post-tests with both trained and novel talkers, as well as everyday identification (ID) tests during the training phase. Results showed that learners with lower inhibitory control (i.e., higher Stroop scores) demonstrated greater improvement over training and increased use of spectral cues—primary acoustic dimension for the target contrasts—when identifying English vowels. These learners also generalized their learning on how to use spectral cues to stimuli from a novel talker. This study suggests that lower (reduced) inhibitory control, possibly allowing for a broader focus of attention, may facilitate a more successful shift from using duration to spectral cues. This is particularly relevant in contexts like Korean EFL education, where learners are often instructed to rely on durational differences between the target English vowels. (University of Kansas · University of Wisconsin-Milwaukee)

Keywords inhibitory control, individual differences, second language training, English vowel contrasts

* We wish to thank two anonymous reviewers for their valuable comments and suggestions.

** First & Corresponding Author

1. Introduction

When learning phonological contrasts in a second language (L2), learners often achieve varying levels of success. Some learners reach a high level of proficiency, while others struggle to make significant progress. This raises the question: What causes these differences in L2 learning? Understanding the sources of variability is crucial for predicting and designing effective L2 training paradigms to enhance learning outcomes. Previous research has identified several contributing factors, including individual differences in cognitive abilities (Darcy et al. 2015), native language (L1) processing (Lengeris 2009), and cue-weighting strategies (Kim et al. 2018). This study focuses on the role of individual differences in inhibitory control ability among L2 learners in learning novel L2 phonological contrasts.

2. Background

2.1 Inhibitory control

Inhibitory control is a key component of executive function, involved in selective attention, self-regulation, and conscious processing, alongside other executive processes such as working memory and cognitive flexibility (Miyake et al. 2000). Inhibitory control is commonly associated with the ability to suppress task-irrelevant or competing information (e.g., Houdé et al. 2000). It is typically assessed using psychological tasks such as the Stroop task or the Flanker task (e.g., Bender et al. 2016).

In bilingual research, inhibitory control is widely recognized as an essential function, and it has been gaining attention as a cognitive factor underlying individual differences among second language (L2) learners (e.g., Gollan et al. 2011; Linck et al. 2012; Lev-Ari and Peperkamp 2013). Inhibitory control is often viewed as the mechanism that enables bilinguals to select the relevant language over the irrelevant one, which is not currently in use. This ability is critical for bilinguals as they switch between languages and manage interference from the non-target language (Green 1998). Thus, strong inhibitory ability enables bilinguals to effectively switch their attention from one language to another and resist attending to the irrelevant language.

2.2 Individual differences in inhibitory control and L2 phonological contrast learning

Recent research in L2 acquisition suggests that L2 learners with strong inhibitory control can reduce native language (L1) interference, thereby enhancing their processing of L2 acoustic-phonetic information and leading to more accurate L2 phonological representations in both perception and production. For example, Darcy, Mora, and Daidone (2016) investigated the role of inhibitory control in the phonological processing of L2 learners. They measured the inhibition control ability of adult L2 learners using the retrieval-induced inhibition (RI) task (Lev-Ari and Peperkamp 2013). Correlation analyses between learners' phonological performance on two tasks - an ABX task for vowel perception and a delayed sentence repetition task for consonant production - and their inhibition scores showed that learners with higher inhibition scores exhibited lower error rates in both tasks. The authors interpreted these results as suggesting that learners with higher (i.e., stronger) inhibitory ability may have used this ability to support their learning of L2 segmental categories.

However, the role of inhibitory control in the learning process, particularly how this domain-general cognitive mechanism influences L2 learning at different stages, remains unclear, and previous studies have shown mixed results. For example, Huensch (2024) conducted a replication study of Darcy et al. (2016) but did not find a clear, strong, or consistent relationship between inhibitory control and L2 perception and production skills. As one possibility of no, or only a weak, relationship between inhibition and L2 phonological skills, Huensch (2024: 1408) suggested that inhibitory control might play a more critical role in acquisition stages or perception category formation and the relationship might not be strongly demonstrated between inhibitory control and more end-state performance of L2 phonological skills. Thus, further research is needed to examine whether individual differences in inhibitory control significantly affect the trajectory of L2 learning. For instance, this could be investigated through L2 training studies with less proficient learners that assess learning outcomes after each training session, or through pre- and post-test design L2 learning studies.

Contrary to the view that strong inhibitory control brings benefits to bilinguals

and L2 learners, it may not always be advantageous. In certain situations, lower (i.e., reduced) inhibitory control can be beneficial, particularly when previously less relevant or inhibited information becomes useful for a task, or when attention is required to multiple sources of information. For instance, in the context of speech processing, Kim et al. (2020) reported a potential advantage of reduced inhibitory control in speech adaptation. They investigated how native Canadian English listeners adapted to unfamiliar speech patterns involving an English vowel contrast. Listeners were presented with auditory stimuli of the English vowel contrast /ɛ/-/æ/ that had ambiguous primary spectral cues but informative secondary duration cues. For a successful adaptation, the listeners had to increase their reliance on a less relevant secondary acoustic cue (i.e., duration) to distinguish English /ɛ/ and /æ/ vowels. The study found a negative correlation between individual differences in inhibitory control, measured by the Stroop task (Stroop 1935), and the degree of adaptation. English listeners with reduced inhibitory control showed better adaptation to unfamiliar L1 vowel patterns by increasing their reliance on the secondary cue. Kim et al. (2020) proposed that reduced inhibitory control may lead to a broader focus of attention to multiple acoustic cues, thereby resulting in better speech adaptation when active use of a less informative acoustic cue is necessary.

3. Current study

Mixed findings from previous research on the role of inhibitory control in L2 phonological skills have resulted in no clear consensus. Moreover, previous studies were mainly cross-sectional studies with proficient L2 learners. For instance, the learners' self-report ratings (1-5) in Darcy et al. (2016) and Huensch (2024) were about 4 (i.e., well) regarding their own ability of speaking, understanding, reading, and writing in L2. This raised important questions about individual differences in inhibitory control and their roles in the course of L2 learning. By conducting a study that examines the learning trajectories of less proficient L2 learners over time, we may gain a better understanding of the role of inhibitory control in L2 phonological contrast learning.

Therefore, in the current study, low- to intermediate-level Korean learners of English received five days of computer-based auditory training to learn L2 phonological

contrasts, specifically the English /ɪ/-/i/ and /ʊ/-/u/ vowel contrasts. Learners' inhibitory control ability was measured using the Stroop task (Stroop 1935), and their learning trajectories were measured after each training session.

In addition to measuring learners' overall improvement before and after training, we examined how individual differences in inhibitory control were related to changes in cue-weighting strategies for identifying the target English vowel contrasts by comparing pre- and post-test results. Phonological contrasts are signaled by multiple acoustic dimensions, each contributing differently to perceptual distinctions. The target English vowel contrasts, for instance, are cued by various acoustic features, including spectral cues, duration, fundamental frequency (f0), and intensity (Hillenbrand et al. 1995; Hillenbrand et al. 2000). These cues, however, do not contribute equally to distinguish a contrast. Some acoustic cues contribute more reliably to signaling phonological contrasts (primary cues), while other cues are more variable and play a less critical role (secondary cues) (e.g., Francis et al. 2000; Francis and Nusbaum 2002; Holt and Lotto 2006; Schertz et al. 2015). English listeners, for example, rely more heavily on spectral than duration cues in distinguishing English tense from lax vowels (e.g., Francis et al. 2008). For successful L2 phonological contrast learning, learners need to learn how to weight these cues in a nativelike way. Accordingly, we measured how learners' use of primary (spectral) and secondary (duration) cues changed throughout training and whether these changes were associated with their inhibitory control ability.

The current study aims to address two key research questions regarding individual differences in inhibitory control and L2 phonological contrast learning:

RQ 1. How are individual differences in inhibitory control ability related to L2 learners' success in acquiring the target L2 phonological contrasts?

RQ 2. How are individual differences in inhibitory control ability related to changes in L2 learners' reliance on primary (spectral) and secondary (duration) acoustic cues in identifying the target L2 contrasts before and after training?

3.1 Target English vowel contrasts

When comparing the English and Korean vowel systems, the most notable difference lies in the distinction between English tense vowels and their lax counterparts. English

has two high front vowels, /i/ and /ɪ/, and two back vowels, /ɔ/ and /u/ (Yang 1996). Since Korean lacks tense-lax vowel contrasts, the English /ɪ/-/i/ and /ɔ/-/u/ contrasts are considered particularly challenging for Korean learners of English to acquire (Han 2001). Lee and Cho (2018) demonstrated the potential confusion Korean learners of English may experience when distinguishing these English vowel contrasts by examining how they map L2 English vowels onto Korean vowel categories. In their study, Korean listeners were asked to label English vowels in an auditory stimulus using Korean vowel orthography. The results showed that the English /ɪ/-/i/ and /ɔ/-/u/ contrasts were consistently mapped onto single Korean vowel categories: /i/ (‘ㅣ’) and /u/ (‘ㅜ’), respectively. These findings suggest that distinguishing these English vowel contrasts poses a significant challenge for Korean listeners.

Regarding the use of multiple acoustic cues to identify English vowels, Korean learners of English often employ cue-weighting strategies that differ from those of native English listeners. English vowel contrasts /ɪ/-/i/ and /ɔ/-/u/ are primarily signaled by spectral cues and secondarily by duration cues, with /i/ and /u/ generally having longer durations than their /ɪ/ and /ɔ/ counterparts (Flege et al. 1997; Kondaurova and Francis 2008). Lee (2008) further suggested that English listeners distinguish these contrasts solely based on spectral cues, regardless of the durational difference between auditory stimuli.

However, unlike native speakers of English, Korean learners of English often fail to use spectral cues as the primary acoustic dimension (Flege et al. 1997; Kim et al. 2018). Several L2 studies have shown that Korean learners of English tend to rely more heavily on duration than on spectral cues when identifying English vowel contrasts. As pointed out in Lee (2009), this may be partly attributable to duration-focused instruction in English as a Foreign Language (EFL) contexts, as well as transcription practices for English tense-lax vowels that emphasize durational differences (e.g., heed transcribed as [hi:d] whereas hid as [hid]). In addition, the psychoacoustic saliency of duration cues may also contribute to the dominant reliance on duration. According to the *Desensitization Hypothesis*, listeners of languages that do not use certain areas of the vowel space become “desensitized” to variations in formants (i.e., spectral cues) within those regions and, therefore, use psychologically more salient temporal cues (i.e., duration) instead (Bohn and Flege 1990). Thus, the primary use of duration cues in distinguishing English tense-lax vowel contrasts may stem from Korean listeners’ reduced sensitivity to formant variations in the regions

of the vowel space occupied by these English vowel contrasts.

The discrepancies between the English and Korean vowel systems, as well as the resulting perceptual challenges, make the training of English tense-lax vowel contrasts an interesting case for investigating the role of inhibitory control ability in L2 learning. This is because successfully acquiring these contrasts requires learners to selectively attend to the most relevant acoustic cues while resisting and suppressing their attention to less-relevant cues. This process may involve domain-general cognitive mechanisms, such as inhibitory control. Specifically, considering that Korean learners often rely on duration rather than spectral cues, shifting their attention from duration to spectral cues is required to correctly identify the English vowel contrast in a nativelike way.

4. Methods

4.1 Participants

Twenty-three native speakers of Korean participated in the experiment, completing all training sessions and tests (ten female, thirteen male; mean age, 22.4 years; range, 19–27 years). Participants were either undergraduate students or recent graduates from a university in Korea¹. All participants graduated from elementary, middle, and high schools in South Korea and had received formal English education from the age of 10 (i.e., third grade in elementary school) until the age of 18 (i.e., senior in high school). The average length of English education was 14.3 years. None of the participants had spent time living or studying abroad in an English-speaking country at the time of the experiment. This criterion was included to minimize potential variability in L2 learning experiences, particularly regarding exposure to different types of formal pronunciation instruction.

On the first day of the experiment, participants completed the LexTALE (Lexical Test for Advanced Learners of English) (Lemhöfer and Broersma 2012) to test their English proficiency. This test consists of a simple, un-speeded visual lexical decision task. During this test, participants were asked to decide whether the presented word

¹ All participants were paid for their participation in the experiment under a protocol approved by the Institutional Review Board (IRB) for the protection of human subjects at the University of Wisconsin-Milwaukee.

is an existing English word or not by clicking a “Yes” or “No” button on a computer screen. There were 60 trials, which took about 4-5 minutes to complete. The LexTALE test is a reliable measure of general English proficiency (see Lemhöfer and Broersma 2012, for a large-scale validation study with Dutch and Korean advanced English learners). Participants in the current study had a slightly lower average LexTALE score ($M = 59.5$, $SD = 8.31$, range = 49-85) compared to the average reported in Lemhöfer and Broersma (2012) ($M = 65.3$, $SD = 10.3$, range = 46-89). Thus, all participants in this study can be considered as low- to intermediate-level learners of English.

4.2 Stimuli

4.2.1 Stimul for the Stroop task

Eight Korean words were used in the Stroop task: the names of four colors - red, blue, yellow, and green - and four other words unrelated to colors (each word has a meaning of either *drill*, *lounge*, *smell*, or *uniform*). These words were arranged to create congruent, incongruent, and neutral conditions. The congruent condition is that the color word written on the computer screen is matched with the color of the ink (e.g., ‘빨강(RED)’ written in red ink). Each color name was presented six times (4 color names $\times$ 6 repetitions = 24 trials). The incongruent condition was created by presenting each color word printed in a color that does not match the name of the color (e.g., ‘빨강(RED)’ written in green ink). Each color name appeared with unmatched three other ink colors with two repetitions (4 color names $\times$ 3 unmatched colors $\times$ 2 repetitions = 24 trials). Finally, the neutral condition was presented with four non-color-related words in four ink colors, each repeated six times (e.g., 4 non-color words $\times$ 6 repetitions = 24 trials).

4.2.2 Stimul for English vowel contrast training

The stimuli for training English /I/-/i/ and /ʊ/-/u/ vowel contrasts consisted of two minimal pairs of English words differing in their vowels: *hid* vs. *heed* and *hood* vs. *who’d*. Four male speakers of American English recorded these words within the carrier sentence, ‘I said ____ again.’ The target words were extracted from the sentence and

manipulated along two acoustic dimensions: spectral and duration. First, we used TANDEM-STRAIGHT (Kawahara et al. 2009) to generate five-step continua for the spectral dimension. The first and second formant (F1, F2) values in Hz for each *hid-heed* and *hood-who'd* continuum from one of the four talkers are provided in Table 1. And then, each of the five spectral steps along the vowel spectral continuum was manipulated using the “To manipulation...” function in Praat (Boersma 2001) to generate five-step continua for the duration dimension. The duration continuum spanned from 140 ms (step 1) to 300 ms (step 5), with each step increasing 40 ms. These acoustic manipulation processes created five-step continua varying in spectral and duration dimensions, resulting in 25 training stimuli for each *hid-heed* and *hood-who'd* continuum for each talker.

Table 1. F1 & F2 values in Hz for the /ɪ/-/i/ (*hid-heed*) and /ʊ/-/u/ (*hood-who'd*) continua from Talker 3's stimuli

Steps	/ɪ/-/i/ (<i>hid-heed</i>)		/ʊ/-/u/ (<i>hood-who'd</i>)	
	F1 (Hz)	F2 (Hz)	F1 (Hz)	F2 (Hz)
1	246	2,250	514	1,317
2	269	2,211	416	1,161
3	341	2,114	390	1,094
4	396	2,011	303	985
5	452	1,892	286	948

The stimuli from three talkers (Talkers 1, 2, and 3) are designated for use only in the training sessions, while those from the remaining talker (Talker 4) will be used for the new talker generalization test, which was completed after the last training session (see Section 4.3.5).

The answers for each set of training stimuli were needed to provide feedback on participants' performances during the training. Therefore, a two-alternative forced-choice task with the training stimuli was conducted and completed by 40 native speakers of American English. After listening to a stimulus, participants were instructed to identify the word they had just heard by clicking one of the orthographic words (e.g., *hid* vs. *heed*) appearing on the computer screen. We selected the most frequent responses as the target answers for the training stimuli. This identification task not only shows native English speakers' overall identification patterns but also their use of spectral and duration dimensions in identification. Visual inspection of heat plots

in Figure 1, which shows the proportion of *hid* and *hood* responses, indicates that overall, training stimuli with lower spectral steps were identified as *hid* or *hood*. In comparison, stimuli with higher spectral steps were more likely to be identified as *heed* and *who'd*. These results suggest that native English listeners predominantly used spectral cues while duration cues had a much weaker effect on their vowel categorization, consistent with previous studies (e.g., Baker and Trofimovich 2005).

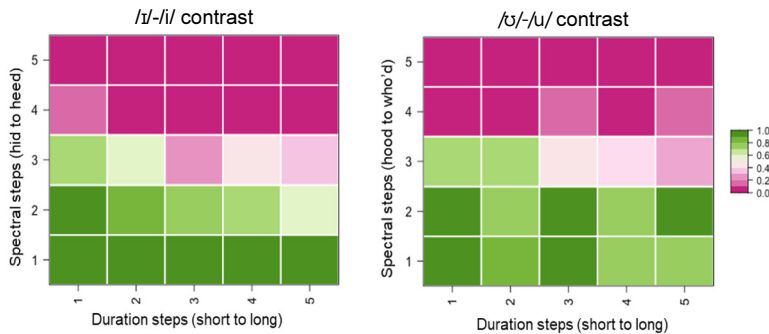


Figure 1. Heatmap plots of native English listeners' identification responses to the training stimuli. Shades closer to green indicate more *hid* or *hood* responses, while shades closer to pink indicate more *heed* or *who'd* responses.

4.3 Procedure

4.3.1 Stroop task

The Stroop task was administered through Paradigm software (Paradigm Stimulus Presentation 2007). Participants first read the detailed instructions that appeared on the computer screen. They were instructed to determine the color of the ink of each word and press the corresponding button on the response box as quickly as possible. Only the rightmost two buttons and the leftmost two buttons on the response box were used for responses, and each button was marked with one of the colors. Participants were asked to place their index and middle fingers from both hands on the response box and use only those fingers to press the button. The order of buttons with colors was counterbalanced. To ensure that participants fully understood the task, a brief practice phase consisting of six trials was provided. After the practice,

participants were presented with the 72 experimental trials. They first saw a fixation mark appearing on the computer screen for 250 ms, and once the fixation mark disappeared, the experimental item was presented. The item disappeared after a maximum of 2,500 ms for the response. The reaction time (RT) for each response was collected. The entire session took about seven minutes.

4.3.2 English vowel contrast training

The entire experiment was conducted over a span of five days, comprising five consecutive online, computer-based training sessions. Once participants started their first training session, no more than two-day intervals were allowed between the training sessions to ensure their consistent participation. The training sessions were administered through PsyToolkit (Stoet 2010, 2017). The training design employed a pretest-posttest procedure closely modeled after the methods used in previous phonetic training studies (e.g., Strange and Dittmann 1984; Logan et al. 1991). The effects of training and participants' improvements were assessed through everyday two-alternative forced-choice (2AFC) identification (ID) tests (hereafter referred to as everyday ID tests), which were administered after each training session. After completing all the training sessions, participants were tested again to assess the degree of generalization of learning to novel stimuli produced by a new talker. The following paragraphs explain the details of each day of participation.

Table 2. Timeline of the 5-day experimental phase with five training sessions

Experiment Day	Day 1	Day 2	Day 3	Day 4	Day 5
Tasks	• Daily familiarization phase • Pre Identification (ID) Test with old talker and new talker • Training Session 1 • Everyday ID test	• Daily familiarization phase • Training Sessions 2-4 • Everyday ID test			• Daily familiarization phase • Training Session 5 • Post ID test with old talker • New talker generalization test • Stroop task
Duration	About 1 hour	About 30 minutes			About 1 hour

4.3.3 Experiment Day 1

All testing and training were conducted at the Phonetics Laboratory of a university in Korea. Participants were seated at individual desks equipped with desktop PCs.

Auditory stimuli were presented over headphones at a comfortable listening level. No more than three participants were allowed to participate in the experiment simultaneously to minimize interruptions or distractions.

Participants started Day 1 with the pre-identification test (pretest, henceforth). The pretest was to measure participants' pre-training states of using spectral and duration dimensions to identify two words for each English vowel contrast. To avoid orthographic bias, photographs were used to represent the target training words. The pretest included stimuli recorded by one of the training talkers (i.e., the old talker) and stimuli recorded by the talker who produced the stimuli for the new talker generalization test (i.e., the new talker). Before the pretest, participants first completed a familiarization phase (Figure 2) to ensure they were familiar with all the target training words, their corresponding photographs, and translated meanings in Korean. Following the familiarization phase, participants heard one stimulus on each trial and were required to press either '1' or '2' on the keyboard. The *hid* or *hood* words always corresponded with the '1' button, and *heed* or *who'd* corresponded with the '2' button. Once participants pressed one of the buttons, a short pause (500 ms) occurred to indicate the start of a new trial. A total of 25 stimuli were repeated three times for each vowel contrast (150 trials with an old talker (25 stimuli $\times$ 2 vowel contrasts $\times$ 3 repetitions) + 150 with a new talker (25 stimuli $\times$ 2 vowel contrasts $\times$ 3 repetitions) = 300 trials).

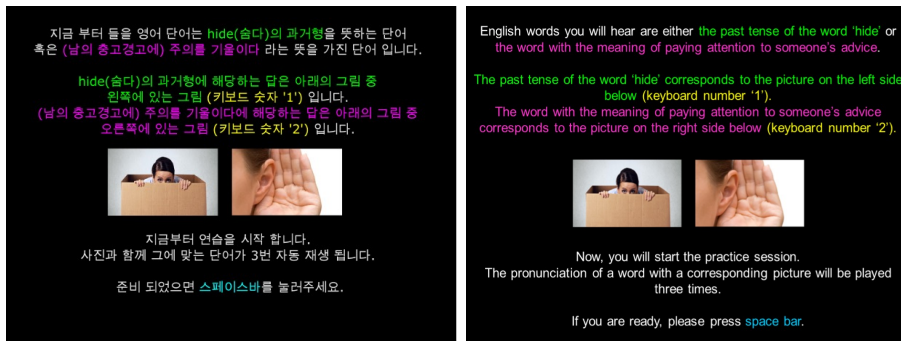


Figure 2. Example instruction page (left) from the daily familiarization phase for the English /ɪ/- /i/ contrast (e.g., *hid* vs. *heed*) and the English translated version (right)

4.3.4 Training days: Session 1 to Session 5

Participants received five days of computer-based auditory training with trial-by-trial feedback from Day 1 to Day 5. Each training session consisted of two blocks, each one for either /ɪ/-i/ (*hid-heed*) or /ʊ/-u/ (*hood-who'd*) contrast. The order of these blocks was counterbalanced. Training stimuli from three talkers were used in each block. It should be noted that stimuli from one talker were randomly presented first before the next talker's set of stimuli was presented. This decision was to reduce the degree of stimulus variability, which may cause excessive confusion for participants (Perrachione et al. 2011). The procedure of the training session was as follows: on each trial, an auditory stimulus was played first, and then two photographs of training words for each English contrast were presented on the computer screen. Participants were then asked to press a button on the keyboard labeled either '1' or '2' to identify the stimulus that had been played. If they chose the wrong picture, the previously played stimulus with the correct picture was replayed as feedback. In each training session, a total of 150 stimuli were presented, derived from 75 stimuli used in each block for each vowel contrast (25 stimuli per contrast × 3 talkers × 2 vowel contrasts). Each training session took approximately 15 minutes. Figure 3 illustrates the steps of each training session and describes the overall process during training, incorporating trial-by-trial feedback.

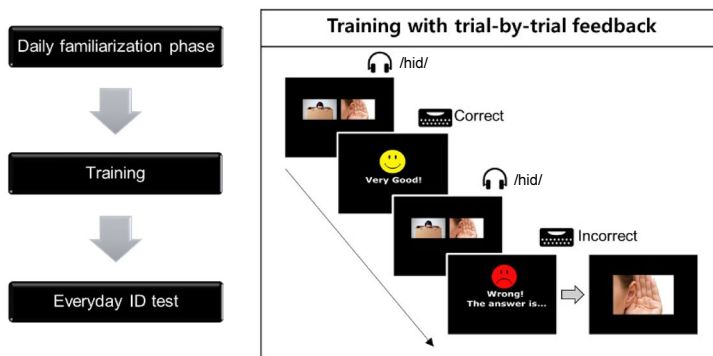


Figure 3. Overview of each training session (left) and example of the trial-by-trial feedback provided during training (right)

Each training session ended with an everyday 2AFC test, which took about 5 minutes. The procedure was identical to the training phase, except no feedback was given. The set of stimuli from Talker 3 was used for this test, yielding a total of 25 trials for each vowel contrast.

4.3.5 Experiment Day 5 (post-training day)

After the last day of the training session (Day 5), participants completed the new talker generalization test and the post-test. The post-test was identical to the pretest. The new talker generalization test presented a total of 50 stimuli (25 stimuli $\times$ 2 contrasts) recorded from one novel talker, whose stimuli were not included as training stimuli, with three repetitions (50 stimuli $\times$ 3 repetitions = 150 trials). The structure of the generalization test was identical to everyday ID tests, except that trial-by-trial feedback was not given.

5. Results

5.1 Inhibitory control and everyday ID tests

Individual differences in inhibitory control were assessed by Stroop interference score, calculated as the average difference in response time (in milliseconds) between incongruent and neutral trials (MacLeod 1991). A higher Stroop score corresponds to lower inhibitory control. The percentage of correct responses on the everyday ID tests, administered on Days 1 through 5, was collected for each participant and used for analysis.

A linear mixed-effects regression model was constructed using the *lmer* function from the *lme4* package (version 1.1-27) (Bates et al. 2015) in *R* (R Core Team 2020) to investigate whether individual differences in inhibitory control predicted participants' performances in the English vowel training. Participants' everyday ID test scores, averaged across both the /ɪ/-/i/ and /ʊ/-/u/ contrasts, were submitted to the model with two fixed effects: Stroop scores and Training day. Stroop scores were standardized by centering and dividing by two standard deviations before they were entered into the model (Gelman 2008). Training day (Day 1, Day 2, Day 3, Day

4, and Day 5) were dummy-coded using the *lizContrasts5* function² with Day 1 as the reference level. This yielded contrasts comparing Day 1 scores with those of the subsequent days: Day 1 versus Day 2, Day 1 versus Day 3, Day 1 versus Day 4, and Day 1 versus Day 5. The model included random intercepts for participants and by-participant random slopes for Training day.

Table 3 shows the results of the regression model. Each coefficient represents the estimated effect when all other predictors are controlled for. To briefly recap, this analysis aimed to examine how individual differences in inhibitory control influence L2 phonological contrasts learning. If inhibitory control plays a role, it was predicted that variability in participants' success in the English vowel training would be associated with their Stroop scores.

The model found a marginally significant main effect of Stroop scores, suggesting that higher overall everyday ID test scores may be associated with lower inhibitory control ability. Figure 4 illustrates this relationship: as standardized Stroop scores increased (i.e., indicating lower/reduced inhibitory control), the average identification accuracy of the target vowel contrasts also increased. Additionally, significant main effects were found for Day1 VERSUS Day3, Day1 VERSUS Day4, and Day1 VERSUS Day5, indicating that participants performed better on Day 3, 4, and 5 compared to Day 1. The increasing estimates for these Training day comparisons reflect the effect of English vowel training and suggest that participants improved in identifying the target English vowel contrasts as they received more training. However, because the model did not find any significant two-way interactions between Stroop scores and Training day contrasts, individual differences in inhibitory control do not appear to be associated with the magnitude of improvements throughout training.

Table 3. Summary of fixed effects from the linear mixed-effects regression model predicting averaged everyday ID test scores

Fixed Effects	Estimate	SE	z	p
Intercept	73.83	1.83	40.29	<0.001
Stroop scores	7.56	3.68	2.05	0.043
Day1 VERSUS Day2	3.30	2.54	1.30	0.197
Day1 VERSUS Day3	5.61	2.33	2.41	0.018
Day1 VERSUS Day4	8.04	2.91	2.77	0.007

² *lizContrasts5* is the modified version of *lizContrasts4* (Wonnacott et al. 2017; Dong et al. 2019) to create four dummy variables that stand in place of a five-way factor.

Day1 VERSUS Day5	11.13	2.11	5.29	<0.001
Stroop scores × Day1 VERSUS Day2	1.78	5.10	0.35	0.729
Stroop scores × Day1 VERSUS Day3	3.84	4.68	0.82	0.413
Stroop scores × Day1 VERSUS Day4	1.60	5.84	0.27	0.785
Stroop scores × Day1 VERSUS Day5	0.95	4.23	0.23	0.822

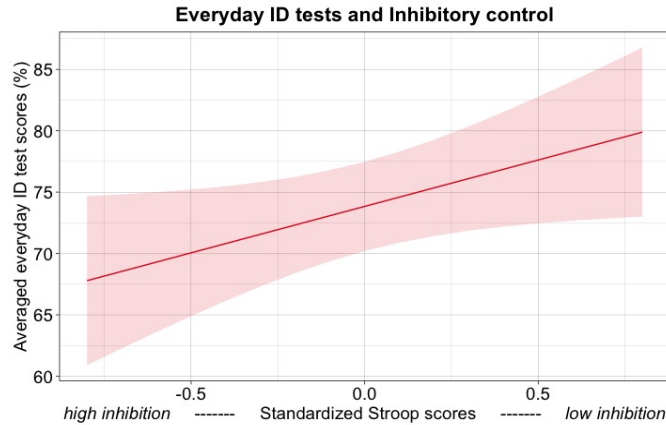


Figure 4. Relationship between individual differences in Stroop scores and averaged everyday ID test scores

5.2 Inhibitory control and primary and secondary cue use

The results of the pre- and post-tests, as well as the new talker generalization test, were analyzed to investigate how participants' use of primary (spectral) and secondary (duration) cues changed after training in identifying the target English vowel contrasts. A mixed-effects logistic regression (Jaeger 2008) was used to investigate how changes in acoustic cue cues are related to individual differences in inhibitory control, using the *glmer()* function from the *lme4* package (version 1.1-27) in *R*. As a reminder, the pre- and post-tests included stimuli recorded by the same talker who also produced the vowel training stimuli (i.e., *old* talker), while the new talker generalization test used stimuli produced by a *new* talker who was not a part of training stimuli set. Therefore, results from the pre- and post-tests were analyzed separately from those of the generalization test.

For each type of test, two separate mixed-effects logistic regression models were constructed: one for the /ɪ/-/i/ and the other for the /ʊ/-/u/ contrasts. Participants'

responses were binary, with *hid* and *hood* responses coded as 0, and *heed* and *who'd* responses coded as 1. Each model included four fixed effects: Stroop scores, Spectral, Duration, and Time predictors. Spectral and Duration were continuous variables consisting of five spectral steps and five duration steps of the auditory stimuli. Both variables were standardized by centering and dividing by two standard deviations. For the pre- and post-test analysis, Time was coded as Pre and Post; for the new talker generalization test analysis, Time was coded as Before and After. In all models, Time was centered (-0.5 and 0.5) to allow the main effects to be interpreted as average effects over both levels of Time. All regression models included random intercepts for participants, along with random slopes for participants for Time variable.

5.2.1 Pre- and Post-tests (old talker)

Figure 5 plots the estimated proportion of *heed* (left-side graphs) and *who'd* (right-side graphs) responses on the pre- and post-tests for the highest and lowest spectral steps of test stimuli, as a function of Stroop scores. Visual inspection of Figure 5 suggests that reliance on spectral cues was more evident in the post-test than in the pretest. Additionally, the greater use of spectral cues in identifying target vowel contrasts appears to be associated with individual differences in inhibitory control. Specifically, participants with relatively low inhibitory control ability (i.e., higher Stroop scores) demonstrated more nativelike use of spectral cues by giving more *heed* and *who'd* responses for the stimuli with the highest spectral step (step 5) than those with the lowest spectral step (step 1).

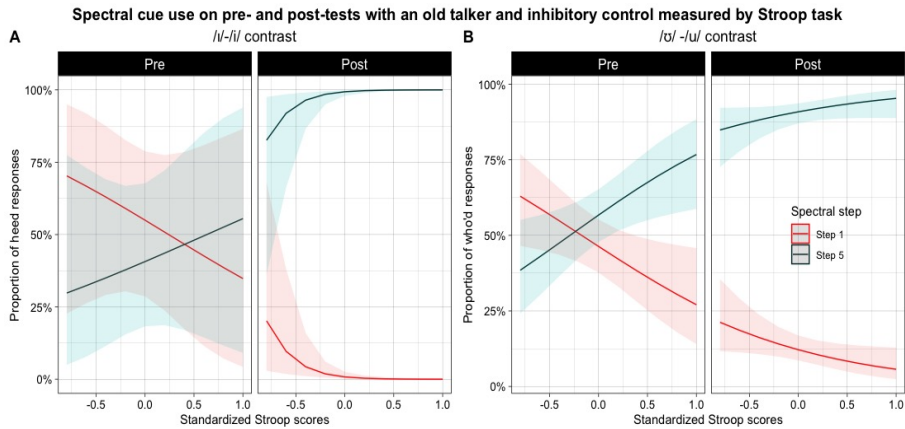


Figure 5. Effects of individual differences in Stroop scores on spectral cue use before and after training. The figure shows results from the pre- and post-tests using stimuli produced by the “old” talker (i.e., the talker used during training)

Tables 4 and 5 summarize the logistic regression models for the /I/-i/ and /ʊ/-u/. Both models found a significant interaction between Spectral and Time, indicating that the use of spectral cues increased from pre- to post-test. Furthermore, both models identified a significant 3-way interaction between Stroop scores, Spectral, and Time. This interaction suggests that the increase in spectral cue use over time was associated with individual differences in inhibitory control. Taken together, these results indicate that participants increased their reliance on spectral cues as training progressed, and this increase was greater for participants with relatively low inhibitory control.

Table 4. Summary of fixed effects from the mixed-effects logistic regression model on listeners’ responses in the pre- and post-test for the /I/-i/ contrast

Fixed Effects	Estimate	SE	z	p
Intercept	0.01	0.13	0.09	0.928
Stroop scores	-0.04	0.25	-0.17	0.861
Duration	0.27	0.34	0.81	0.421
Spectral	3.27	0.79	4.17	<0.001
Time	0.20	0.11	1.88	0.060
Stroop scores × Duration	-0.07	0.67	-0.11	0.913
Stroop scores × Spectral	3.56	1.57	2.27	0.023
Stroop scores × Time	0.14	0.22	0.64	0.522
Duration × Time	-0.47	0.22	-2.19	0.029
Spectral × Time	7.36	0.36	20.30	<0.001

Stroop scores $\times$ Spectral $\times$ Time	5.11	0.74	6.91	<0.001
Stroop scores $\times$ Duration $\times$ Time	0.69	0.44	1.58	0.114

Table 5. Summary of fixed effects from the mixed-effects logistic regression model on listeners' responses in the pre- and post-test for the /ɔ/-/u/ contrast

Fixed Effects	Estimate	SE	z	p
Intercept	0.08	0.10	0.73	0.463
Stroop scores	0.01	0.21	0.05	0.961
Duration	0.74	0.28	2.69	0.007
Spectral	1.72	0.34	5.11	<0.001
Time	0.35	0.09	4.04	<0.001
Stroop scores $\times$ Duration	-0.32	0.55	-0.57	0.567
Stroop scores $\times$ Spectral	-0.24	0.67	-0.36	0.722
Stroop scores $\times$ Time	0.57	0.18	3.17	0.002
Duration $\times$ Time	0.20	0.18	1.11	0.265
Spectral $\times$ Time	2.62	0.19	13.58	<0.001
Stroop scores $\times$ Spectral $\times$ Time	2.78	0.41	6.80	<0.001
Stroop scores $\times$ Duration $\times$ Time	-1.11	0.37	-3.00	0.003

5.2.2 New talker generalization test

The new talker generalization test employed a set of stimuli produced by a novel talker to assess whether participants could generalize their learning of how to use acoustic cues, particularly using spectral cues as the primary dimension, to untrained stimuli. Figure 6 illustrates the estimated proportion of *heed* (left-side graphs) and *who'd* (right-side graphs) responses during the new talker generalization test conducted before and after the training.

When comparing two vowel contrasts, the generalization test results depicted in Figure 6 revealed a notable influence of individual differences in inhibitory control on the use of spectral cues after training, particularly evident in the /ɔ/-/u/ contrast. Initially, prior to the training, participants showed minimal reliance on spectral cues for the /ɔ/-/u/ contrast, irrespective of their inhibitory control ability. After the training, however, the significance of spectral cues in identifying new talker stimuli notably increased, especially as participants' Stroop score increased (i.e., lower inhibitory control). This is evident in the greater difference between the estimated proportion of *who'd* responses for stimuli with spectral steps 1 versus 5 in the after-training graphs as Stroop scores increased. This trend suggests that the /ɪ/-/i/

contrast may be easier to acquire and more susceptible to changes in cue weights over time for all learners, regardless of their differences in inhibitory control. Conversely, for contrasts like /ʊ/-/u/, which are relatively more difficult to acquire, the relationship between learning this contrast and individual traits, such as inhibitory control, may become clearer.

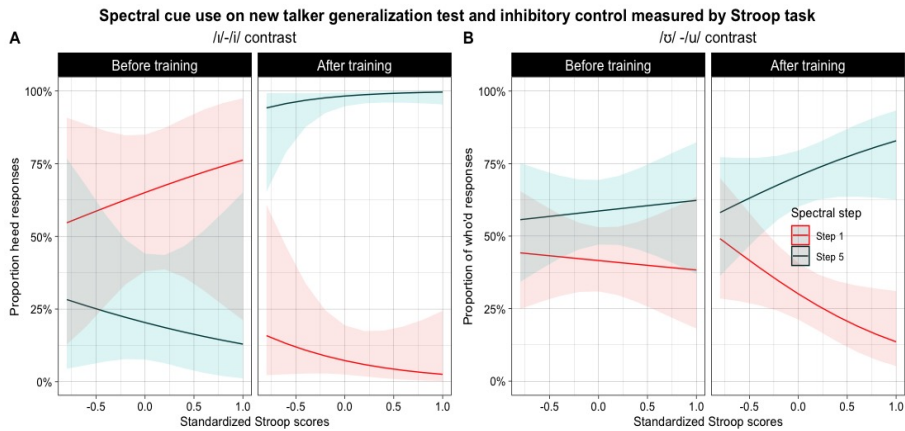


Figure 6. Effects of individual differences in Stroop scores on spectral cue use before and after training. The figure shows results from the new talker generalization test

The results of logistic regression analyses for the /ɪ/-/i/ and /ʊ/-/u/ contrasts (Tables 6 and 7) mirrored the patterns observed in the pre- and post-tests. Both models found a significant two-way interaction between Spectral and Time, as well as a significant three-way interaction between Stroop scores, Spectral, and Time. These interactions indicate that the spectral cue use changed after training and that the magnitude of this change was contingent on participants' inhibitory control abilities. When comparing the two contrasts, the models showed that the increase in spectral cue reliance after training was greater for the /ɪ/-/i/ than for the /ʊ/-/u/ (/ɪ/-/i/: $\beta_{\text{Spectral} \times \text{Time}} = 6.04$ vs. /ʊ/-/u/: $\beta_{\text{Spectral} \times \text{Time}} = 0.73$). It is important to note that the larger increase for the /ɪ/-/i/ contrast stemmed from participants' initial opposite use of spectral cues prior to training (i.e., more *heed* responses for spectral step 1 stimuli over step 5 stimuli). Consequently, the after-training use of spectral cues for the /ɪ/-/i/ contrast is inevitably increased to a larger degree than the /ʊ/-/u/ contrast.

Table 6. Summary of fixed effects from the mixed-effects logistic regression model on listeners' responses in the new talker generalization test for the /ɪ/-/i/ contrast

Fixed Effects	Estimate	SE	z	p
Intercept	0.19	0.18	1.07	0.285
Stroop scores	0.12	0.36	0.34	0.737
Duration	0.53	0.29	1.84	0.065
Spectral	1.62	0.76	2.14	0.032
Time	1.12	0.11	10.65	<0.001
Stroop scores × Duration	-0.32	0.58	-0.55	0.584
Stroop scores × Spectral	0.56	1.51	0.37	0.712
Stroop scores × Time	0.24	0.22	1.09	0.275
Duration × Time	-0.54	0.20	-2.69	0.007
Spectral × Time	6.04	0.26	22.98	<0.001
Stroop scores × Spectral × Time	2.64	0.54	4.93	<0.001
Stroop scores × Duration × Time	0.89	0.41	2.16	0.030

Table 7. Summary of fixed effects from the mixed-effects logistic regression model on listeners' responses in the new talker generalization test for the /ʊ/-/u/ contrast

Fixed Effects	Estimate	SE	z	p
Intercept	0.01	0.11	0.11	0.916
Stroop scores	-0.07	0.23	-0.33	0.744
Duration	0.38	0.27	1.42	0.154
Spectral	0.85	0.28	3.07	0.002
Time	0.02	0.08	0.19	0.850
Stroop scores × Duration	-0.19	0.54	-0.35	0.730
Stroop scores × Spectral	0.70	0.56	1.26	0.206
Stroop scores × Time	-0.17	0.16	-1.04	0.300
Duration × Time	1.14	0.16	6.93	<0.001
Spectral × Time	0.73	0.16	4.44	<0.001
Stroop scores × Spectral × Time	1.00	0.34	2.97	0.003
Stroop scores × Duration × Time	-1.37	0.33	-4.11	<0.001

6. Discussion

This study examined how individual differences in inhibitory control influence Korean learners' acquisition of two English vowel contrasts, /ɪ/-/i/ and /ʊ/-/u/, through five days of auditory training. Our results showed that L2 learners can learn to prioritize more relevant acoustic dimensions (i.e., spectral cues) to English vowel contrasts after training. However, these results were found to be associated with individual differences

in inhibitory control ability: learners with relatively lower inhibitory control, as measured by the Stroop task, demonstrated greater improvement in identifying the target vowel contrasts and increased their reliance on spectral cues. Taken together, the current findings suggest that reduced inhibitory control may be advantageous in certain L2 learning situations, especially when learners need to shift their attention away from initially biased cues to more L2-relevant ones.

The results from the everyday ID tests revealed that learners with relatively lower inhibitory control showed greater improvement across training days. Although the interactions between Stroop scores and training days were not statistically significant, the main effect of Stroop scores suggests that reduced inhibitory control may be associated with more successful L2 learning outcomes. Analyses of pre- and post-test performance further demonstrated that participants with lower inhibitory control showed a greater increase in their reliance on spectral cues over time. When comparing the two vowel contrasts in the generalization test results (section 5.2.2), the effect of inhibitory control was more evident for the /ʊ/-/u/ contrast than for the /ɪ/-/i/ contrast. Regardless of individual differences in inhibitory control, all participants appeared to generalize their learning of spectral cue use for the /ɪ/-/i/ contrast. In contrast, only participants with relatively low inhibitory control successfully extended their generalization of learning to the /ʊ/-/u/ contrast. Consistent with previous studies on the relative difficulty between these two vowel contrasts (e.g., Baker et al. 2002; Nishi and Kewley-Port 2007, 2008; Lee 2009), the current study confirms that not all learners successfully acquired the ability to use spectral cues to distinguish the /ʊ/-/u/ contrast, and learners' inhibitory control may partially contribute to these individual differences in L2 phonological contrast learning. According to PAM/PAM-L2 (Best 1995; Best and Tyler 2007), both target English contrasts are predicted to be challenging as category goodness assimilation cases. The findings of this study reconfirm the effectiveness of short-term laboratory training demonstrated in previous studies and further suggest that perceptually difficult L2 contrasts may not be perceived with the same degree of difficulty partially depending on learners' inhibitory control ability.

The results of the current study are somewhat surprising, given previous findings in bilingualism and L2 acquisition, which suggest that stronger inhibitory control is beneficial, as it reduces interference from a language that is not in use or learners' L1 (e.g., Darcy et al. 2016). As suggested in Amer et al. (2016), reduced inhibitory

control may enhance performance under certain circumstances. For example, if some tasks require the use of previously irrelevant information, which is therefore inhibited, or can benefit from drawing on various bits of information from multiple sources, reduced inhibitory control may be beneficial for success in these tasks. In the case of our participants, Korean learners of English, all of them completed their formal English education in Korea, which often emphasizes the durational differences between the English tense-lax vowel contrasts. Both instructional influences, focusing on durational differences and reduced perceptual sensitivity to formant differences in unfamiliar areas of the vowel space (Bohn and Flege 1990; Lee 2009), may have contributed to their initial primary reliance on duration cues—a cue-weighting strategy that deviates from that of English native listeners. In this study, participants' reliance on duration cues before training was reflected in the low pretest scores (average identification accuracy: 51.5%). Assuming learners' attention was initially tuned to duration, the key to successful identification of English vowels would involve shifting attention to spectral cues, which were previously irrelevant in their identification, and suppressing the use of duration cues.

Crucially, before learners can effectively shift their attention and change cue-weighting strategies, they must first recognize which acoustic cues are available and their relevant importance in distinguishing L2 sounds. Without this exploratory phase, learners may persist in relying on duration or may shift attention to other non-optimal (i.e., less-relevant) cues. In this respect, learners with low inhibitory control may be in a better position because their broad focus of attention, which is a possible by-product of their ability, enables them to pay attention to various sources of acoustic information and explore them.

Kim et al. (2020) discussed the possible benefits of reduced inhibitory control in speech adaptation. In their study, listeners adapted to unfamiliar native vowel contrasts by up-weighting a secondary cue, and individual differences in inhibitory control (measured by the Stroop task) predicted the degree of adaptation: individuals with reduced inhibitory control showed greater adaptation (see section 2.2 for more details). Although Kim et al. (2020) is not an L2 study, their findings, along with those of Amer et al. (2016), support the idea that learners' broader focus of attention—possibly induced by low inhibitory control—may help L2 learners identify which cues are available and most relevant for distinguishing L2 contrasts. This process likely precedes and facilitates later modification in their cue-weighting strategies.

However, it is worth noting that the stage of L2 learning may be a crucial factor to consider when discussing possible benefits of reduced inhibitory control in L2 phonological contrast learning. In earlier or intermediate stages of learning, a broader focus of attention may be beneficial, as explained above. However, once learners have identified the appropriate primary acoustic cues, stronger inhibitory control may become advantageous. Learners' strong inhibitory control may lead them to suppress the influence of less-relevant acoustic cues, resulting in more nativelike and reliable L2 speech perception. Previous research with native speakers provides relevant insights. For instance, Kong and Lee (2019) investigated the relationship between individual differences in cognitive abilities, such as inhibitory control, and the use of primary and secondary acoustic cues. Their results revealed that Korean listeners with stronger inhibitory control were better at suppressing irrelevant acoustic information when identifying Korean stop laryngeal categories. Assuming that native listeners fully understand which cues are more relevant than others, L2 learners may exhibit a similar pattern like native listeners as they gain more L2 experience. In the current study, all participants had similar L2 proficiency levels at the time of participation. Therefore, further research should include learners at different L2 proficiency levels or employ a longitudinal design to examine how the relationship between inhibitory control and L2 learning changes over time.

Although inhibitory control can be broadly understood as the ability to suppress task-irrelevant information, Friedman and Miyake (2004) discussed various types of inhibition-related functions and measurements (i.e., tasks) that may primarily measure one of these functions. For example, Friedman and Miyake (2004) used the Stroop task to measure prepotent response inhibition, which refers to the ability to deliberately suppress dominant, automatic, or prepotent responses, as the task requires participants to inhibit such responses based on the lexical information of the stimuli. The Retrieval-induced (RI) inhibition task used in Darcy et al. (2016) may be more appropriate for measuring resistance to proactive interference, which refers to the ability to resist information that was previously relevant to the task but has since become irrelevant. However, as pointed out in previous research (e.g., Huensch et al. 2015; Linck and Weiss 2015), the question of which inhibition measures best capture which types of inhibitory-related functions remains unclear. Thus, it is possible that different choices of inhibition measures may influence the examination of the relationship between inhibitory control and L2 learning, as evidenced by the mixed

results in previous studies. Friedman and Miyake (2004) suggested that using multiple measures of inhibition might be a practical solution to the issues of low reliability associated with inhibition measures. The current study employed only a single task—the Stroop task—to measure inhibitory control and found that a broader focus of attention, possibly resulting from reduced inhibitory control, may bring benefits during training. Thus, future research utilizing multiple inhibition measures within the same study is essential to investigate whether this pattern is observable across other tasks and which measures (or type of inhibitory functions) best capture the relationship between inhibitory control and L2 phonological learning.

Given the wide range of variability in L2 learning, understanding sources of these variations is crucial for providing more effective training tailored to individual learners, thereby avoiding a one-size-fits-all approach. To achieve this, we need reliable pre-instructional assessments to predict L2 learners' achievements before receiving training. The current study highlights the pedagogical potential of using inhibition measures as pre-instructional assessments, especially the Stroop task. Furthermore, our findings suggest that L2 phonetic training, which emphasizes the acoustic characteristics of target contrasts and raises learners' awareness of their relative importance, may lead to more effective learning outcomes. However, it remains unclear whether the robust training effects observed in laboratory settings can be replicated in real-world contexts, such as classroom pronunciation teaching. As highlighted in Wang and Munro (2004: 540), further research examining the pedagogical application of L2 phonetic training paradigms in a classroom setting will help bridge the significant gap between the key findings from laboratory studies and the techniques implemented in practice.

7. Conclusion

The current study contributes to growing evidence that individual cognitive traits—particularly inhibitory control—can influence the acquisition of L2 phonological contrasts. The findings suggest that lower inhibitory control may facilitate improved identification of the target L2 contrasts. Furthermore, the results indicate that participants demonstrated a significant increase in reliance on relevant acoustic cues over time, particularly those with lower inhibitory control. This study suggests that

a broader focus of attention—a possible by-product of low inhibitory control—may assist language learners in noticing the existence of more relevant acoustic cues in L2 and successfully shifting their attention to those cues. In conclusion, this study highlights the importance of understanding the possible sources of individual differences among L2 learners, including cognitive factors such as inhibitory control.

References

- Amer, Tarek, Karen L. Campbell, and Lynn Hasher. 2016. Cognitive control as a double-edged sword. *Trends in Cognitive Sciences* 20(12): 905-915.
- Baker, Wendy and Pavel Trofimovich. 2005. Interaction of native-and second-language vowel system(s) in early and late bilinguals. *Language and Speech* 48(1): 1-27. <https://doi.org/10.1177/00238309050480010101>.
- Baker, Wendy, Pavel Trofimovich, Molly Mack, and James E. Flege. 2002. The effect of perceived phonetic similarity on non-native sound learning by children and adults. *Proceedings of the Annual Boston University Conference on Language Development* 26(1): 36-47. Somerville: Cascadilla Press.
- Bates, Douglas, Martin Mächler, Ben Bolker, and Steve Walker. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67: 1-48. <https://doi.org/10.18637/jss.v067.i01>.
- Bender, Angela D., Hannah L. Filmer, K. G. Garner, Claire K. Naughtin, and Paul E. Dux. 2016. On the relationship between response selection and response inhibition: An individual differences approach. *Attention, Perception, and Psychophysics* 78(8): 2420-2432.
- Best, Catherine T. 1995. A direct realist view of cross-language speech perception. In Strange, Winifred (eds.), *Speech perception and linguistic experience*, 171-204. Timonium: York Press.
- Best, Catherine T. and Michael Tyler. 2007. Nonnative and second-language speech perception: Commonalities and complementarities. In Murray, J. Munro and Ocke-Schwen. Bohn (eds.), *Second language speech learning: The role of language experience in speech perception and production*, 13-34. Amsterdam: John Benjamins Publishing Company.
- Boersma, Paul. 2001. Praat, a system for doing phonetics by computer. *Glott International* 5(9/10): 341-345. Retrieved from <https://ci.nii.ac.jp/naid/10026090047/>.
- Bohn, Ocke-Schwen and James Emil Flege. 1990. Interlingual identification and the role of foreign language experience in L2 vowel perception. *Applied Psycholinguistics* 11(3): 303-328.
- Darcy, Isabelle, Hanyong Park, and Chung-Lin Yang. 2015. Individual differences in L2 acquisition of English phonology: The relation between cognitive abilities and phonological processing. *Learning and Individual Differences* 40: 63-72.

- Dong, Hanyu, Meghan Clayards, Helen Brown, and Elizabeth Wonnacott. 2019. The effects of high versus low talker variability and individual aptitude on phonetic training of Mandarin lexical tones. *PeerJ* 7: e7191. <https://doi.org/10.7717/peerj.7191>.
- Flege, James Emil, Ocke-Schwen Bohn, and Sunyoung Jang. 1997. Effects of experience on non-native speakers' production and perception of English vowels. *Journal of Phonetics* 25(4): 437-470.
- Francis, Alexander L. and Howard C. Nusbaum. 2002. Selective attention and the acquisition of new phonetic categories. *Journal of Experimental Psychology: Human Perception and Performance* 28(2): 349-366.
- Francis, Alexander L., Kate Baldwin, and Howard C. Nusbaum. 2000. Effects of training on attention to acoustic cues. *Perception & Psychophysics* 62(8): 1668-1680. <https://doi.org/10.3758/BF03212164>.
- Francis, Alexander L., Natalya Kaganovich, and Courtney Driscoll-Huber. 2008. Cue-specific effects of categorization training on the relative weighting of acoustic cues to consonant voicing in English. *The Journal of the Acoustical Society of America* 124(2): 1234-1251.
- Friedman, Naomi P. and Akira Miyake. 2004. The relations among inhibition and interference control functions: A latent-variable analysis. *Journal of Experimental Psychology: General* 133(1): 101-135.
- Gollan, Tamar H., Tiffany Sandoval, and David P. Salmon. 2011. Cross-language intrusion errors in aging bilinguals reveal the link between executive control and language selection. *Psychological Science* 22(9): 1155-1164.
- Green, David. W. 1998. Mental control of the bilingual lexico-semantic system. *Bilingualism: Language and Cognition* 1(2): 67-81.
- Han, Jong-Im. 2015. *English phonetics and teaching English pronunciation*. Seoul: Hankookmunwhasa.
- Hillenbrand, James M., Michael J. Clark, and Robert A. Houde. 2000. Some effects of duration on vowel recognition. *The Journal of the Acoustical Society of America* 108(6): 3013-3022.
- Hillenbrand, James M., Laura A. Getty, Michael J. Clark, and Kimberlee Wheeler. 1995. Acoustic characteristics of American English vowels. *The Journal of the Acoustical Society of America* 97(5): 3099-3111.
- Holt, Lori L. and Andrew J. Lotto. 2006. Cue weighting in auditory categorization: Implications for first and second language acquisition. *The Journal of the Acoustical Society of America* 119(5): 3059-3071.
- Houdé, Olivier, Laure Zago, Emmanuel Mellet, Sylvain Moutier, Arlette Pineau, Bernard Mazoyer, and Nathalie Tzourio-Mazoyer. 2000. Shifting from the perceptual brain to the logical brain: The neural impact of cognitive inhibition training. *Journal of Cognitive Neuroscience* 12(5): 721-728.
- Huensch, Amanda. 2024. Clarifying the role of inhibitory control in L2 phonological processing: A preregistered, close replication of Darcy et al. (2016). *Studies in Second Language*

- Acquisition* 46(5): 1392-1412.
- Jaeger, T. Florian. 2008. Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Memory and Language* 59(4): 434-446.
- Kawahara, Hideki, Toru Takahashi, Masanori Morise, and Hideki Banno. 2009. Development of exploratory research tools based on TANDEM-STRAIGHT. *Proceedings of APSIPA ASC 2009: Asia-Pacific Signal and Information Processing Association, 2009 Annual Summit and Conference*, 111-120. Asia-Pacific signal and information processing association.
- Kim, Donghyun, Meghan Clayards, and Eun Jong Kong. 2020. Individual differences in perceptual adaptation to unfamiliar phonetic categories. *Journal of Phonetics* 81: 100984. <https://doi.org/10.1016/j.wocn.2020.100984>.
- Kim, Donghyun, Meghan Clayards, and Heather Goad. 2018. A longitudinal study of individual differences in the acquisition of new vowel contrasts. *Journal of Phonetics* 67: 1-20. <https://doi.org/10.1016/j.wocn.2017.11.003>.
- Kondaurova, Maria V. and Alexander L. Francis. 2008. The relationship between native allophonic experience with vowel duration and perception of the English tense/lax vowel contrast by Spanish and Russian listeners. *The Journal of the Acoustical Society of America* 124(6): 3959-3971.
- Kong, Eun Jong and Hyunjung Lee. 2021. Relationship between executive function and cue weighting in Korean stop perception across different dialects and ages. *Phonetics and Speech Sciences* 13(3): 21-29.
- Lee, Ji-Yeon. 2008. Perception of English high vowels: Duration as a cue by Korean speakers of English. *Kansas Working Papers in Linguistics* 30: 195-204. <https://doi.org/10.17161/KWPL.1808.3915>.
- Lee, Ji-Yeon. 2009. *The effects of pronunciation instruction using duration manipulation on the acquisition of English vowel sounds by pre-service Korean EFL teachers*. PhD Dissertation. University of Kansas.
- Lee, Shinsook and Mi-Hui Cho. 2018. Predicting L2 vowel identification accuracy from cross-language mappings between L2 English and L1 Korean. *Language Sciences* 66: 183-198. <https://doi.org/10.1016/j.langsci.2017.09.006>.
- Lemhöfer, Kristin and Mirjam Broersma. 2012. Introducing LexTALE: A quick and valid lexical test for advanced learners of English. *Behavior Research Methods* 44(2): 325-343.
- Lengeris, Angel. 2009. *Individual differences in second-language vowel learning*. PhD Dissertation. University College London.
- Lev-Ari, Shiri and Sharon Peperkamp. 2013. Low inhibitory skill leads to non-native perception and production in bilinguals' native language. *Journal of Phonetics* 41(5): 320-331.
- Linck, Jared A. and Daniel J. Weiss. 2015. Can working memory and inhibitory control predict second language learning in the classroom? *Sage Open* 5(4): 2158244015607352.
- Linck, Jared A., John W. Schwieter, and Gretchen Sunderman. 2012. Inhibitory control predicts language switching performance in trilingual speech production. *Bilingualism: Language and*

- Cognition* 15(3): 651-662.
- Logan, John S., Scott E. Lively, and David B. Pisoni. 1991. Training Japanese listeners to identify English /r/ and /l/: A first report. *The Journal of the Acoustical Society of America* 89(2): 874-886.
- MacLeod, Colin M. 1991. Half a century of research on the Stroop effect: An integrative review. *Psychological Bulletin* 109(2): 163.
- Miyake, Akira, Michael J. Emerson, and Naomi P. Friedman. 2000. Assessment of executive functions in clinical settings: Problems and recommendations. *Seminars in Speech and Language* 21(2): 169-183.
- Nishi, Kanae and Diane Kewley-Port. 2007. Training Japanese listeners to perceive American English vowels: Influence of training set. *Journal of Speech, Language, and Hearing Research: JSLHR* 50(6): 1496-1509. [https://doi.org/10.1044/1092-4388\(2007/103\)](https://doi.org/10.1044/1092-4388(2007/103)).
- Nishi, Kanae and Diane Kewley-Port. 2008. Non-native speech perception training using vowel subsets: Effects of vowels in sets and order of training. *Journal of Speech, Language, and Hearing Research: JSLHR* 51(6): 1480-1493. [https://doi.org/10.1044/1092-4388\(2008/07-0109\)](https://doi.org/10.1044/1092-4388(2008/07-0109)).
- Paradigm Stimulus Presentation. 2007. Perception research systems. Retrieved from Retrieved from <http://www.paradigmexperiments.com>.
- Perrachione, Tyler K., Jiyeon Lee, Louisa YY Ha, and Patrick Wong. 2011. Learning a novel phonological contrast depends on interactions between individual differences and training paradigm design. *The Journal of the Acoustical Society of America* 130(1): 461-472. <https://doi.org/10.1121/1.3593366>.
- R Core Team. 2020. R Core Team R: A language and environment for statistical computing. Vienna: R Foundation for Statistical Computing.
- Schertz, Jessamyn, Taehong Cho, Andrew Lotto, and Natasha Warner. 2015. Individual differences in phonetic cue use in production and perception of a non-native sound contrast. *Journal of Phonetics* 52: 183-204. <https://doi.org/10.1016/j.wocn.2015.07.003>.
- Stoet, Gijsbert. 2010. PsyToolkit: A software package for programming psychological experiments using Linux. *Behavior Research Methods*, 42(4): 1096-1104.
- Stoet, Gijsbert. 2017. PsyToolkit: A novel web-based method for running online questionnaires and reaction-time experiments. *Teaching of Psychology* 44(1): 24-31. <https://doi.org/10.1177/0098628316677643>.
- Strange, Winifred and Sibylla Dittmann. 1984. Effects of discrimination training on the perception of /r-l/ by Japanese adults learning English. *Perception & Psychophysics* 36(2): 131-145.
- Stroop, J. Ridley. 1935. Studies of interference in serial verbal reactions. *Journal of Experimental Psychology* 18(6): 643-662.
- Wang, Xinchun and Murray J. Munro. 2004. Computer-based training for learning English vowel contrasts. *System* 32(4): 539-552.
- Wonnacott, Elizabeth, Helen Brown, and Kate Nation. 2017. Skewing the evidence: The effect

of input structure on child and adult learning of lexically based patterns in an artificial language. *Journal of Memory and Language* 95: 36-48. <https://doi.org/10.1016/j.jml.2017.01.005>.

Yang, Byunggon. 1996. A comparative study of American English and Korean vowels produced by male and female speakers. *Journal of Phonetics* 242: 245-261.

Jieun Lee

Visiting Assistant Professor
Department of Linguistics
University of Kansas
1541 Lilac Lane
Lawrence, KS 66045 U.S.A.
E-mail: jieunlee@ku.edu

Hanyong Park

Associate Professor
Department of Linguistics
University of Wisconsin-Milwaukee
2522 East Hartford Ave.
Milwaukee, WI 53211 U.S.A.
E-mail: park27@uwm.edu

Received: 2025. 07. 28.

Revised: 2025. 08. 12.

Accepted: 2025. 08. 27.