CROSS-LINGUISTIC INFLUENCE IN L1 PHONETIC CATEGORIES IN KOREAN HERITAGE SPEAKERS AND LONG-TERM IMMIGRANTS

by

Yuhyeon Seo

A Dissertation

Submitted to the Faculty of Purdue University

In Partial Fulfillment of the Requirements for the degree of

Doctor of Philosophy

Department of Linguistics

West Lafayette, Indiana

May 2024

THE PURDUE UNIVERSITY GRADUATE SCHOOL
STATEMENT OF COMMITTEE APPROVAL

Dr. Olga Dmitrieva, Chair

School of Languages and Cultures

Dr. Alejandro Cuza Blanco

School of Languages and Cultures

Dr. Alexander L. Francis

Center of Health and Human Sciences

Dr. Daniel J. Olson

School of Languages and Cultures

Approved by:

Dr. Olga Dmitrieva

To my beloved family

ACKNOWLEDGMENTS

The journey through the Ph.D. program at Purdue University has not only elevated my professional career but also facilitated my personal growth. During the journey, I am very fortunate to have met so many wonderful people who have provided me with invaluable support, guidance, and advice.

First and foremost, I cannot thank my advisor, Dr. Olga Dmitrieva, enough for her unconditional dedication to my growth as an independent researcher. Her professional training has been invaluable in my career development as a speech scientist and a linguist specializing in bilingualism. More than the professional training, she has always been a model of a good scholar and leader. The lessons and memories from our shared experiences of presenting and publishing research together and organizing academic events will remain in my heart in the future.

My special thanks go to Dr. Daniel J. Olson for his invaluable support and guidance for my career and professional development. His training ignited my research interests in short-term cross-linguistic influence and spurred my learning of data analysis. I truly enjoyed our collaboration and look forward to the next. I am also sincerely grateful for his invaluable assistance with my preparation for an academic career and his advice to become a good mentor.

Both professionally and personally, I am also indebted to my other Ph.D. committee members, Dr. Alejandro Cuza and Dr. Alexander Francis. I am very grateful to Dr. Cuza for his guidance and support in my research on heritage speakers. Without his support, I could not have reached the current level of accomplishment in my research on heritage language. I also appreciate Dr. Francis' advising and guidance on statistics, acoustic analysis, and speech synthesis. I will not forget the hours we spent together exploring different methods for speech synthesis.

There are also many colleagues whom I thank a lot. I thank the members and alumni of the Purdue Experimental Phonetics and Phonology lab, including Amy Hutchinson, Ye-Jee Jung, Alexis Zhou, Chien-Min Kuo, and Jenna Conklin who have contributed to my learning and growth throughout my graduate program. I also thank the members and alumni of the Purdue Bilingualism lab with whom it has been my pleasure to work, practice, and discuss, including Bruno Staszkiewicz Garcia, Yi-Fang Cheng, Itzel Citlalli Matamoros Santos, Santiago Revelo Castillo, Paige Doyle, Adam Buehrer, April Grotberg, and Karen Fullin.

There are many individuals outside Purdue University who have provided invaluable assistance with my research during the graduate program. I thank Dr. Jennifer Cabrelli (University of Illinois at Chicago) and Dr. Joo-Kyeong Lee (University of Seoul) for not only assisting with data collection but also providing personal advice on my research and academic career. I am also grateful to Dr. Eun Jong Kong (Korea Aerospace University) for her invaluable advice and feedback on experimental methodology. Many thanks to Dr. Anna Schmidt, who retired from Kent State University, for generously sharing her data, which ignited my research interest in Korean laryngeal stops. I also thank many other professionals, including those I met at conferences and anonymous reviewers, who gave me constructive feedback in developing my dissertation project.

The data collection for this dissertation project would not have been possible without the help of several Korean communities in the Indiana and Illinois states. For privacy reasons, I am not disclosing their names, but they deserve a section of gratitude. Most of my participants recruited in the US volunteered to participate in my study not because of the small payment from my grants but because of their altruistic desire to help a young Korean student researcher study their underrepresented language in the US. The data collection process involved driving six hours back and forth between my school and Chicago, which was physically challenging. Yet, their encouragement, support, and cheering elevated my motivation to finish this project and, hopefully, contribute to now elevating the representation of the Korean community and language in the US.

My thanks also go to the Department of Linguistics and the Purdue community for their professional training and the warm and welcoming environment. I am also very grateful for the funding opportunities that enabled the complement of the graduate program, including the Department of Linguistics, School of Interdisciplinary Studies, College of Liberal Arts, Office of Interdisciplinary Graduate Programs, Graduate School, and Purdue University.

Now, I want to thank my family. First and foremost, I thank my parents, Mijin Chu and Jaimo Seo for their endless love and support for the completion of the program and my personal growth. I thank my younger brother, Kanghyeon, a good son to our parents in my absence in South Korea. I also thank my relatives who have cheered for my success. My full heartfelt gratitude goes to my two grandfathers who became angels during my program. Joonpyo Seo was my model as a good educator. God bless you in heaven.

Lastly but certainly not least, I am indebted to my partner Rena Naganuma and her family in Japan for their endless support and love. You are the only person in the world who has witnessed every success and failure of mine throughout my graduate program. I cannot thank you enough for your support and love.

Portions of the current thesis have been published in *Journal of Phonetics* (Seo & Dmitrieva, 2024).

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREIVATIONS

3AFC	Three-alternative forced-choice
95%CI	95% Confidence/Credible interval
CLI	Cross-linguistic influence
f0/F0	Fundamental frequency
HPD	Highest posterior density
HS	Heritage speaker
L1	First language
L1-i	L1-immersed
L2	Second language
LTI	Long-term immigrant
M	Mean
MINT Sprint	Multilingual Naming Test Sprint
ms	Millisecond
n	Number
pd	Probability of direction
RT	Reaction time
SD	Standard deviation
SE	Standard error
SLM	Speech Learning Model
SLM-r	The revised Speech Learning Model
VOT	Voice onset time

ABSTRACT

Upon acquiring or learning another language, cross-linguistic influence (CLI) is an inevitable phenomenon with which a bilingual speaker lives. One key aspect of CLI is its bidirectionality, flowing between both the first (L1) and second languages (L2) mutually affecting each other. However, investigations of L1 CLI on L2 have dominated previous bilingual studies, and despite the increasing amount of research on L2 CLI on L1, the phonetic and phonological domains remain relatively underexplored. The primary goal of the present study is to expand our understanding of the underlying mechanisms governing L2 CLI on L1 phonetics and phonology.

The present study investigates L2 CLI on L1 phonetics and phonology by examining both the speech perception and production of L1 sound categories among two different groups of bilinguals, Korean heritage speakers (HSs, *n* = 30) and long-term immigrants (LTIs, *n* = 27) group participants in the US, in comparison to L1(Korean)-immersed (L1-i) native speakers residing in South Korea (*n* = 30). Participants completed a series of three experimental tasks: (1) a three-alternative forced-choice (3AFC) identification task, (2) an AX discrimination task, and (3) a controlled reading paradigm task.

Experiment 1 (3AFC task) was conducted to investigate the extent and direction of L2 CLI in perceptual cue weighting to L1 speech categories. In this task, participants listened to a Korean word in each trial, potentially differing in the word-initial stop, and decided which word they heard from a real-word Korean minimal triplet /pul/ 'fire,' /p^hul/ 'grass,' and /p^*ul/ 'horn.' Specifically, the word-initial stop consisted of an eight-by-eight orthogonal voice onset time (VOT)–onset f0 continuum, created through a speech resynthesis technique. Based on the similarities and differences in the use of the two acoustic parameters between Korean (either onset f0 or VOT is a primary cue) and English stops (VOT is the primary cue), bilingual participants were expected to exhibit different cue-weighting patterns, as compared to L1-i speakers. The results from the mixed-effects logistic regression model analyses indicated that while HSs were less sensitive to the Korean primary cue, onset f0, compared to L1-i speakers—suggesting assimilation to L2 in the perceptual domain—LTIs exhibited greater sensitivity to this cue, indicating dissimilation from L2. It was also found that bilingual participants' Korean dominance significantly influenced their cue weighting in the perception of Korean stops.

Experiment 2 (AX discrimination task) was administered to assess participants' perceptual accuracy for L1 stop categories and the potential impact of L1 cue weighting, as estimated in Experiment 1, on their discrimination performance. Notably, the VOT in the stop stimuli used in the AX task were resynthesized to have a consistent VOT of 70 ms across all stimuli. This setup created a condition where participants had no choice but to rely solely on the onset f0 cue—the primary cue to the Korean lenis-aspirated stop contrast, rendering VOT, the primary cue for the voicing contrast in English stops, uninformative. The results from mixed-effects logistic regression models showed that HSs were significantly less accurate in discriminating their L1 stop categories without the VOT cue, while LTIs outperformed the L1-i speakers. That is, the LTI group, the most balanced group in terms of language dominance, had the highest accuracy in discriminating L1 contrasts among the participant groups. Furthermore, individual sensitivity to the onset f0 cue was found to be positively correlated with discrimination performance.

Experiment 3 (Controlled reading paradigm) aimed to examine L2 CLI on the implementation of acoustic parameters for L1 Korean stops, as well as the potential impact of proficiency and dominance on these parameters. Participants read aloud a list of minimal triplet stimuli differing in the word-initial stop within a carrier phrase. A machine-learning-based audio signal detection system was used to analyze the acoustic parameters, and Bayesian mixed-effects linear regression models, along with quadratic polynomial regression models, were implemented for statistical analysis of the processed data. The results of the production task mirrored the perception task (Experiment 1): HSs demonstrated assimilation to L2 via onset f0, while LTIs showed dissimilation, as compared to L1-i speakers. The analysis also revealed that the degree of bilingual balance in dominance and proficiency significantly influenced the implementation of onset f0, with more balanced bilinguals exhibiting greater category contrasts than less balanced bilinguals, regardless of whether they were Korean-dominant or English-dominant.

The findings from these experiments provide concrete evidence of L2 CLI in L1 phonetics and phonology. Importantly, the results demonstrate that not only the timing of L2 acquisition and the quantity and quality of L2 input but also the quality and quantity of L1 acquisition and bilingual balance contribute to the direction and the degree of L2 CLI in L1 speech. These findings align with the predictions of the revised Speech Learning Model (SLM-r, Flege & Bohn, 2021) and expand its scope of application to include both HSs and LTIs. In particular, the evidence of category assimilation and dissimilation lends support to the bidirectional CLI hypothesis proposed by SLM-r. To conclude, the present dissertation expands our understanding of the nature of L2 CLI in L1 phonetics and phonology in bilingual speakers.

1. INTRODUCTION

1.1 Cross-linguistic influence in bilingual speakers

According to Speech Learning Model and its revised version (Flege, 1987, 1995, 2003; Flege & Bohn, 2021), it is posited that bilingual speakers may produce second language (L2) sounds by assimilating them to similar first language (L1) sounds. This interaction is referred to as cross-linguistic influence (CLI) or bidirectional CLI. Although the term, "cross-linguistic influence", is often interchangeably used with cross-linguistic interference or as a hypernym of cross-linguistic interference, the current study makes a distinction between the two concepts in order to differentiate between long-term and short-term CLI (Seo & Olson, in press; Simonet, 2014). CLI refers to cognitive interactions between two languages that arise in bilinguals' long-term mental representations. For example, L1 transfer to L2 phonetics, such as L2 category assimilation (e.g., Flege, 1987), is an outcome of L1 CLI in L2 phonetics. On the other hand, cross-linguistic interference refers to short-term (or transient) interactions between two language systems, which manifests in certain contexts, such as code-switching, power mismatch between a speaker and an addressee, and more, which does not necessarily result in changes in the long-term mental representations. For instance, Seo and Olson (in press) examined L2 (English) vowel productions by advanced Korean–English bilinguals who can distinguish between English [æ] and its similar Korean vowel [e] when speaking only English or Korean. The study found evidence for short-term cross-linguistic interference in L2 phonetics by demonstrating that the contrast between the two different vowels diminishes during code-switching, compared to isolated contexts.

One key-aspect of CLI is its bidirectionality flowing between two languages, such that not only does the L1 impact an L2, but the L2 can also affect the L1 (Baker & Trofimovich, 2005; Flege & Bohn, 2021, among others). Despite a growing number of recent studies on L2 CLI on L1, most previous bilingual phonetic studies focused on L1 CLI on L2. Moreover, investigations of the morphosyntactic domain dominate the existing studies exploring L2 CLI on L1. For instance, Polinsky (2011) suggests that L1 grammar can deteriorate, resulting in L1 attrition, due to increased dominance in L2. However, relatively little research has examined the mechanisms that govern L2 CLI on L1 in the phonetic and phonological domains. Nevertheless, existing evidence lends support to the bidirectionality of CLI in speech. For instance, Chang (2012) demonstrated

L2 CLI in L1 vowels, which shifted towards similar L2 vowels during an L2 immersion program. In the perceptual domain, L2 CLI can even change cue weighting (relative importance in acoustic parameters to perceive a speech sound) for L1 categories such that a primary cue for similar L2 sounds is relied on by bilinguals to a greater extent than by monolingual speakers (Dmitrieva, 2019). Yet, less is known about the extent and the direction of L2 CLI in L1 speech perception and production. The present study aims to bridge this gap by examining L1 speech perception and production by heritage speakers (HSs) and long-term immigrants (LTIs) who follow unique developmental patterns of language acquisition.

1.2 Heritage speakers and long-term immigrants

HSs can be broadly defined as bilinguals who use an immigrant language at home, while speaking a different language, dominantly spoken in their country of residence (Polinsky, 2011; Valdés, 2005). HSs undergo a unique language development, whereby their later acquired L2 eventually becomes more dominant and proficient than their early acquired L1 (heritage language). For instance, the National Heritage Language Survey reports that Korean HSs in the US, who exclusively use their heritage language at home until the age of 5, show a sharp decrease in the use of the heritage language compared to English (L2) thereafter, with 98.7% of them becoming more dominant in English by the age of 18 (Carreira & Kagan, 2011). This switch in language dominance implies that HSs are expected to experience a strong degree of L2 CLI on their L1, making them an ideal population to study L2 CLI. However, previous studies have focused on examining the morphosyntactic domain in HSs, with phonetic and phonological domains remaining "among the least understood properties of heritage languages" (Polinsky, 2018, p. 162).

LTIs are another type of bilinguals who are assumed to experience a great degree of L2 CLI on the L1 since they have a prolonged period of residence in an L2-speaking country and presumably use the L2 dominantly. However, existing literature varies in terms of the definition of LTIs. For instance, the criteria for defining the period of residence in a foreign country to be considered LTIs ranges from a few years to more than eleven years (e.g., Dmitrieva, 2019: 11 years). In addition, the age of immigration to a foreign country, as it can serve as a proxy measure for L1 proficiency, also varies depending on the literature. Regarding the age of immigration, it is postulated that, the older the age of immigration, the more proficient an LTI's L1 is likely to be, given the amount of time they spent being immersed in their L1-speaking country. The current study defines LTI bilinguals as those who have resided in a foreign country (the US) for five years or longer and who immigrated at 18 years of age or older. These criteria were set to ensure that LTI bilinguals acquired and mastered their L1 before immigrating to the US as adult L2 learners, but with a significant amount of exposure to the L2 (English). In addition, this age limit excluded college students who might not necessarily become immigrants, which could affect their attitude towards English. Thus, LTI bilinguals differ from HSs not only in terms of the timing of L2 acquisition (i.e., HSs acquire L2 earlier than LTIs), but also in terms of the quality of L2 acquisition (i.e., formal education, social and professional contexts). Importantly, LTIs and HSs also differ in the quality and quantity of L1 input, with HSs' experience in L1 acquisition not being comparable to that of LTIs due to their immersion in an L2 speaking country during childhood. This assumption regarding the quality and quantity of L1 input of HSs and LTIs in the present study is limited to the extent to which the methods used to document and measure their linguistic backgrounds reflect these factors. The assumption that the quality of input for HSs is not comparable to that of LTIs is based on the fact that HSs' exposure to L1 is usually limited to their family or expatriate communities (Polinsky, 2018).

1.3 The current study and the outline of the dissertation

The overarching goal of the current study is to investigate the nature of L2 CLI in the phonetic and phonological domains. To achieve this goal, the study conducted three experiments, leveraging the differences in Korean and English laryngeal stops. These experiments aimed to examine the degree and the direction of L2 CLI in L1 perception and production by two groups of Korean–English bilinguals (HS and LTI groups) residing in the US, as compared to Korean-immersed (L1-i) native speakers residing in South Korea. The study corroborates the revised Speech Learning Model (SLM-r, Flege & Bohn, 2021), with the goal of contributing to informing and expanding our understanding of CLI in bilingual speakers. The structure of the dissertation is outlined below.

Chapter 2 provides a review of literature, including the outline of the relevant theoretical model, SLM-r (Flege & Bohn, 2021). The chapter also discusses previous studies on CLI in speech perception and production, and investigations of Korean and English stop categories.

Chapter 3 presents the designs, methodology, results, and interim discussion from a three-alternative forced-choice (3AFC) identification task and an AX discrimination task, designed to investigate CLI in speech perception.

Chapter 4 presents the designs, methodology, results, and interim discussion from the controlled reading task designed to investigate CLI in speech production.

Chapter 5 summarizes the results from the three experiments and discusses the implications of the current dissertation.

Chapter 6 provides the conclusion of the dissertation and future directions.

Appendices can be found after Chapter 6.

2. LITERATURE REVIEW

Portions of the current chapter have been published in *Journal of Phonetics:*

Seo, Y., & Dmitrieva, O. (2024). L2 cross-linguistic influence on L1 perception: Evidence from heritage speakers and long-term immigrants. *Journal of Phonetics*, *104*, 101314. DOI: https://doi.org/10.1016/j.wocn.2024.101314.

2.1 Introduction

Previous research on the linguistic aspects of bilingualism suggests that CLI in bilingual speech is all but inevitable. For example, a substantial body of literature provides evidence that the L1 impacts L2 speech in both the production and perception domains (Abrahamsson, 1999; Bergmann et al., 2016; Broselow et al., 1998; Caramazza et al., 1973; Dmitrieva et al., 2010; Eckman, 1987; Flege & Eefting, 1987a, 1987b; Flege et al., 2003; Guion, 2003; Hancin-Bhatt & Bhatt, 1997; Harada, 2003; Kong & Kang, 2023; Lang & Davidson, 2019; MacKay et al., 2001; Picard, 2002; Trofimovich et al., 2007). Similarly, a growing body of research explores CLI of L2 on L1 speech, contributing evidence for the bidirectional nature of cross-linguistic phonetic influence (Baker & Trofimovich, 2005; Chang, 2012; Chang & Mandock, 2019; Chang & Weiss-Cowie, 2021; Cheng, 2017; Flege, 1987; Flege et al., 1997; Flege & Eefting, 1987a; Guion, 2003; Harada, 2003; Kang & Guion, 2006; Kang & Nagy, 2016; Kong & Kang, 2023; Lord, 2008; Mack, 1990; Major, 1992, 1996).

Bidirectional CLI is among the major predictions posited by one of the most influential models of L2 speech, the Speech Learning Model (SLM, Flege, 1987, 1995, 2003) and the recently revised version, SLM-r (Flege & Bohn, 2021). This prediction derives from the core assumption of SLM, which remains intact in SLM-r, that L1 and L2 sounds exist as separate or composite L1/L2 phonetic categories in a shared mental space of a bilingual speaker and thus are expected to interact and mutually affect each other.

Another important assumption of SLM-r is that a strong mutual connection exists between speech production and speech perception. Although SLM-r relinquishes the idea that development of perceptual skills in L2 necessarily precedes and predicts the development of L2 production skills, it underscores the concurrent evolution of L2 perception and production, thus maintaining this mutual connection. A simple and logical extrapolation based on this premise is that a similar perception-production connection must be maintained in L1 as it undergoes changes under the effect of L2.

While the effect of L2 on L1 production has been relatively well documented (e.g., Baker & Trofimovich, 2005; Chang, 2012; Chang & Mandock, 2019; Chang & Weiss-Cowie, 2021; Cheng, 2017; Flege, 1987; Flege et al., 1997; Flege & Eefting, 1987a; Guion, 2003; Harada, 2003; Kang & Guion, 2006; Kang & Nagy, 2016; Kong & Kang, 2023; Lord, 2008; Mack, 1990; Major, 1992, 1996), the corresponding changes in the perception of L1 sound categories have not been given equal attention in the phonetics literature on CLI. The present work aims to bridge this gap by investigating the production and perception of L1 sound categories in populations whose L1 speech is expected to be subject to L2 CLI, thereby probing the extension of the theoretical postulates of SLM-r. A second contribution of the present work is the extension of predictions of SLM-r, developed primarily to account for L2 speech of language learners, to another population of bilinguals, specifically HSs.

The predictions of SLM have been tested predominantly on the populations of long-term, immersed L2 learners, who began acquiring the additional language in adulthood (late sequential adult bilinguals). However, applying the model to a diverse range of bilinguals allows for a more rigorous corroboration of its predictions regarding the nature and direction of CLI in light of the differences in timing of L2 acquisition, quantity and quality of L1 and L2 input, and language dominance among the bilingual populations.

Specifically, the present study investigates the perceptual identification and acoustic realization of lenis, aspirated, and fortis stops in Korean and discrimination of the lenis-aspirated contrast by adult Korean HSs residing in the United States of America (USA) and adult long-term immigrants (LTIs) from South Korea, in comparison to adult L1-immersed (L1-i) speakers of Korean (i.e., speakers who spent the majority of their lives in South Korea and learned English only through school education). The present study refers to the group residing in South Korea as L1-immersed instead of native or monolingual because all participants in the current study identify themselves as native speakers of Korean with knowledge of and exposure to English (Seo et al., 2022).

The main focus of this study is on determining whether the relative reliance on the two perceptual cues, voice onset time (VOT) and onset f0, which carry a different functional load in cueing laryngeal contrasts in English and Korean, differs across the three groups of participants.

In particular, the study aims to uncover any potential differences that would arise under the assumption that CLI from English affects bilinguals' production and perception of stop categories in their L1, Korean.

2.2 Revised Speech Learning Model (SLM-r)

Several models have been proposed in the field of L2 phonetics to account for the acquisition of L2 phonetics, notably the Perceptual Assimilation Model (PAM, Best, 1994, 1995), the Second Language Perception Model (L2LP, Escudero, 2005, 2009), and SLM (Flege, 1987, 1995, 2003). These models focus on different aspects of the acquisition of L2 speech. For instance, PAM addresses the issue of perceptual discrimination of L2 contrasts, suggesting that L2 contrast discrimination is affected by patterns of perceptual assimilation between the relevant L2 categories and perceptually similar L1 categories. PAM-L2 (Best & Tyler, 2007), proposes six possible types of category assimilation patterns between L1 and L2 sounds, such as two-category and single-category assimilation, which guide the patterns of L2 contrast discrimination. On the other hand, L2LP focuses on predicting the perception of similar L2 sounds. This model predicts that an L2 listener will use the optimal L1 perception to categorize L2 sounds. Specifically, the model posits that L2 learners are "optimal perceivers of their native language" and that L2 learners make use of their L1 categories when perceiving similar L2 sounds (Escudero, 2009, p. 184). For example, the model predicts and demonstrates a "full copy" of L2 vowels in L2 initial perception when L1 vowels similar to the target L2 vowels exist (e.g., Schwartz & Sprouse, 1996).

While both PAM and L2LP primarily address the perceptual aspects of L2 category acquisition, SLM extends its predictions to include both perception and production. This comprehensive approach makes SLM particularly relevant for the present study that investigates CLI in both speech perception and production. The fundamental assumption of SLM and SLM-r is that similar L1 and L2 categories coexist within the same cognitive space, jostling for a position (Flege, 1987, 1995, 2003; Flege & Bohn, 2021). According to the model, the ability to sever this perceptual link between target L2 and similar L1 categories is the key to the successful acquisition of L2 sounds. Relevant to one key assumption of the present study, the revised model proposes the "co-evolution" hypothesis predicting that "L2 segmental production and perception coevolve without precedence" (Flege & Bohn, 2021, pp. 28–29).

One of the advantages of SLM-r as a model of speech development in the context of L2 learning and bilingualism is the fact that it makes explicit predictions regarding the bidirectional CLI. According to the model, the proposed mechanism behind CLI is the creation of distinct phonetic categories for L2 sounds. One important condition that determines the likelihood and speed of L2 category creation is the amount of L2 input speakers receive. In the initial stage, having received little input, learners are not yet attuned to acoustic differences between L1 and L2 speech sounds. Consequently, they tend to create composite or merged L1/L2 categories, whereby an L2 sound and a perceptually similar L1 sound are grouped together. This leads to L2 sounds being pronounced in a manner similar to the corresponding L1 categories, demonstrating assimilatory CLI of L1 on L2, and vice versa. Empirical studies demonstrate that even beyond the beginner stage, L2 speakers often produce perceptually similar L1 and L2 categories as acoustically distinct but not identical to their monolingual renditions, primarily because categories continue to assimilate to each other (Bergmann et al., 2016; Caramazza et al., 1973; Dmitrieva et al., 2010; Flege & Eefting, 1987a, 1987b; Flege et al., 2003; Guion, 2003; Harada, 2003; Lang & Davidson, 2019; MacKay et al., 2001). This suggests a possibility that the link between related L1 and L2 categories can persist long into the L2 acquisition process.

SLM-r also posits that once a truly separate L2 category is established, dissimilatory CLI is expected rather than assimilatory, as a bilingual speaker strives to minimize the overlap between acoustically adjacent but distinct phonetic categories. However, despite abundant research on the topic, few empirical examples of this type of phonetic dissimilation have been documented (Flege & Eefting, 1988; Flege et al., 2003; Mack, 1990; Yusa et al., 2010), suggesting that either L2 learners rarely create completely separate L2 categories or that cross-linguistic category dissimilation is not a necessary consequence of L2 category creation. In light of these observations, the study hypothesized that the likeliest outcome for L2 speakers, such as the immigrant population tested in the present study, is assimilatory CLI of L2 on L1.

While SLM-r primarily targets adult L2 learners, its assumptions and predictions can be extended relatively straightforwardly to other types of bilinguals. The present study proposes to apply them to HSs, broadly defined as bilingual speakers who use their L1 at home while using predominantly their L2 in broader social contexts, such as school, work, and socialization (Valdés, 2005). Unlike sequential adult bilinguals, HSs are typically exposed to both of their languages in early childhood, although individual details can vary (Polinsky, 2018, p. 9). Exposure to the first, family language, typically precedes exposure to the second, community language, and both types of input are immersive, natural, and authentic, being provided by native speakers of the corresponding languages. While heritage language often dominates during the early years, a dominance reversal typically occurs around the age when primary schooling begins, as a result of which HSs become and remain majority-language dominant (Carreira & Kagan, 2011; Polinsky, 2018). Thus, the key differences of HSs from adult L2 learners are the significantly earlier exposure to L2 as well as greater cumulative L2 exposure and ultimate L2 dominance. An additional important consideration concerns the fact that HSs' L1 input is often provided by a limited number of speakers—members of the family and sometimes members of the extended expatriate community. This input does not compare to the quantity, diversity, and dynamics of L1 input received by those growing up surrounded by L1 spoken as the majority language (Chang, 2021).

Applying the predictions of SLM-r to HSs, one may conclude that given their extensive experience with and dominance in an L2, HSs have had plenty of opportunities to develop distinct L2 speech categories, and, therefore, only dissimilatory CLI could be expected. However, SLM-r proposes that the likelihood of forming separate L2 categories depends not only on the quantity and quality of L2 input but also on the nature of existing L1 phonetic representations. For example, less precision in individual L1 phonetic categories, measured as the amount of acoustic variability in production, is predicted by SLM-r to counteract the creation of distinct and well-defined L2 categories, meaning that the link between perceptually similar L1 and L2 phonetic categories may persist, giving rise to assimilatory CLI.

While L1 category precision in SLM-r is an endogenous factor unique to individuals, it can also be a characteristic of groups of speakers. Individuals may come to share properties of phonetic representations due to shared linguistic experiences, such as the quality and quantity of L1 input. For example, adults in general exhibit greater category precision in their L1 than children do (Lee et al., 1999). Similarly, HSs, who have had limited L1 input, are likely to have less robust L1 categories than their peers who grew up in L1-immersed environments. Thus, extending SLM-r predictions based on individual characteristics to the group of HSs, assimilatory CLI can be expected for HSs due to relatively imprecise L1 phonetic categories.

The final condition determining the likelihood of L2 category creation is the perceived degree of similarity of an L2 sound to the closest L1 sound. SLM-r posits that the more dissimilar

L1 and L2 sounds are, the stronger the likelihood of rapid category formation, precluding at least the assimilatory CLI (Flege & Bohn, 2021). Parallel laryngeal categories across languages (e.g., Spanish voiceless stops with a short-lag VOT and English voiceless stops with a long-lag VOT) have been widely used as examples of sound categories sufficiently similar to be linked into composite L1-L2 categories and trigger CLI (French-English: Caramazza et al., 1973; Korean-English: Schmidt, 1996; Spanish-English: Flege & Eefting, 1987b, among others). The present study aims to corroborate the SLM-r prediction regarding the likelihood and the direction of CLI by examining properties of acoustic realization, patterns of perceptual cue weighting, and perceptual discrimination of Korean stops by Korean–English bilinguals.

2.3 CLI in L1 speech production

Previous literature provides evidence of L2 CLI on L1 speech production, suggesting that L1 phonetics and phonology are not as fixed but rather flexible and malleable systems (Baker & Trofimovich, 2005; Chang, 2012; Chang & Mandock, 2019; Chang & Weiss-Cowie, 2021; Cheng, 2017; Flege, 1987; Flege et al., 1997; Flege & Eefting, 1987a; Guion, 2003; Harada, 2003; Kang & Guion, 2006; Kang & Nagy, 2016; Lord, 2008; Mack, 1990; Major, 1992, 1996). Notably, while these studies suggested evidence of L2 CLI on L1 speech, its direction and effect varied across studies. One of the most common findings among these studies is the pattern of assimilation of L1 speech categories to similar L2 sounds, which is interpreted in SLM-r as a consequence of forming "equivalence classes" (Flege & Bohn, 2021). For instance, Major (1992) investigated L2 (Portuguese) CLI on L1 (English) among adult L2 learners. The study examined VOT in laryngeal stops in both L1 and L2 produced by American adult bilinguals in Brazil with the length of residence of at least 12 years, compared to monolingual control groups for each language. The study found that while the VOT of Portuguese (L2) stops produced by the bilingual group became significantly longer compared to that of the Portuguese control group, the VOT of their English (L1) stops was also characterized by a significant shortening compared to the English control group. This finding provides a representative case of category assimilation between L1 and L2 via the mechanisms of bidirectional CLI.

With relevance to the present study, previous studies found evidence of L2 (English) CLI on L1 (Korean) phonetics in speech production (Chang & Mandock, 2019; Cheng, 2017; Kang & Nagy, 2016; Lee & Iverson, 2012; Oh, 2019; Oh & Daland, 2011; Yoon, 2015). A common finding in the literature is that Korean–English bilinguals tend to rely on onset f0, the primary cue to the lenis-aspirated stop contrast in Korean, to a lesser degree, compared to Korean-immersed speakers in South Korea. For instance, Kang and Nagy (2016) investigated the variation in the implementation of the two acoustic cues, VOT and onset f0, in the production of Korean laryngeal stops among first- and second-generation Korean immigrants in Canada in comparison to Korean-immersed speakers in South Korea. The study found a significant decrease in the use of onset f0 in the production of Korean stop contrasts among Korean–English bilinguals of both generations, compared to Korean-immersed speakers. However, they showed a relatively greater reliance on VOT than on onset f0 in the production of Korean stops, although there were generation- and gender-related differences. More specifically, the HS group exhibited increased VOT values compared to the control group in both lenis and aspirated stops. The study interprets this finding as a manifestation of CLI from English, leading to prolonged VOT for both lenis and aspirated stops, assimilating to English voiceless stops. These findings suggest L2 CLI from English in which onset f0 plays a secondary role in cueing the voicing contrast in laryngeal stops, thereby decreasing its weight even in participants' L1, Korean, providing evidence of assimilation.

A similar finding was reported in Cheng (2017), which examined the production of Korean and English stops produced by different generations of Koreans in the US in comparison to Korean-immersed speakers in South Korea. Specifically, the study found distinct behavior in the use of onset f0 in the production of Korean stops among second-generation Korean Americans (HSs in the definition of the present study), with their reliance on this tonal cue significantly decreasing compared to the first- and 1.5-generation Koreans in the US as well as the Korean-immersed speakers. The study also pointed out that there was no significant difference between male and female speakers within the second-generation group, indicating that the observed reduction in the weight of onset f0 was consistent across genders. Regarding VOT, the HS group participants showed prolonged VOTs for both lenis and aspirated stops, compared to the other groups, which is interpreted as "a result of contact with English" (Cheng, 2017, p. 301). This interpretation is compatible with Kang and Nagy (2016), providing additional evidence of category assimilation in Korean lenis and aspirated stops to their similar English voiceless stops.

A more recent study examining the three-way laryngeal contrast in Korean stops produced by Korean HSs in the US found consistent patterns in the application of VOT and onset f0 as acoustic correlates of the stop contrasts (Chang & Mandock, 2019). That is, HSs tended to rely less on onset f0 and more on VOT, compared to Korean-immersed speakers, corroborating the influence of L2 (English) CLI on the phonetic system of the L1 (Korean). However, the study also found that HSs produced all types of Korean stops with significantly higher VOT and onset f0 compared to both Korean-immersed and English-immersed speakers, which is distinct from the findings of other previous studies (e.g., Cheng, 2017; Kang & Nagy, 2016). The study suggested that enhanced VOT and onset f0 within the HS group may be attributed to a clear speech effect. The enhancement effect is not an uncommon observation in the examination of VOT and onset f0 in previous studies (e.g., Jung, 2023; Kang & Guion, 2008). Regarding the predictions of the present study on L2 CLI on L1 Korean in the production of Korean stops, it is expected that both HSs and LTIs, who are immersed in an L2-speaking environment, will exhibit a pattern of assimilation in their Korean stops due to the L2 (English) influence.

2.4 CLI in L1 speech perception

Comparatively few studies have investigated the issue of CLI in L1 speech perception, in contrast to speech production (Caramazza et al., 1973; Chang & Ahn, 2023; Dmitrieva, 2019; Elman et al., 1977; Garcia-Sierra et al., 2009; Lee-Ellis, 2012; Llanos et al., 2013; Williams, 1977). Yet, the assumption of a strong intra-linguistic link between production and perception presupposes that changes in one domain should be mirrored by corresponding modifications in the other. Therefore, in the most general terms, if we expect CLI-related changes in the production of certain L1 sound categories, we should also expect relevant changes in their perception.

As it relates to the specifics, L2-driven changes in speech perception can manifest in a variety of ways. Traditionally, bilingual perceptual abilities have often been assessed via evaluating the location of the category boundary in the category identification paradigm, or percent correct discrimination in the discrimination paradigm (e.g., Lee-Ellis, 2012; Oh et al., 2003). These approaches have not consistently revealed strong evidence for CLI in L1 perception. One explanation is that these methodologies may not be sufficiently sensitive to detect subtle, subphonemic modifications in the perceptual representations of L1 categories due to CLI. For instance, several studies examined perceptual accuracy in the discrimination of L1 categories by Korean HSs. The findings demonstrated a high degree of accuracy, comparable to that of L1-i participants, which has been interpreted as evidence against CLI on L1 perception in HSs (Chang, 2016; Lee-Ellis, 2012; Oh et al., 2003; Seo et al., 2022).

An alternative, a possibly more sensitive measure of perceptual CLI, can be found in a cue-weighting approach. If a given acoustic dimension could be expected to turn into a more (or less) reliable and pronounced production marker of a sound category as a result of CLI, correspondingly, it should come to play a more determinative role in identifying and discriminating the category perceptually. Previous findings suggest that it is indeed the case. For instance, data reported by Dmitrieva et al. (2010) indicated that, under the effect of CLI from English, where vowel duration is an important cue for the word-final voicing contrast, vowel duration became a stronger correlate of final voicing in Russian speech of Russian-English bilinguals. Later, Dmitrieva (2019) provided evidence for L2 CLI on L1 perception by examining perceptual cue weighting in Russian LTIs in the USA. The study investigated whether Russian LTIs, who had resided in the USA for an average of ten years, differed from Russian and English monolingual speakers in their weighting of two acoustic cues – vowel duration and glottal pulsing – when identifying Russian and English voiced and voiceless stops. The results indicated that in the perception of Russian word-final stops, Russian LTIs assigned a greater weight to vowel duration than Russian monolingual speakers, who were more sensitive to glottal pulsing. There were, however, no significant differences between English monolinguals and Russian LTIs in using vowel duration while identifying English word-final stops.

Although there has been no clear evidence as to the perceptual behavior of the target population, LTI and heritage Korean–English bilinguals, clues to the inquiry about CLI in perception can be found in prior production studies. Specifically, Korean–English bilinguals, including HSs, have been reported to demonstrate evidence of English CLI on Korean production of laryngeal categories. For instance, bilinguals were shown to produce greater VOT differences and smaller onset f0 differences between Korean lenis and aspirated stops compared to Korean monolinguals (Cheng, 2017; Kang & Nagy, 2016; Lee & Iverson, 2012; Oh, 2019; Oh & Daland, 2011; Yoon, 2015), which suggests CLI from English, where VOT is a more important cue for voicing distinctions. It is reasonable to expect that bilinguals' reliance on onset f0 would also decrease while their reliance on VOT would increase in the perception of the same contrasts in Korean, as a result of influence from English.

The present study pursues this line of inquiry, applying the cue-weighting approach, in order to determine the extent to which Korean–English bilinguals exemplify the CLI effect in the perception of Korean sound categories. The study also explores whether individual and group cue-weighting patterns have implications for category discrimination scenarios that require utilization of a specific cue.

2.5 Korean and English laryngeal stop categories

The current study leverages the differences and similarities in laryngeal stop categories between Korean and English in investigating L2 (English) CLI on L1 (Korean) perception. As shown in Table 2.1, stop consonants in Korean exhibit a unique three-way laryngeal contrast. In the word-initial position, the contrast can be distinguished by two acoustic parameters: VOT (the period between the release of a stop and the onset of vocal fold vibration of a following segment) and onset f0 (the fundamental frequency at the onset of the vowel following the stop). The three Korean stop categories include (1) lenis /p, t, k/ (also referred to as plain or lax), (2) aspirated /p^h, t^h, k^h/, and (3) fortis stops /p^*, t^*, k^*/ (also referred to as tense) (/pul/ 'fire', /p^hul/ 'grass', and /p^*ul/ 'horn'). While fortis stops can also be distinguished from lenis and aspirated stops via lower H1−H2, the amplitude difference between the first and second harmonic (Cho et al., 2002), the present study focuses on VOT and onset f0 because the main research questions are centered on the difference between lenis and aspirated stops, both of which are characterized by similarly breathy phonation.

Production studies, including corpus research, have investigated this three-way distinction in speech, particularly focusing on the Seoul-Gyeonggi variety (Bang et al., 2018; Cho et al., 2002; Han & Weitzman, 1970; Hardcastle, 1973; Kagaya, 1974; Kang, 2014; Kang & Guion, 2006, 2008; Kang & Nagy, 2016; Kim, 1965; Kong et al., 2011; Lisker & Abramson, 1964; Silva, 2006). Findings showed that lenis stops were distinguished from aspirated stops primarily via onset f0 in Seoul-Gyeonggi Korean, with aspirated stops characterized by greater onset f0 than lenis stops, while both types of stops were realized with a long lag VOT. Several studies documented the prominent role of onset f0 in implementing the distinction between lenis and aspirated stops, particularly among young speakers (Bang et al., 2018; Kang, 2014; Kang & Guion, 2006, 2008; Kim & Jongman, 2022; Kong et al., 2011; Lee et al., 2020; Lee & Jongman, 2012, 2019; Schertz et al., 2015; Silva, 2006). A recent study provides additional evidence that the onset f0 difference between the two types of stops is most pronounced in the phrase-initial and prominent (focused) condition (Choi et al., 2020). The lenis-fortis stop contrast was found to be implemented via both VOT and onset f0, with lenis stops (long-lag VOT, low f0) having greater VOT and lower onset f0 values, compared to fortis stops (short-lag VOT, high f0). Fortis and aspirated stops were shown to be distinguished primarily via VOT, with fortis stops being characterized by shorter VOT than aspirated stops.

Perception studies reported corresponding patterns in the use of the acoustic cues, onset f0 and VOT, in the identification of the three-way contrast in Korean stops (Kim et al., 2002; Kong, 2012; Kong et al., 2011, 2022; Kong & Kang, 2023; Kong & Yoon, 2013; Lee et al., 2013; Schertz et al., 2015). Onset f0 was reported to play a dominant role in cueing the lenis-aspirated stop contrast, especially among young listeners of the Seoul-Gyeonggi variety. For instance, Kong et al. (2022) conducted a three-alternative forced-choice identification task (3AFC) to investigate the role of dialectal variety and age (elementary, high school, and university students) in Korean speakers' cue weighting of onset f0 and VOT for the three types of stops. They found that Seoul listeners weighed onset f0 over VOT for the lenis-aspirated stop contrast.

Table 2.1. Korean and English stops and corresponding acoustic parameters.

	Korean			English	
Acoustic cues	Aspirated	Lenis	Fortis	Voiceless	Voiced
VOT	long	long	short	long	short
Onset f0	high	low	high	high	low

In contrast to Korean, English exemplifies a more typical two-way laryngeal contrast. Prior studies examining the production of English stops in the word-initial position documented that voiced and voiceless stops are differentiated robustly via VOT, with voiceless stops characterized by a long-lag VOT and voiced stops by a short-lag VOT, although prevoicing is also relatively common in word-initial voiced stops (Abramson & Lisker, 1985; House & Fairbanks, 1953; Ohde, 1984; Shultz et al., 2012, among others). Onset f0 also covaries with voicing in English: voiced stops, including short-lag and prevoiced ones, are characterized by lower onset f0 than voiceless stops (Dmitrieva et al., 2015; House & Fairbanks, 1953; Ohde, 1984).

Perceptual voicing identification studies on English suggested that VOT functioned as a primary cue to the two-way laryngeal contrast, with onset f0 being a secondary cue (Abramson & Lisker, 1985; Llanos et al., 2013; Shultz et al., 2012; Whalen et al., 1993, for studies using visual analogue scaling (VAS) to investigate categorization gradiency and cue integration, see Kapnoula et al., 2017; Kong & Edwards, 2016).

Therefore, both Korean and English use VOT and onset f0 as acoustic cues to the laryngeal contrasts in the production and perception of stops. Despite the differences in cue primacy, laryngeal categories across the two languages are perceived as considerably similar by speakers of both languages. Previous research examining perceptual assimilation between Korean and English stops showed that Korean speakers perceive English voiceless stops as similar to Korean aspirated stops, while English voiced stops are assimilated to either lenis or fortis stops (Kang & Guion, 2006; Park & de Jong, 2008; Schmidt, 1996).

Based on the assumptions of SLM-r (Flege & Bohn, 2021), it can be predicted that certain conditions, such as similar L1 and L2 sound categories, low L2 input, and/or insufficiently robust L1 categories (e.g., due to limited L1 input), may lead to the formation of composite categories between corresponding L1 and L2 categories (Baker & Trofimovich, 2005; Chang, 2012; Chang & Mandock, 2019; Cheng, 2017; Flege, 1987; Flege et al., 1997; Flege & Eefting, 1987b; Guion, 2003; Harada, 2003; Kang & Guion, 2006; Kang & Nagy, 2016; Lord, 2008; Mack, 1990; Major, 1992, 1996). The experimental designs of the present study align with these conditions, as Korean and English laryngeal stops share sufficiently similar acoustic characteristics. Moreover, LTIs may not have established separate L2 categories due to insufficient L2 input, and HSs' L1 categories may be insufficiently robust due to limited L1 input. Therefore, the study predicts that Korean–English bilinguals will exhibit bidirectional assimilatory changes in the perception of stop categories in each language. Specifically, with respect to CLI from English (L2) on Korean (L1), the study expects to observe a greater reliance on VOT and a less pronounced reliance on onset f0 cue in the perception of Korean stops, as a result of CLI from English, given that this is the pattern characteristic of the perception of English stops.

2.6 Perception experiments

2.6.1 3AFC identification task

The purpose of the 3AFC task is to investigate CLI in L1 speech perception by examining perceptual cue weighting in the identification of the three laryngeal stops in Korean among Korean HSs and LTIs in the USA, compared to L1-i speakers residing in South Korea. Cue weighting is the result of assigning relative importance to acoustic parameters as cues to a phonetic category (Holt & Lotto, 2006). According to attention-to-dimension warping models from which the cue-weighting approach is derived, the process of cue weighting can be linked to the attentional operations in the mental space, where learning of phonetic categories is treated as a modification of the perceptual space through selective attention mechanisms (Francis et al., 2000; Francis & Nusbaum, 2002). In the cue-weighting approach, focused attention on a specific sensory dimension (temporal or spectral), that is, an increased weight on an acoustic parameter (e.g., VOT), leads to a salient contrast between categories (voiced vs. voiceless stops), while withdrawn attention from a dimension compresses the space, diminishing the contrast. Cue weighting is considered a crucial element in the mechanism of perceptual learning (Gibson, 1969; Goldstone, 1998).

Previous studies have shown that languages differed in how much weight their speakers assigned to specific acoustic correlates in categorizing speech sounds and that cue weighting was pliable and subject to change not only in an L2 (Hazan & Boulakia, 1993; Kajouj & Kager, 2019; Kim & Tremblay, 2021; Kondaurova & Francis, 2010; Kong & Kang, 2023; Kong & Yoon, 2013; MacKain et al., 1981; Yamada & Tohkura, 1992) but also in the L1 (Dmitrieva, 2019; Idemaru & Holt, 2011; Kong & Kang, 2023). For instance, Kong and Kang (2023) found that Korean–English bilinguals adjusted their reliance on acoustic cues, onset f0 and VOT, for language-specific contrasts (Korean stop contrast vs. English stop contrast). Specifically, when perceiving Korean stops, participants assigned a greater weight to the onset f0 cue than to the VOT cue, in contrast to the perception of English contrasts. Therefore, cue weighting could be susceptible to CLI, as category boundaries and other aspects of speech perception are. Examining perceptual cue weighting to L1 stops can elucidate the perceptual strategies bilinguals employ when making categorical decisions.

The study seeks to determine the degree to which the amount and timing of L2 experience, as well as the quality and quantity of L1 experience, which differ for HSs and LTIs, affect the likelihood and direction of L2 CLI on L1 speech perception. SLM-r hypothesized that these factors play an important role in the CLI process, via the mechanism of L2 category creation. LTIs are expected to be subject to assimilatory CLI because their L2 input began in adulthood and may not have been sufficient for establishing separate L2 categories. HSs, on the other hand, are expected to be susceptible to assimilatory CLI because their L1 categories may not be sufficiently robust, precluding the creation of separate L2 categories.

2.6.2 Speeded AX discrimination task

In addition to the 3AFC task, implemented in order to investigate CLI in perceptual cue weighting, the study also administered a discrimination paradigm, using acoustically manipulated lenis-aspirated stop stimuli. The stimuli were resynthesized, such that only onset f0, but not VOT, could be used to distinguish the two categories. If the hypothesis is correct—that CLI from English leads to a decrease in reliance on onset f0 as a cue to laryngeal categories in Korean—the ability to use this cue in perceptual discrimination should also be diminished. By removing VOT as a cue in the stimuli, the study created a condition in which we can focus on examining CLI on the reliance on onset f0 in the perceptual discrimination of the lenis-aspirated stop contrast. Finally, the present study examines whether individual listeners' reliance on primary cue measured in the identification task predicts their discrimination performance.

2.6.3 Research questions and hypotheses

Research questions and corresponding hypotheses of the perception experiments are as follows:

- RQ1. To what extent do Korean–English bilinguals (HSs and LTIs) in the USA differ in their perceptual cue weighting to the three-way laryngeal contrast in Korean stops from L1-i speakers residing in South Korea?
- RQ2. To what extent can HSs and LTIs distinguish between lenis and aspirated stops in an AX discrimination task when they must rely solely on onset f0?
- RQ3. To what extent is participants' auditory sensitivity to onset f0 as a cue to the lenis-aspirated stop contrast in a 3AFC task related to their perceptual accuracy in discriminating lenis and aspirated stops in an AX discrimination task?
- H#1. Given the more pronounced role that VOT plays in the perception of laryngeal categories in English, compared to onset f0, the study expects that both groups of Korean–English bilinguals will demonstrate a greater reliance on VOT and a lesser reliance on f0 in identifying Korean laryngeal categories, compared to L1-i Koreans.
- H#2. With VOT unavailable as a cue, the study predicts that both HSs and LTIs will be less accurate in their discrimination of the lenis and aspirated stops compared to L1-i speakers.

This prediction is based on the assumption that these groups of speakers are less sensitive to onset f0 as a result of CLI from English.

- H#3. The study hypothesizes that participants' weighting of onset f0 in the identification task will be a significant predictor of their discriminatory performance in discriminating lenis and aspirated stops.

2.7 **Production experiment**

2.7.1 Controlled reading task

The aim of the production portion of the study is to substantiate the SLM-r predictions regarding bidirectional CLI in speech production. By focusing on VOT and onset F0 in L1 (Korean) stop categories, known to be susceptible to CLI from English stop categories, the study compares the acoustic patterns of HSs and LTIs in the US with those of L1-i speakers in South Korea. The primary goal is to unravel how and to what extent bilingual dominance and proficiency influence CLI on L1 phonetic categories, particularly among bilingual speakers with a spectrum of language dominance from English (L2) dominance (HSs) to more balanced bilingualism (LTIs).

2.7.2 Research questions and hypotheses

- RQ1. To what extent do HSs and LTI speakers differ in their use of VOT in the production of Korean stops, compared to L1-i speakers?
- RQ2. To what extent do HSs and LTI speakers differ from L1-i speakers in implementing onset F0 in Korean stops?
- RQ3. In what ways and to what extent does bilingual dominance modulate the CLI in the realization of VOT and onset F0 in Korean stops among bilingual speakers?
- H#1. While VOT is a secondary cue for the Korean lenis-aspirated stop contrast, it is the primary cue in the voicing contrast in English stops. Therefore, it is hypothesized that HSs and LTIs with a high degree of English proficiency and dominance will exhibit a more pronounced VOT distinction among Korean stop categories compared to L1-i speakers.

- H#2. Onset F0 is a secondary cue in the distinction of English stops, while being the primary cue for the Korean lenis-aspirated stop. Thus, HSs and LTIs are anticipated to demonstrate a less pronounced onset F0 contrast between Korean stop categories relative to L1-i speakers.
- H#3. It is hypothesized that both HSs and LTIs will show patterns of assimilation via both VOT and onset f0 in the production of Korean stops, with bilingual dominance modulating the degree of assimilation.

3. CROSS-LINGUISTIC INFLUENCE IN L1 PERCEPTION

Portions of the current chapter have been published in *Journal of Phonetics:*

Seo, Y., & Dmitrieva, O. (2024). L2 cross-linguistic influence on L1 perception: Evidence from heritage speakers and long-term immigrants. *Journal of Phonetics*, *104*, 101314. DOI: https://doi.org/10.1016/j.wocn.2024.101314.

3.1 Methods

3.1.1 Participants

A total of 86 participants took part in the study, divided into three groups: the HS group (*n* = 30), the LTI group (*n* = 26), and the L1-i group (*n* = 30; baseline group). Two participants were excluded from data analyses, one from the HS group and one from the L1-i group, due to voluntary withdrawal from participation, resulting in a total of 84 participants. HS and LTI participants were recruited in the states of Illinois and Indiana, including the Chicago metropolitan area. The HS group consisted of 29 second-generation Korean Americans (Age: *M* = 26.3, range = 20 – 50, SD = 6.7), born and raised in the USA. These second-generation participants had first-generation Korean immigrant parents on both sides. Importantly, their parents, born post-1950, would have been exposed to the sound change in the lenis-aspirated stop contrast prior to immigration, affording HS participants early childhood exposure to the onset f0-based Korean stop distinction (Bang et al., 2018). Their initial age of immersion in English as an L2 was 3 years on average, coinciding with the start of preschool. The LTI group was composed of 26 participants (Age: *M* = 39.5, range = 28 – 50, SD = 6.6) who were born and raised in South Korea. LTI participants immigrated to the USA after the age of 18, with an average length of residence of 11.6 years (*M* = 11.6, SD = 6.6). This suggests that LTI participants had sufficient time to acquire and use their L1 (Korean) before becoming immersed in English. Therefore, LTIs in this study were not used as a baseline for HSs but as a distinct participant group. While our hypotheses posit that any differences between HSs and LTIs are due to distinct patterns of L1 and L2 acquisition and exposure, given this limitation, it is possible that some differences could be due to the generational gap between LTIs and HSs' parents. The L1-i group included 29 Korean-immersed native speakers who were recruited in the Seoul-Gyeonggi areas of South Korea (Age: *M* = 28.1, range = 19 – 40, SD = 6.5).

The rationale for recruiting L1-i participants from these specific regions is based on the evidence that the sound change relevant for the identification and discrimination of Korean stops first emerged and spread in the Seoul-Gyeonggi areas (Bang et al., 2018; Han & Weitzman, 1970; Lee et al., 2020; Silva, 2006). Moreover, recent studies showed that speakers in these areas rely most strongly on the onset f0 cue in the perception of lenis and aspirated stops (Kim, 2023; Kong et al., 2022; Lee et al., 2013). The study aimed to recruit individuals without extensive L2 experience (e.g., a short visit to an English-speaking country), although a certain degree of exposure to English was unavoidable to the mandatory English education that starts as early as the age of eight in South Korea (Korea Ministry of Education, 2015).

All participants completed a Bilingual Language Profile questionnaire (BLP; Birdsong et al., 2012; see also Olson, 2023 for a discussion on the reliability of BLP in speech science) (Appendix D) in their preferred language, either Korean or English, as a means to document their linguistic background and quantify their language dominance. The BLP questionnaire was used because it was designed to assess the quality and frequency of L1 and L2 use and exposure in various contexts, such as home, work, and social settings. The BLP provides relative dominance scores ranging from −218 (English-dominant) to 218 (Korean-dominant), calculated by summing equally weighted scores in the four subcategories, including language history, use, proficiency, and attitudes. In addition to this subjective measure, the present study employed an objective measure of bilingual proficiency and dominance, using a modified version of the Multilingual Naming Test Sprint (MINT Sprint; see Garcia & Gollan, 2021 for more information). In this test, participants named pictures shown on a computer screen in Korean and English. The MINT Sprint provides a proxy measure of bilingual proficiency and dominance with a maximum score of 80 in each language. Table 3.1. summarizes participants' background information, and Figure 3.1 and Figure 3.2 show the aggregated BLP and MINT Sprint scores by group, with pairwise comparison statistics (*t*-tests) where applicable.

As depicted in Figure 3.1, the BLP score indicated that the HS group was moderately English-dominant, as well as the most English-dominant among the participant groups ($M = -81.6$, SD = 28.4), with their BLP scores being significantly different from those of other groups (L1-i: $t(58.0) = 32.1$, $p < .001$; LTI: $t(37.0) = -12.2$, $p < .001$). Consistent with the BLP scores, the MINT Sprint results, as shown in Figure 3.2, confirmed that HSs were English-dominant speakers, as their scores were significantly higher in English ($M = 70.7$, SD = 4.3) than in Korean ($M = 32$, SD

= 12.5) ($t(33.4) = 15.4$, $p < .001$). Besides, their Korean score in the MINT Sprint was significantly lower than those of other groups (L1-i: $t(38.2) = -12.8$, $p < .001$; LTI: $t(46.8) = -11.4$, $p < .001$), while their English score was significantly higher than the other groups (L1-i: $t(37.5) = 15.2$, $p < .001$; LTI: $t(30.3) = 8.2$, $p < .001$). Previous studies on heritage language showed that HSs were highly variable in their heritage language dominance and proficiency (see, Polinsky, 2018, for a discussion on linguistic variability of HSs). The results of the MINT Sprint demonstrate that Korean HSs also exhibited considerable individual variability in terms of their Korean proficiency, with some participants scoring in the range of L1-i speakers and others situated at the lowest end of the proficiency scale.

LTI participants showed moderate dominance in Korean in their BLP score, situated in-between the HS and L1-i groups (M = 60.3, SD = 53.0) (L1-i: $t(37.3) = -8.1$, $p < .001$). The results of the MINT Sprint also indicated that LTI participants were Korean-dominant, although their English score was almost equally high, suggesting a relatively balanced proficiency between English (M = 54.3, SD = 8.8) and Korean (M = 63.4, SD = 7.4). When comparing the difference in the proficiency scores between Korean and English among the three groups, it was found that the LTI group had the smallest difference in proficiency scores between the two languages (M = 9.1, SD = 12.8). When compared to the L1-i group, LTI participants were not significantly different in their MINT Sprint score for Korean ($p = .8$), but their English score was significantly higher than that of L1-i participants (L1-i: $t(51.0) = 6.2$, $p < .001$).

The BLP questionnaire confirmed that L1-i participants were Korean-dominant, indicated by the highest BLP score among the three groups (M = 154.9, SD = 28.8). In addition, their Korean MINT Sprint score (M = 63.9, SD = 5.6), the highest among the groups, was significantly higher than their English MINT Sprint score (M = 37.4, SD = 11.5) ($t(40.4) = -11.2$, $p < .001$), the lowest among the groups.

Table 3.1. Participants' linguistic backgrounds.

	Group		
	HS	LTI	L1-i
Mean age (year)	26.3 (6.7)	39.5 (6.6)	28.1 (6.5)
Place of birth	US	South Korea	South Korea
Start of L2 immersion	Early childhood (Mean age = 3.0)	Adulthood (Mean age = 28.7)	Not applicable
BLP score	−81.6 (28.4)	60.3 (53.0)	154.9 (28.8)
MINT Sprint English	70.7 (4.3)	54.4 (8.8)	37.3 (11.5)
MINT Sprint Korean	32.0 (12.5)	63.4 (7.4)	63.9 (5.5)

NOTE: Values in parentheses are standard deviations (*SDs*).

Figure 3.1. BLP score by group.

Negative values indicate English dominance, while positive values indicate Korean dominance. Each box in the boxplot displays the interquartile range of scores for a group, with the horizontal line within the box denoting the median score. The vertical lines outside each box (also known as whiskers) extend from the first and third quartiles (hinges) to the largest value no further than the value of 1.5 times the interquartile range from each hinge. The colored dots represent individual data points, with each dot corresponding to a participant's BLP score within each group. The asterisk indicates statistical significance for each comparison between groups.

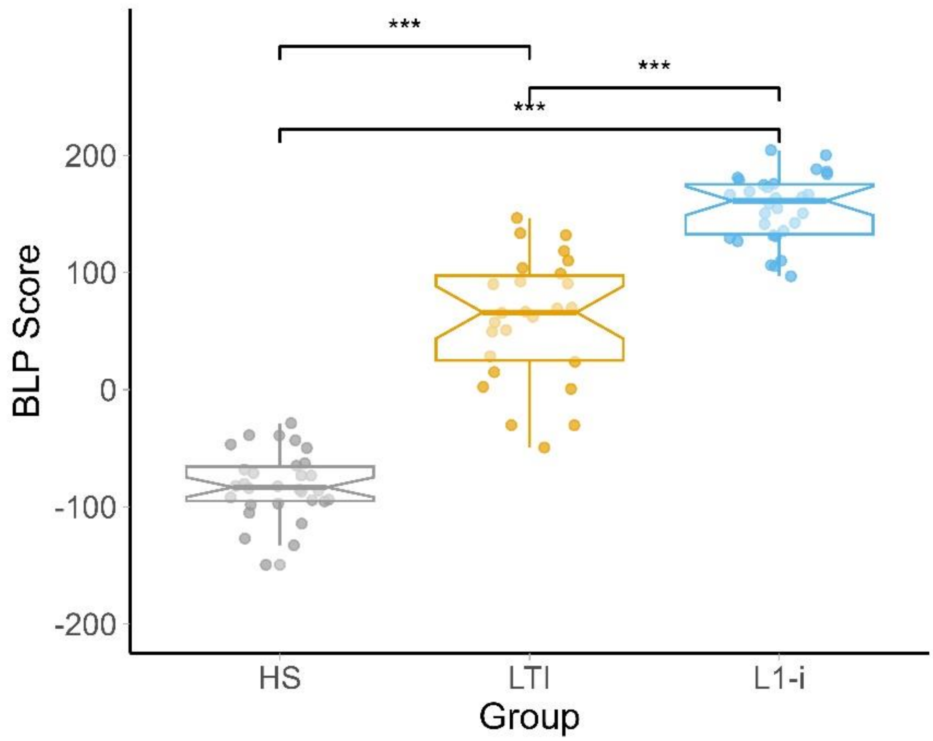
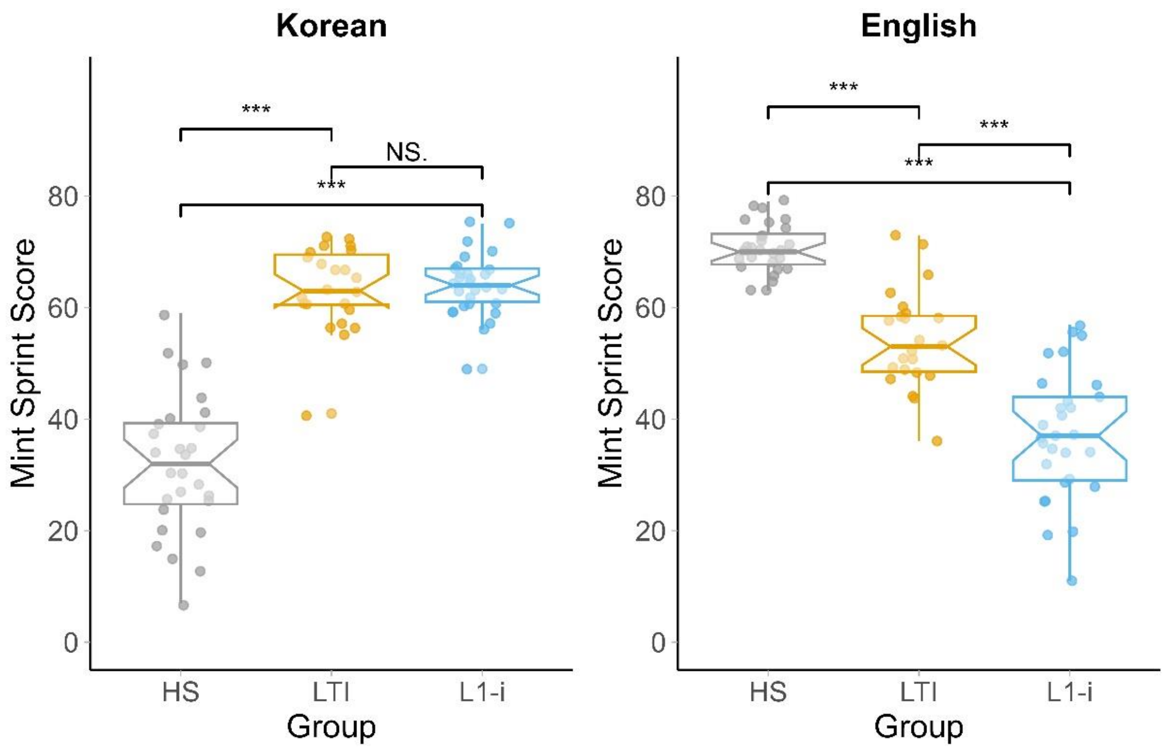


Figure 3.2. The MINT Sprint score by group and language.



3.1.2 Experiment 1: Three-alternative forced-choice (3AFC) identification task

Materials

The 3AFC task used a real-word triplet /pul/ (‘fire’, with a lenis stop), /p^{h}ul/ (‘grass’, with an aspirated stop), and /p*ul/ (‘horn’, with a fortis stop) created based on the resynthesis of the naturally recorded words pair /pul/, ‘fire’ and /p^{h}ul/, ‘grass’. Recordings were provided by a male Korean speaker in his twenties, originally from a Seoul-Gyeonggi area, who had spent two years in the US as a graduate student. The speaker read the triplet three to five times per word in a sound-attenuated booth, while being recorded using a cardioid condenser microphone. Recordings were digitized at a sample rate of 44.1 kHz and quantized at 16 bits. Recordings were resynthesized using a script (Winn, 2020) in Praat (Boersma & Weenink, 2022) in order to construct an eight-by-eight orthogonal VOT–f0 continuum in the word-initial stop. Intensity of the aspiration portion was controlled to covary with VOT (for more details on the functions of the script, please see Winn, 2020). The intensity of the vocalic portion of the stimuli was normalized to 80 dB to ensure consistent auditory perception across trials. VOT steps were created by modifying the VOT portion of the lenis stop to create eight log-scaled steps (10, 14, 19, 27, 38, 53, 74, and 104 ms), according to the approach adopted in the previous studies (Escudero et al., 2009; Kong et al., 2022; Kong & Kang, 2023). To create these steps, one token of /pul/ ('fire') was chosen as the basis. The initial unmodified VOT of the base recording was either lengthened or shortened. The use of log-scaled steps instead of equidistant steps was based on the observation that human perception of temporal differences in speech operates in a log-scale-like fashion, such that the same absolute time difference between two shorter sounds is perceived to be greater than between two longer sounds, (i.e., the difference between 10 and 15 ms is perceived to be greater than the difference between 100 and 105 ms (Buhusi & Meck, 2005; Escudero et al., 2009; Gibbon, 1977)).

Onset f0 ranged from 120 to 190 Hz in eight equidistant steps of 10 Hz for each VOT step, with f0 remaining constant (equal to the onset value) throughout the remainder of the vocalic portion of the word, which was achieved using a built-in function of the same Praat script (Winn, 2020). The resynthesis resulted in 64 unique stimuli for the identification task.

Procedure

Participants completed a 3AFC identification task using PsychoPy (version 3.8, Peirce et al., 2019). All L1-i participants performed the task in a speech perception and production laboratory at a university located in Seoul, South Korea, while HS and LTI participants completed the task either in the phonetics laboratories located in two midwestern universities in the USA or in their residence, provided the ambient noise was minimal. Participants were situated in front of a computer screen, wearing a headset. Participants could choose English or Korean as language of written instructions (most chose Korean). Oral interactions with the experimenter were also conducted mostly in Korean. English was used sparingly when clarification was needed for English-dominant participants. Participants completed five practice trials before proceeding to the main session. In each trial, participants listened to a word and identified it by choosing among the three response options displayed on the screen, using a mouse. The response options, '불'/pul/ ('fire'), '풀'/p^{h}ul/ ('grass'), and '뿔'/p*ul/ ('horn'), were arranged as a triangle, mitigating positional biases that might arise from a linear arrangement, and displayed in Korean orthography (see Appendix A). Participants were advised to respond as quickly as possible, and, after registering a decision, they proceeded to the next trial automatically with an inter-trial interval of

1.5 s. To mitigate potential sequential or carry-over effects, the location of response options was randomized across trials (i.e., /pul/ might appear on the top in one trial and on the left, right, or top in subsequent trials). The experiment consisted of a total of 384 trials, where 64 unique stimuli were repeated six times each. The task took approximately 15 minutes to complete, including instructions and practice.

Analysis

The 3AFC identification task generated a total of 32,640 data points. Three mixed-effects logistic regression models were implemented using R (version 4.2.1, R Core Team, 2022) and the 'glmer' function of the 'lme4' package (Bates et al., 2015). The models were designed to investigate the extent to which participants' phonetic categorization of stimuli was influenced by the two acoustic parameters: onset f0 and VOT. For each model, the binary dependent variable was participants' response (lenis-aspirated, lenis-fortis, and fortis-aspirated, coded as '0' and '1', respectively). The present study opted for models with a binomial function instead of a multinomial function in order to include random effects into the model, an option not supported for the multinomial regression in R. Each model included Group (HS, LTI, L1-i: reference, treatment coded), onset f0, VOT, and two-way interactions of Group with onset f0 and VOT as fixed effects to measure group variability in the use of the two acoustic cues in the categorization of each stimulus. The three-way interaction was excluded because it did not improve model fit and to avoid overfitting and the resulting convergence issues. In addition, each model included by-subject intercepts and slopes for both onset f0 and VOT as random effects to account for individual variability in cue weighting of onset f0 and VOT—individual sensitivity to acoustic cues in identifying laryngeal categories in Korean. Therefore, the models can quantify listeners' reliance on the acoustic dimensions of onset f0 and VOT via the fixed effect coefficients. The magnitude and significance of these coefficients directly reflect the extent to which each acoustic parameter influenced the participants' identification responses. For instance, a larger coefficient for an acoustic parameter indicates a stronger reliance on that cue for the categorization of one stop type over the other. When both acoustic cues, onset f0 and VOT, were significant in the model, the study compared their beta coefficients to determine the relative strength of each cue, following the method used in Tremblay et al. (2021).

3.1.3 Experiment 2: Speeded AX discrimination task

Materials

The stimuli for the speeded AX discrimination task were based on nine unique minimal pairs of monosyllabic and disyllabic real Korean words produced by the same male speaker (see Appendix B for the full list of stimuli). The speaker pronounced words in isolation, using the same equipment in the same phonetics lab where the identification stimuli were recorded. A single-speaker design was used in this task in order to eliminate the potential confounding variable of speaker differences inherent in discrimination tasks involving multiple speakers (Lee-Ellis, 2012). Each minimal pair contrasted stops in the word-initial position, starting with either a lenis or aspirated stop, followed by the same vowel, /a/. The target stops included three different places of articulation (/p/, /t/, /k/), with three words for each place of articulation. Recordings were resynthesized using the PSOLA method in Praat (Boersma & Weenink, 2022), creating a consistent VOT of 70 ms across all stimuli—the average VOT value for both lenis and aspirated stops in Korean (Kang & Guion, 2008)[1]. In addition, the intensity of the stimuli was normalized at 80 dB. The purpose of this resynthesis was to examine participants' ability to perceive the lenis-aspirated stop contrast without relying on VOT, the primary cue for the voicing contrast in English stops but a secondary cue for the Korean lenis-aspirated contrast. The mean value of onset f0 in the recorded words was 136.9 Hz (SD = 5.1) and 183.9 Hz (SD = 6.4) for lenis and aspirated stops, respectively, and this difference was statistically significant ($t(15.2) = 17.19$, $p < .001$). There was no significant effect of place of articulation on onset f0 within the same laryngeal category ($p > .05$). The 'same' stimuli pairs (AA or XX) were created by pairing different recordings of the same words. This approach served to increase the complexity of the otherwise relatively simple task by requiring participants to decide whether the 'same' stimuli were phonologically noncontrastive rather than acoustically identical. 'Different' pairs were created by pairing together members of the same minimal pair, such as /tal/ ('the moon', a lenis stop) vs. /t^hal/ ('a mask', an aspirated stop).

[1] In the original, naturally recorded words, the mean VOT for alveolar stops was the longest (74.5 ms), followed by velar and bilabial stops (73 and 64 ms, respectively), with only the difference between alveolar and bilabial stops being significant ($t(9.7) = 2.6$, $p = .02$). The mean VOT of lenis (69.1 ms) was not significantly different from that of aspirated stops (72.1 ms) ($p > .05$).

Procedure

Upon completion of the 3AFC task, participants were offered an optional break before starting the AX discrimination task implemented in PsychoPy (version 3.8, Peirce et al., 2019). Once ready to proceed, participants were seated in front of a computer monitor wearing a headset and were given written and oral instructions. Participants were aurally presented with pairs of words that could potentially differ in the laryngeal specification of the word-initial stop (lenis or aspirated), and their task was to determine whether the two words were the same or different on each trial. They registered their responses by clicking on one of the two circle-shaped buttons labeled 'different' and 'same' in Korean orthography (see Appendix A). Importantly, participants were asked to decide as quickly as possible within up to three seconds to minimize potential lexical effects, including frequency effects (Johnson & Babel, 2010). The screen automatically advanced to the next trial upon participants' button click or once the three-second interval elapsed. An interval of 200 ms separated Word A from Word X, and each trial was followed by a period of 500 ms of a blank screen. The interval of 200 ms was used to provide the minimum time required to recognize and process a sound signal as well as to prevent mutual masking between two tokens (Gerrits & Schouten, 2004). The stimuli order and the order of the response options were randomized for each participant. Each of the nine unique minimal pairs appeared an equal number of times in both orders, AX and XA, and an equal number of 'same' (XX and AA) and 'different' pairs were used. There were a total of 216 trials presented to each participant: 18 test words (three words for each unique combination of laryngeal type (lenis or aspirated) and place of articulation (/p/, /t/, /k/)) × 2 orders (AX and XA) × 3 repetitions × 2 pairing/trial types ('same' and 'different').

Analyses

Participants' responses, a total of 18,144 data points, were coded as either '1' correct or '0' incorrect and submitted to statistical analyses. A mixed-effects logistic regression model was implemented with the binary dependent variable of accuracy. The model included Group (HS, LTI, L1-i), Trial Type (same and different trials), interactions between Group and Trial Type, f0 weight, and an interaction between f0 weight and Trial Type. The model also included by-subject and by-item intercepts as random effects, based on the maximal random effect structure (Bates et al., 2015), resulting in the following glmer function in R: Response ~ Group*TrialType +

TrialType*F0Weight + (1|Subject) + (1|Item). The present study selected the best fit model by employing a stepwise model comparison approach: using likelihood ratio tests by comparing the full model to the nested models. The f0 weight parameter was a continuous variable representing each participant's relative weight of the onset f0 coefficient in the identification of lenis vs. aspirated stops in the 3AFC task. This weight was calculated for each participant based on the method used in previous studies (Kong et al., 2022; Kong & Kang, 2023). Specifically, each participant's random effect coefficient for onset f0 from the mixed-effects logistic regression model examining lenis and aspirated pair was added to its corresponding fixed effect coefficients (onset f0 and group) extracted from the same model. Therefore, this variable enabled the model for the AX discrimination to determine to what extent individual weighting of the f0 cue over the VOT cue, estimated in the perception of aspirated stops vs. lenis stops in the 3AFC task, was associated with perceptual accuracy in discriminating between lenis and aspirated stops. In other words, the model can reveal whether a greater f0 weight in categorical perception leads to more accurate discrimination of the phonemic contrast between lenis and aspirated stops. The categorical variables, Group and Trial Type, were sum-coded to create matrices of 3×2 and 2×1, respectively, by assigning 1 to comparison groups, −1 to groups never compared to other levels, and 0 to all the other comparisons. Therefore, the intercept equals the grand mean of the means of all levels of the factors, and the beta coefficients indicate the deviations (main effects) of a level of a factor from the grand mean (the intercept), averaged for the other factors. The choice of the sum coding was to ensure a balanced and symmetrical comparison across all levels of each factor, such that the model reflects the research question regarding the effect of f0 weight on participants' discriminatory performance across all groups, without using any group as a baseline.

To further investigate the impact of CLI on auditory processing in the perception of the lenis-aspirated stop contrast, the present study also implemented a supplementary mixed-effects linear regression model with the dependent variable of log-transformed RT. Only the data from 'correct' responses in different-word trials (n = 8,902) were analyzed in this model, and RT was calculated by measuring the interval from the onset of the response buttons' appearance on the screen to the moment response was registered on each trial, following the approach in previous studies examining RT in perception tasks (Johnson & Babel, 2010; Stevenson, 1973; Wrembel et al., 2019, among others). The model included Group, f0 weight, and their interaction as fixed effects and by-item and by-subject intercepts as random effects, following the maximally allowed random effect structure. Despite the potential differences in latencies due to varying experiment environments among participants, the use of PsychoPy software for RT measurement provides sufficiently reliable data, given its precision in gauging response times (under 3.5 ms, according to Bridges et al., 2020).

3.2 Results

3.2.1 3AFC identification task

Distributions of participants' responses in the identification of the three Korean stops

Figure 3.3 displays the distributions of participants' responses representing the extent of reliance on onset f0 (y-axis) and VOT (x-axis) in identifying the three Korean stops (lenis, aspirated, fortis) in the stimuli matrix (eight f0 steps crossed with eight VOT steps).

For the identification of lenis stops, a primary reliance on onset f0 was evident in the L1-i and LTI groups. There was also a non-monotonic association between intermediate VOT values and lenis stop responses across all groups, with the association increasing from the first VOT step to the intermediate step and decreasing from there to the last step. HSs showed a more variable distribution of responses as a function of onset f0, suggesting a less categorical use of this cue, compared to the other groups.

Aspirated stop responses were characterized by a reliance on both longer VOT and higher onset f0, with HSs displaying a more gradient pattern of reliance on these cues. Especially conspicuous was the occasional identification of stops with high f0 but short VOT as aspirated, not observed in the other groups.

The identification of fortis stops was predominantly influenced by VOT, but higher onset f0 also played a role. HSs again were more variable in their use of VOT values to identify fortis stops.

The distinct patterns across groups are more evident in Figure 3.4. This heatmap plot shows aggregated responses for stop types within one matrix for each group. Each cell is labeled with the most frequently identified stop type (L: lenis, A: aspirated, F: fortis). The color transparency represents the proportion of times this stop type was chosen, relative to VOT and onset f0 cues, with the color intensity levels corresponding to three levels of categorization: ambiguous (0–0.49), categorical (0.5–0.74), and highly categorical (0.75–1.0) identification, following Lee et al. (2013).

The heatmap for the HS group displays a smaller number of cells with proportions above 0.75 compared to those of the LTI and L1-i groups, suggesting that HSs categorized Korean laryngeal stops with less certainty in the 3AFC task. Instead, the HS group demonstrated a higher incidence of ambiguity, particularly for lenis stops. LTIs categorized a wider range of VOT continuum as lenis stops than L1-i speakers. Moreover, LTIs were less likely to identify the stimulus at the 8th VOT step and the 4th f0 step as aspirated stops as compared to both HSs and L1-i speakers, suggesting a more pronounced f0-based selection for aspirated stops than the other groups' participants. The following sections present the results of the statistical analyses for the 3AFC identification task.

Figure 3.3. Heatmaps representing the distributions of participants' responses in the 3AFC identification task on a stimuli matrix by group and stop type.

Color transparency in each block represents proportions of participants' responses of each stop category: the darker the cell, the more corresponding stop responses.

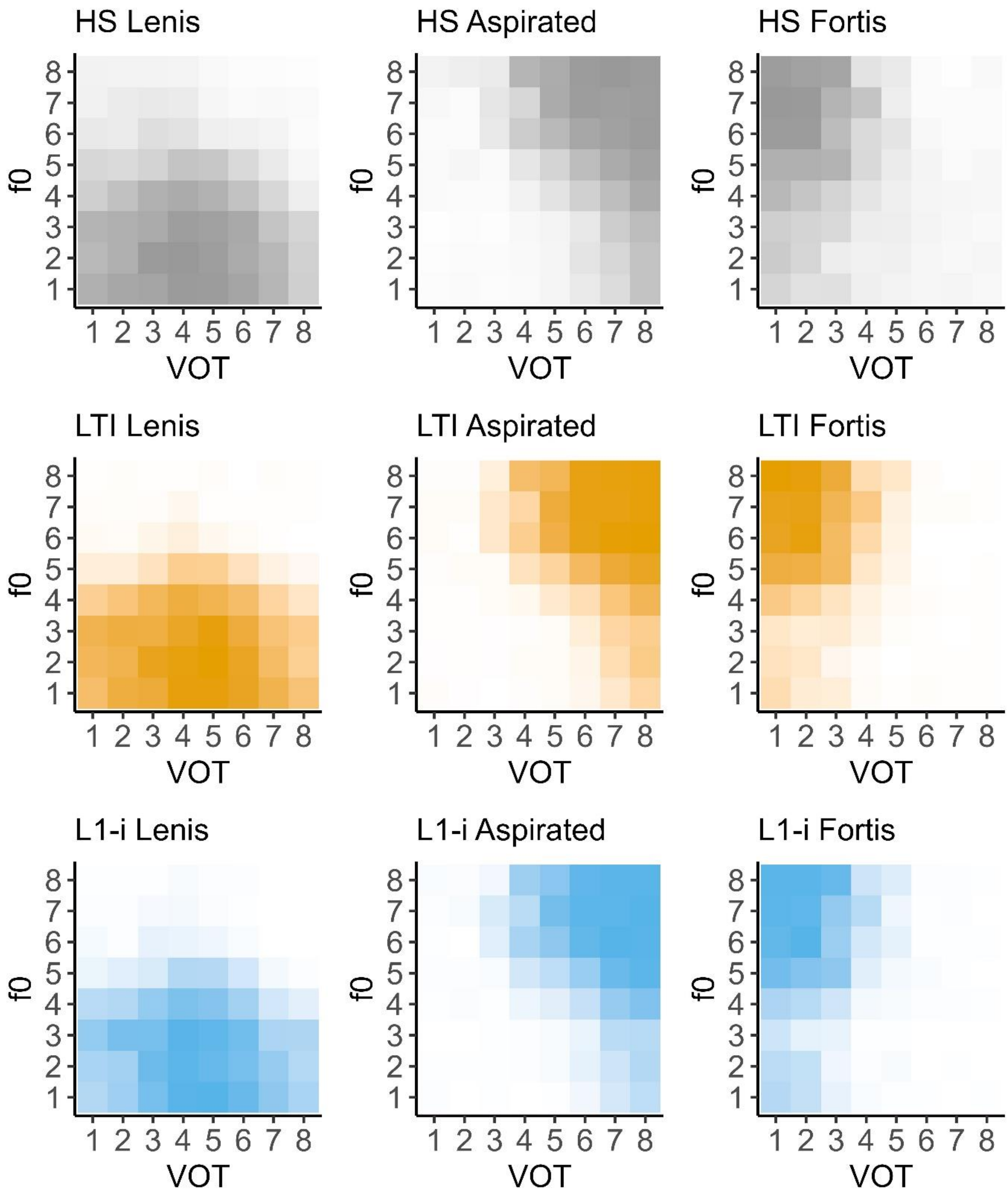
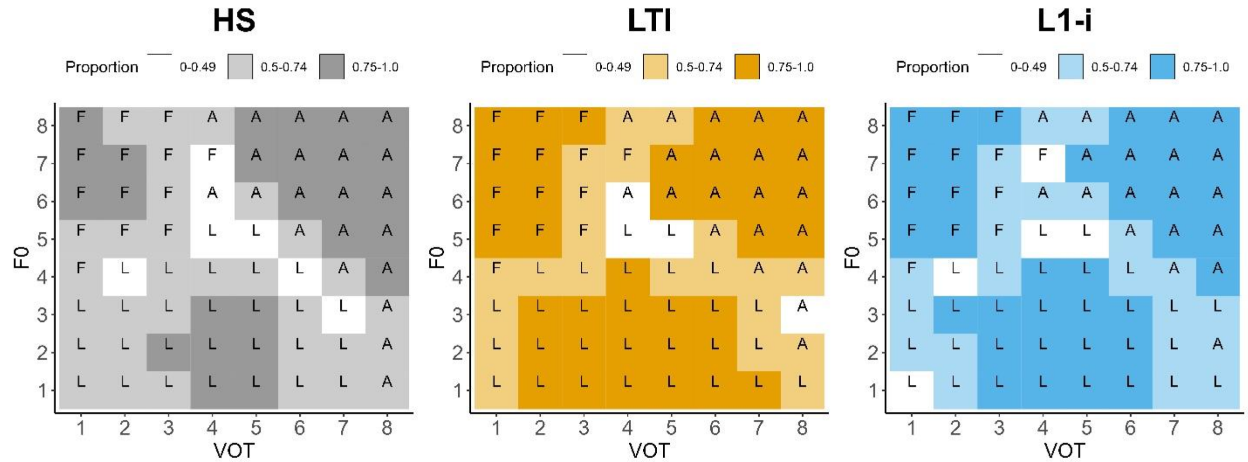


Figure 3.4. Participants' distributions of stop type responses in the 3AFC identification task, presented in a stimuli matrix categorized by group.

Each matrix cell labels the most frequently chosen stop type, and the color transparency reflects the proportion of selections within three defined ranges (ambiguous: 0–0.49, categorical: 0.5–0.74, highly categorical: 0.75–1.0).



Cue weighting to lenis-aspirated contrast

The primary purpose of the 3AFC identification task was to investigate to what extent the two bilingual groups (HS and LTI) differed from L1-i participants in their weighting of the two acoustic parameters, onset f0 and VOT, in identifying lenis, aspirated, and fortis stops. The statistical results indicated that both onset f0 and VOT significantly predicted the categorization of the two types of stops for L1-i speakers, with onset f0 serving as the primary cue. Specifically, for each incremental step in onset f0, the odds ratio of selecting the 'aspirated' response significantly increased by a factor of 3.36 ($\beta = 1.21$, SE = 0.07, $z = 16.88$, $p < .001$). A one-step increase in VOT also significantly increased odds ratio, though less than by onset f0, by a factor of 2.57 ($\beta = 0.945$, SE = 0.062, $z = 15.186$, $p < .001$). Comparing the beta coefficients of onset f0 and VOT cues revealed that the perception of the lenis-aspirated stop contrast for L1-i speakers was determined primarily by onset f0, while VOT played a secondary role ($t = -2.75$, $p < .001$).

The interaction between Group and onset f0 revealed that a significantly different onset f0 weight was used by the HS and LTI groups, as compared to the L1-i group. Specifically, the LTI group showed greater sensitivity to the f0 cue, demonstrating a higher odds ratio by onset f0 than the L1-i group by a factor of 1.28 (LTI vs. L1-i:f0, $\beta = 0.24$, SE = 0.11, $z = 2.28$, $p = .02$). In contrast, the HS group was less sensitive compared to the L1-i group, with a decreased odds ratio by a factor of 0.74 (HS vs. L1-i:f0, $\beta = -0.30$, SE = 0.10, $z = -2.97$, $p = .003$). HSs' weaker reliance on the tonal cue can be observed in Figure 3.5 (top-left): the logistic curve for HS is visibly less steep in comparison to the LTI and L1-i curves. An examination of the coefficients for VOT and onset f0 parameters, as shown in the left panel of Figure 3.6, further unveils that the auditory processing pattern of HS differed from both LTI and L1-i groups. Notably, the data points representing coefficients of onset f0 for LTI participants clustered predominantly in the upper regions of the y-axis, indicating greater sensitivity to the f0 cue compared to the other groups. The analysis of the estimated marginal means of trends also confirmed these group patterns, with the LTI group showing the strongest effect of onset f0 in the identification of aspirated stops and the HS group showing the weakest reliance on onset f0 cue (L1-i: $\beta = 1.21$, SE = 0.07, 95% confidence interval (CI) = [1.07, 1.35]; LTI: $\beta = 1.46$, SE = 0.08, 95% CI = [1.30, 1.61]; HS: $\beta = 0.92$, SE = 0.07, 95% CI = [0.78, 1.05]). Furthermore, these estimated marginal means of trends revealed that LTIs' reliance on onset f0 was significantly greater than that of HSs (LTI vs. HS:f0, $\beta = 0.54$, SE = 0.11, $z = 5.11$, $p < .001$).

Although data visualization suggested that the HS group relied more strongly on VOT compared to the other two groups (HS group's VOT slope was somewhat steeper in Figure 3.5 and positive in Figure 3.6, left panel), the interaction between Group and VOT was not statistically significant in the perceptual identification of lenis and aspirated stops. Therefore, the results suggest that CLI affects the use of onset f0 in both bilingual groups rather than VOT in the identification of lenis and aspirated stops.

In summary, these findings demonstrated that both the HS and LTI groups implemented a distinct cue-weighting strategy in the perception of lenis and aspirated stops, as compared to the L1-i group, but in different directions. The HS group leaned less heavily on onset f0, the primary cue for differentiating the two types of stops, than the L1-i group. Interestingly, the statistical analyses revealed the LTI group as the most sensitive to the onset f0 cue, followed by L1-i and HS groups. The results of all mixed-effects logistic regression models employed to examine perceptual cue-weighting data from the 3AFC task are summarized in Table 3.2.

Figure 3.5. Logistic curves representing the probabilities of identifying 'aspirated', 'fortis', and 'aspirated' responses from left to right by column plotted on the y-axes.

The x-axes of the first and second rows indicate f0 and VOT steps, respectively. The steepness of the curves corresponds to the beta coefficients from each regression model.

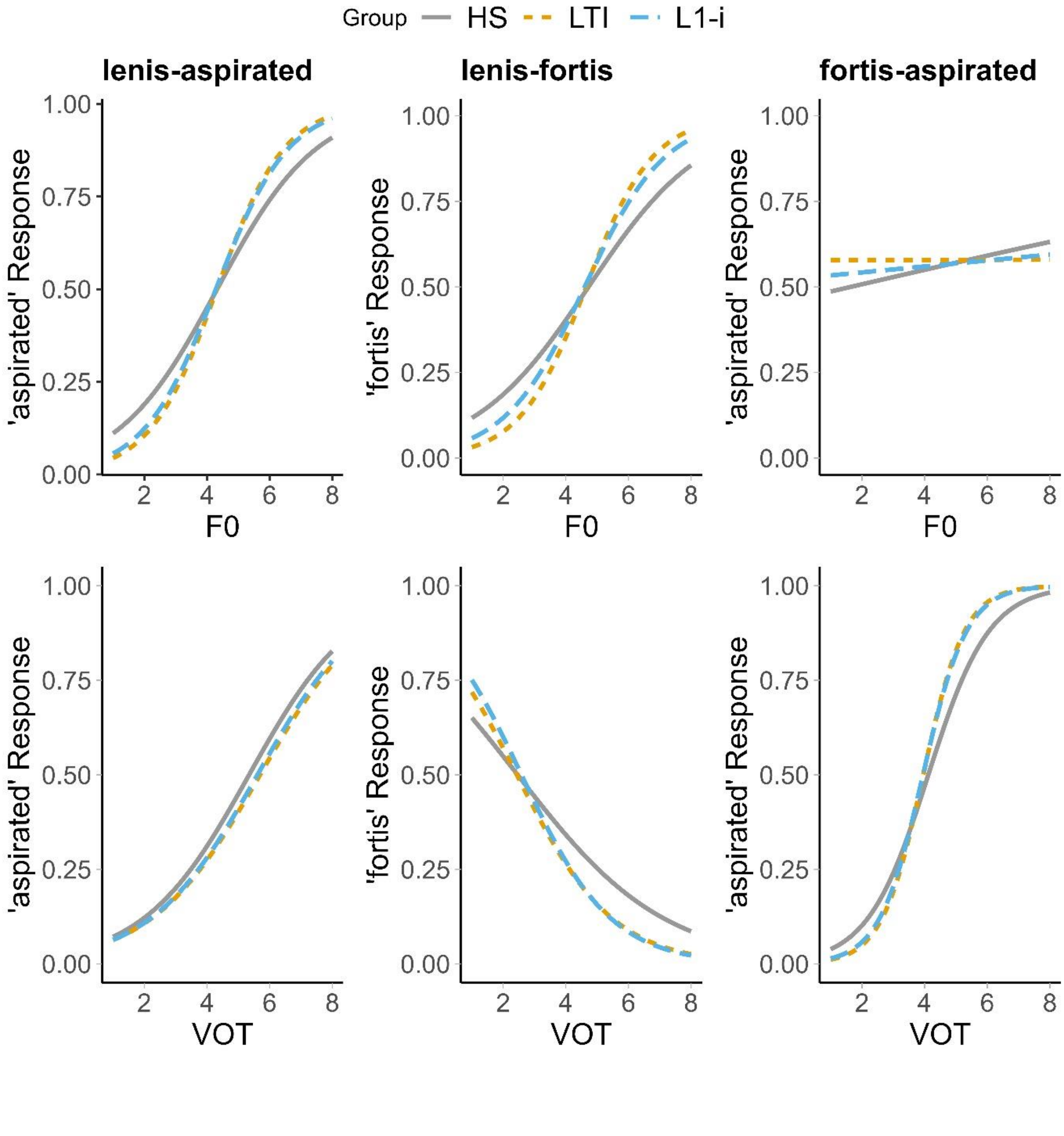


Figure 3.6. Scatter plots showing onset f0 (y-axis) and corresponding VOT (x-axis) coefficients for each participant.

Each data point was calculated by adding random slopes of each mixed-effects logistic regression model for the 3AFC identification task to correlated fixed effects. Points with higher y-values indicate a stronger reliance on f0, while points further along the x-axis suggest a greater dependence on VOT. The orientation and slope of the lines indicate the trade-offs, if any, in terms of individual reliance on the two cues.

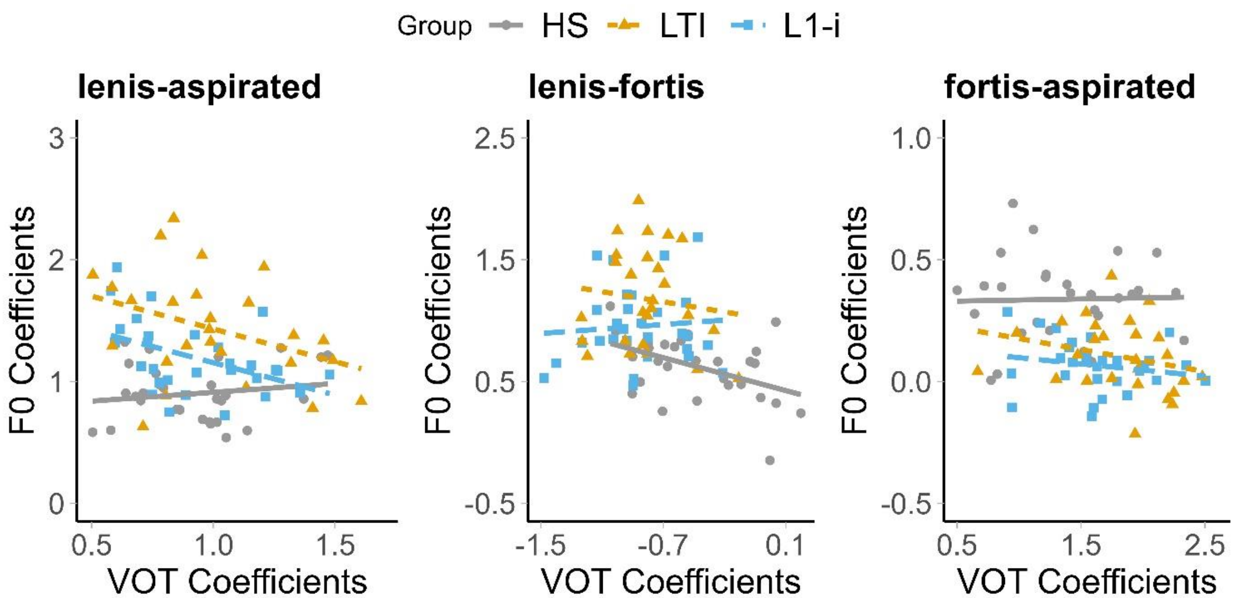


Table 3.2. A summary of the mixed-effects logistic regression models examining perceptual cue weighting to Korean stops in the 3AFC identification task.

	(a) lenis-aspirated			(b) lenis-fortis			(c) fortis-aspirated		
	β (SE)	*z*	*p*	β (SE)	*z*	*p*	β (SE)	*z*	*p*
Intercept	−10.384 (0.529)	−19.633	< .001	−1.802 (0.453)	−3.977	< .001	−7.046 (0.524)	−13.442	< .001
LTI	**−1.606 (0.788)**	**−2.039**	**.042**	−1.287 (0.662)	−1.944	.052	−0.967 (0.777)	−1.244	.213
HS	1.397 (0.737)	1.896	.058	0.326 (0.638)	0.512	.609	−0.691 (0.737)	−0.938	.348
f0	**1.212 (0.072)**	**16.884**	**< .001**	**0.976 (0.072)**	**13.637**	**< .001**	0.068 (0.046)	1.464	.143
VOT	**0.945 (0.062)**	**15.186**	**< .001**	**−0.876 (0.068)**	**−12.851**	**< .001**	**1.681 (0.102)**	**16.491**	**< .001**
LTI:f0	**0.243 (0.107)**	**2.280**	**.023**	**0.229 (0.105)**	**2.168**	**.030**	0.039 (0.069)	0.562	.574
HS:f0	**−0.297 (0.1)**	**−2.971**	**.003**	**−0.345 (0.1)**	**−3.47**	**< .001**	**0.274 (0.063)**	**4.363**	**< .001**
LTI: VOT	0.069 (0.092)	0.755	.450	0.049 (0.1)	0.489	.625	0.169 (0.15)	1.131	.258
HS:VOT	−0.009 (0.087)	−0.102	.919	**0.4 (0.094)**	**4.219**	**< .001**	−0.266 (0.142)	−1.874	.061

Cue weighting to lenis-fortis contrast

The results from the logistic regression analysis confirmed that both parameters were significantly predictive for L1-i speakers, with a greater weight assigned to onset f0. L1-i participants were significantly more likely to identify a stop as fortis, rather than lenis, by a factor of 2.67, for each incremental step in onset f0 ($\beta = 0.98$, SE = 0.07, $z = 13.64$, $p < .001$). In contrast, an increase of one step in VOT significantly decreased the likelihood of fortis stop identification by L1-i speakers by a factor of 0.41 ($\beta = -0.88$, SE = 0.07, $z = -12.85$, $p < .001$). The comparison of the coefficients of onset f0 and VOT also revealed that L1-i speakers' reliance on onset f0 was greater than on VOT for the perception of lenis and fortis stops ($t = 17.11$, $p < .001$)[2].

Group-wise analysis of onset f0 weight showed similar patterns to those observed for the lenis-aspirated comparisons, as shown in the top-center plot of Figure 3.5. Specifically, the odds ratio of selecting fortis stops over lenis stops significantly decreased in the HS group by a factor of 0.71 compared to the L1-i group for each one-step increase in onset f0 ($\beta = -0.345$, SE = 0.10, $z = -3.47$, $p < .001$). In contrast, the LTI group displayed a significant increase in the odds ratio by a factor of 1.26, compared to the L1-i group, with a one-step increase in onset f0 ($\beta = 0.23$, SE = 0.11, $z = 2.17$, $p = .03$). This indicates that the LTI group was the most reliant on the f0 cue, while the HS group was the least reliant on f0 in differentiating between lenis and fortis stops among the participant groups. This comparative reliance on onset f0 versus VOT is visually represented in the center plot of Figure 3.6, where the data points for the LTI group are concentrated higher on the y-axis (onset f0) than those of the L1-i and HS groups. The greater reliance on onset f0 by the LTI group than by the HS group was also supported by the analysis of the estimated marginal trends, evidenced by the significant difference in the beta coefficient between the two groups ($\beta = 0.57$, SE = 0.10, $z = 5.50$, $p < .001$) (estimated marginal trends: L1-i: $\beta = 0.98$, SE = 0.07, 95% CI = [0.84, 1.12]; LTI: $\beta = 1.21$, SE = 0.08, 95% CI = [1.05, 1.36]; HS: $\beta = 0.63$, SE = 0.07, 95% CI = [0.50, 0.77]).

Assessment of the interaction effects between group and VOT revealed a significantly increased VOT effect compared to the L1-i group solely within the HS group ($\beta = 0.40$, SE = 0.09, $z = 4.22$, $p < .001$), while there was no significant difference in the VOT weighting between the

[2] This result should be interpreted with caution as the directions of the two coefficients being compared are different. Nevertheless, the greater absolute value of the onset f0 coefficient than that of the VOT coefficient suggests that onset f0 plays a more important role than VOT in the lenis-fortis stop contrast.

LTI and L1-i groups (p = .63). This finding reveals a greater effect of VOT on fortis identification within the HS group, due to an unexpected tendency to categorize stimuli with long VOT as fortis stops, despite the fact that fortis stops in Korean are associated with a short-lag VOT. In contrast, the LTI group mirrored the pattern of the L1-i group by considerably decreasing the probability of fortis identification as VOT step increased. The post-hoc analysis also revealed that LTIs' reliance on VOT was lesser in degree compared to HSs ($\beta = -0.35$, SE = 0.10, $z = -3.57$, p = .001). This distinct pattern of VOT weighting by the HS group is illustrated in the bottom-center plot of Figure3.5, where the shallow negative slope for the HS group is contrasted with steep slopes of the LTI and L1-i groups extending towards high VOT steps on the x-axis. HSs' greater likelihood of selecting fortis over lenis stops is also supported by their greatest beta coefficient obtained from the analysis of estimated marginal means of trends (L1-i: $\beta = -0.88$, SE = 0.07, 95% CI = [−1.01, −0.74]; LTI: $\beta = -0.83$, SE = 0.07, 95% CI = [−0.97, −0.68]; HS: $\beta = -0.48$, SE = 0.07, 95% CI = [−0.61, −0.35]).

To summarize, these findings illustrate the divergent cue-weighting strategies employed by each group in their perceptual identification of lenis and fortis stops. The HS group demonstrated a distinct approach characterized by a decreased reliance on onset f0 compared to L1-i and LTI speakers and an unexpected tendency to perceive longer VOT stimuli as fortis stops. Conversely, the LTI group, much like the L1-i group, relied more heavily on onset f0 and showed a decreased likelihood of identifying fortis stops with an increase in VOT.

Cue weighting to fortis-aspirated contrast

The statistical analysis confirmed that for the L1-i group, VOT was the sole predictive cue in the identification of fortis and aspirated stops (β = 1.68, SE = 0.10, z = 16.49, $p < .001$), with onset f0 not being a significant predictor (p = .14). The examination of the interaction effects of Group with onset f0 revealed significant deviation within the HS group in comparison to the L1-i group due to the increased odds ratio of selecting aspirated stops over fortis stops by a factor of 1.32 (β = 0.27, SE = 0.06, z = 4.36, $p < .001$) associated with a one-step increase in onset f0. On the contrary, the LTI group's pattern of no significant reliance on onset f0 was consistent with that of the L1-i group (p = .57). The post-hoc analysis of the estimated marginal trends revealed that the LTI group's reliance on onset f0 was significantly lower compared to that of the HS group (β

= −0.24, SE = 0.07, *z* = −3.53, *p* = .001). This finding indicated a unique tendency of the HS group to make an onset f0-based distinction between fortis and aspirated stops. The visualization of the model data in the right panel of Figure 3.5 illustrates this finding, demonstrating a steeper slope for HSs than for LTI and L1-i participants. Furthermore, as shown in the right panel of Figure 3.6, the data points for the HS group are clustered higher on the y-axis, indicating a greater effect of onset f0 in the identification of fortis and aspirated stops. In contrast, data points for the LTI and L1-i groups are primarily clustered around zero on the y-axis, indicating that onset f0 was not a cue to this contrast. The post-hoc analysis also showed that the HS group had the greatest slope for onset f0 among the groups, supporting their distinct use of the onset f0 cue (L1-i: β = 0.07, SE = 0.05, 95% CI = [−0.03, 0.16]; LTI: β = 0.11, SE = 0.05, 95% CI = [0.01, 0.21]; HS: β = 0.34, SE = 0.04, 95% CI = [0.26, 0.43]). The interaction effects of Group with VOT did not show any significant deviation in either the LTI or HS groups from the L1-i group.

Effects of bilingual profiles on the perception of lenis-aspirated stop contrast

The study conducted a post-hoc analysis to examine the effects of the BLP and Korean MINT Sprint scores on the use of onset f0 cue by bilingual speakers (HSs and LTIs) in the perception of lenis and aspirated stops. Specifically, the present study investigated the effects of onset f0, relative language dominance (BLP score) and language proficiency (Korean MINT Sprint score) as fixed effects on the binary dependent variable (1: 'aspirated', 0: 'lenis' responses) in the 3AFC task (a 'glm' model without random effect structures due to convergence issues). The rationale for examining the lenis-aspirated stop contrast is based on the observation from the 3AFC identification task that onset f0 is the locus of CLI in the perception of in this contrast. This analysis indicated a significant interaction effect between BLP score and f0 cue weight, suggesting that the role of f0 in the identification of aspirated stops vs. lenis stops was modulated by participants' language dominance (β = 0.003, SE = 0.000, *z* = 6.61, *p* < .01). Specifically, the effect of f0 cue (f0 weight) on the log odds of the 'aspirated' response increased by 0.003 units for each additional unit increase in the BLP scores towards Korean dominance. In other words, bilingual speakers with higher Korean dominance had a stronger positive relationship between f0 weight and the 'aspirated' response, compared to those who were less dominant in Korean.

3.2.2 Speeded AX discrimination task

Effects of group and type

The examination of the descriptive statistics shows that participants across all groups demonstrated relatively high accuracy on the AX discrimination task, but with noticeable differences by group and trial type. The HS group demonstrated a mean accuracy of 92.5% (SD = 25.6) for same trials and 96.5% (SD = 18.4) for different trials. The LTI group had average accuracies of 94.9% (SD = 22.1) and 99.5% (SD = 7.3) for same and different trials, respectively. The L1-i group showed accuracies of 92.9% (SD = 25.6) and 98.6% (SD = 11.9) in the same and different trials, respectively. Figure 3.7 represents the estimated log-ratio for all participants by group and trial type, based on the random effects for subjects and their corresponding fixed effects from the mixed-effects logistic regression model. As shown in the figure, all participant groups performed better in the different-word trials than in the same-word trials. Furthermore, it is manifest that HSs' accuracy was lower than that of the other groups, especially in the different-word trials. Finally, each group exhibited individual variability in their accuracy performance in the task.

Table 3.3. A summary of the mixed-effects logistic regression model examining perceptual accuracy in the AX discrimination task.

	Estimate (SE)	*z*	*P*
Intercept (grand mean)	4.083 (0.155)	26.342	< .001
HS vs. grand mean	**−0.622 (0.166)**	**−3.752**	**< .001**
LTI vs. grand mean	**0.687 (0.185)**	**3.706**	**< .001**
different vs. same	**0.775 (0.106)**	**7.334**	**< .001**
f0weight	0.034 (0.077)	0.441	.66
HS:different vs. same	**−0.377 (0.066)**	**−5.700**	**< .001**
LTI: different vs. same	**0.353 (0.094)**	**3.736**	**< .001**
different vs. same:f0weight	**0.068 (0.031)**	**2.176**	**.03**

Table 3.3 provides a summary of the logistic regression analysis of perceptual accuracy in the AX discrimination task data. The statistical results indicated a significant deviation from the grand mean, averaged for Trial Type, for both the HS and LTI groups in their discriminatory

performance (HS, $\beta = -0.62$, SE = 0.17, $z = -3.75$, $p < .001$; LTI, $\beta = 0.69$, SE = 0.19, $z = 3.71$, $p < .001$). Using these group coefficients, the L1-i group's deviation from the intercept was calculated at 0.07 (= $\beta_1 + \beta_2$). This suggests that while the HS group's discriminatory accuracy lagged behind the L1-i group, the LTI group surpassed the L1-i group in terms of discrimination accuracy, averaged for the Trial Type. A post-hoc analysis, conducted using the Tukey method for comparing a family of three estimates by averaging the levels of Type, evidenced that the LTI group surpassed both the HS and L1-i groups in accuracy, averaged for Trial Type (HS vs. LTI, $\beta = -1.31$, SE = 0.31, $z = -4.25$, $p < .001$; LTI vs. L1-i, $\beta = 0.75$, SE = 0.31, $z = 2.40$, $p = .04$). However, when averaged for Trial Type, there was no significant difference in estimated accuracy between the HS and L1-i groups ($p = .11$).

The effect of Trial Type, averaged for Group, was significant, as all groups demonstrated superior accuracy in 'different' trials compared to 'same' trials ($\beta = 0.78$, SE = 0.11, $z = 7.33$, $p < .001$). This 'false alarm' bias indicated that participants were prone to responding 'different' when the words were actually the same.

The logistic regression analysis also indicated significant interaction effects between Group and Trial Type, pointing to a different impact of Trial Type on the discrimination performance among groups. For the HS group, the effect of Trial Type was less pronounced, as reflected by a decreased difference in performance from 'different' to 'same' trials, compared to the average decrease from the grand mean across all groups ($\beta = -0.38$, SE = 0.07, $z = -5.70$, $p < .001$). This suggests that the HS group's performance was less influenced by Trial Type compared to other groups. In contrast, the LTI group's performance was more affected by Trial Type. This pattern of the LTI group is shown in Figure 3.7 by a greater decrease in performance from 'different' to 'same' trials ($\beta = 0.35$, SE = 0.09, $z = 3.74$, $p < .001$).

To further examine these interaction effects, a post-hoc analysis using Tukey multiplicity adjustment method was performed. The results of pairwise comparisons, displayed in Figure 3.8, revealed that in 'different' trials, the HS group was less likely to select correct responses by a factor of 0.13 compared to the LTI group, and this difference was significant ($\beta = -2.04$, SE = 0.38, $z = -5.32$, $p < .001$). Similarly, the HS group significantly deviated from the L1-i group in 'different' trials with a decrease in the odds ratio by a factor of 0.38 ($\beta = -0.96$, SE = 0.32, $z = -3.04$, $p = .01$). Interestingly, the LTI group exhibited significantly higher odds of correct responses in 'different' trials, outperforming the L1-i group by a factor of 2.95 (β = 1.08, SE = 0.40, z = 2.71, p = .018).

However, none of the pairwise comparisons in 'same' trials demonstrated significant differences (HS-LTI: p = .118; HS-L1-i: p = .841; LTI-L1-i: p = .319), indicating that the overall difference in accuracy between groups was driven by 'different' rather than 'same' trials. To summarize, the post-hoc analysis confirmed that, in the absence of the VOT cue, the LTI group displayed the highest accuracy in discriminating the lenis-aspirated stop contrast, while the HS group was the least accurate.

Figure 3.7. Violin plot of adjusted effects for the AX discrimination task, separated by participant group and Trial Type.

The plot represents the distribution of the estimated log-odds of correct responses as predicted by the mixed-effects logistic regression model. The individual data points reflect the predicted log-odds for each participant, adjusted for the mean effects of respective Group and Trial Type, based on the random effects. The dashed line at 4.1 indicates the grand mean, the intercept across all participants. The width of the violins shows the density of data points at different levels of the adjusted effect.

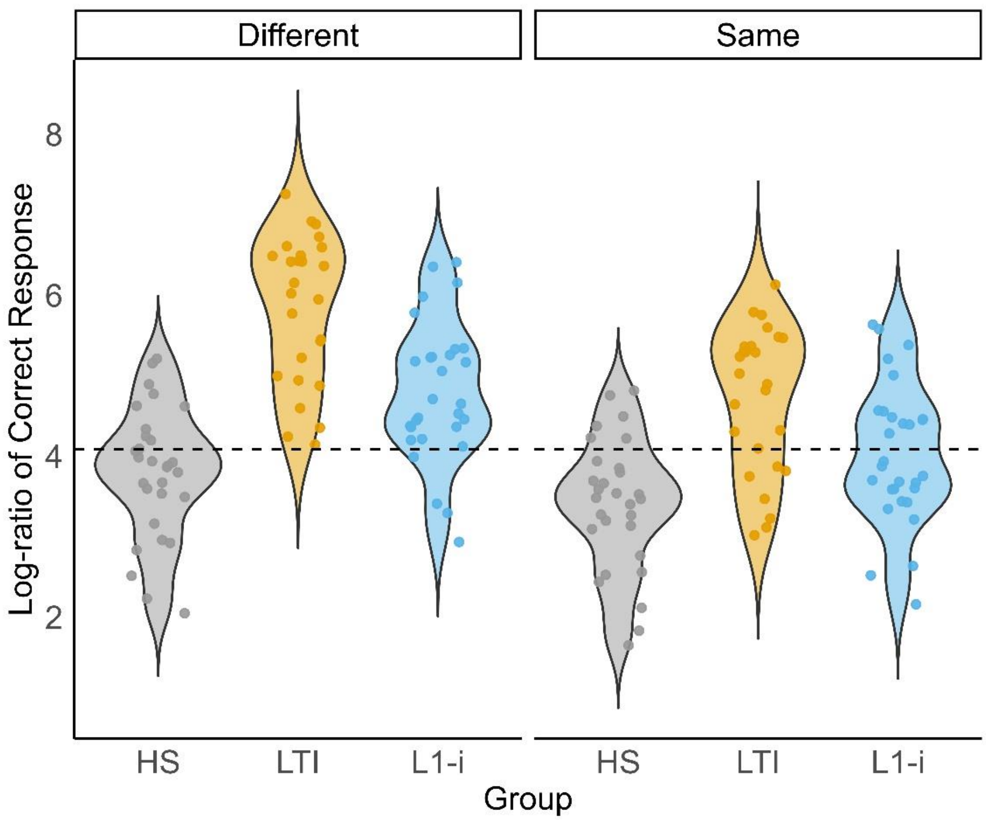
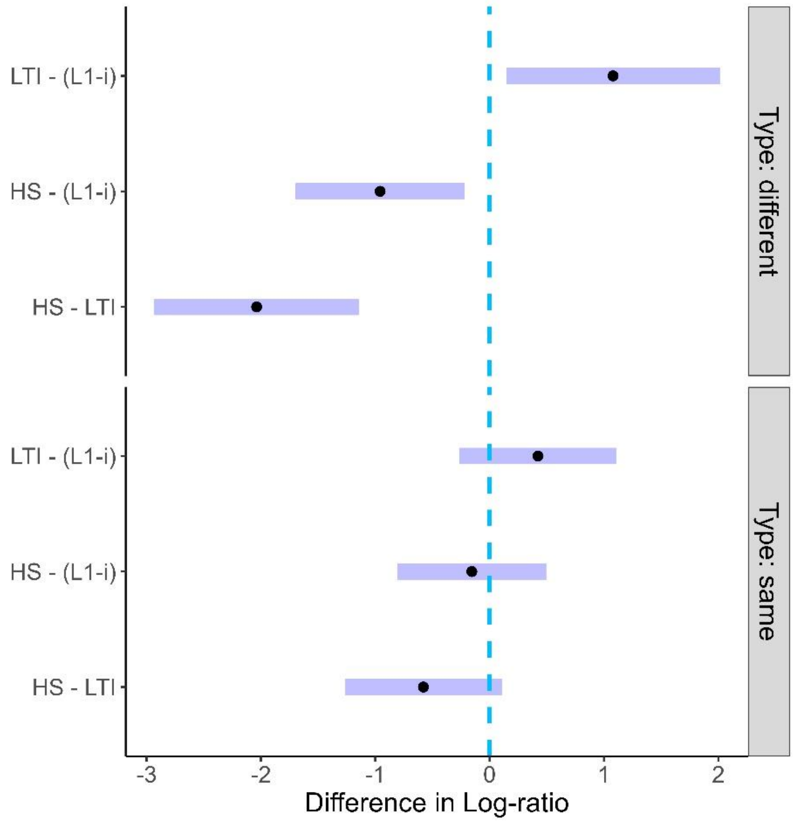


Figure 3.8. Visualization of estimated mean differences in log-odds ratio between participant groups by Trial Type, derived from the mixed-effects logistic regression model.

Zero on the x-axis indicates no difference between the two compared groups. The dashed line at zero serves as a reference point for visual interpretation: differences above zero indicate the first group named in the comparison outperformed the second, while differences below zero indicate the opposite.



Effect of cue weighting

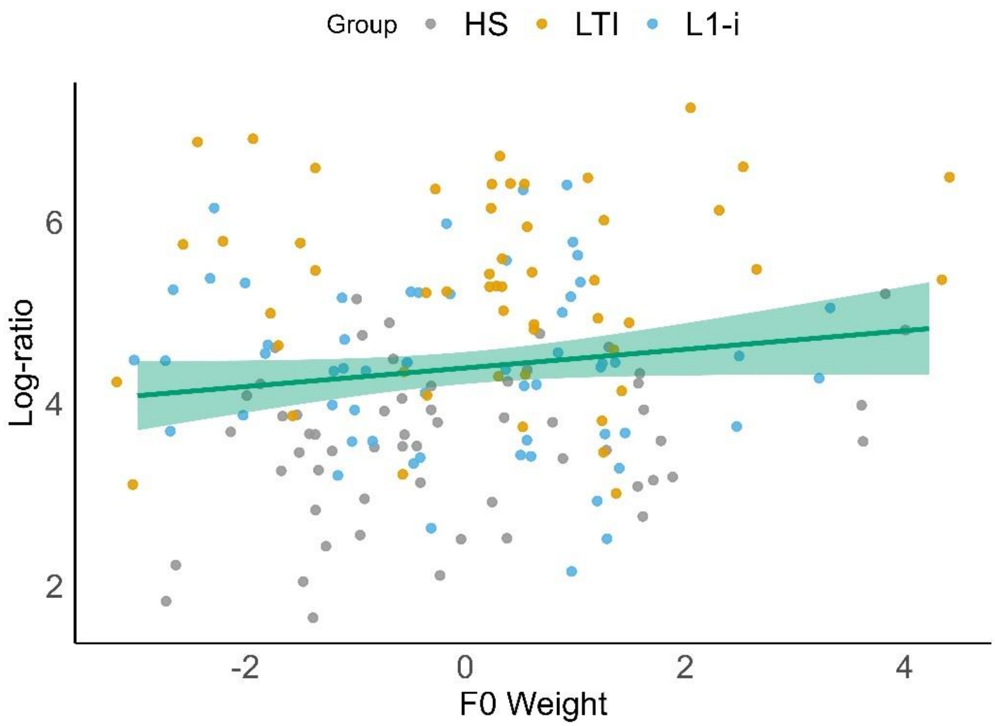
Another key objective of the AX discrimination task was to investigate the impact of participants' relative weighting of onset f0 as a cue in the 3AFC identification task on their ability to discriminate the lenis-aspirated stop contrast. Essentially, the model examined the extent to which individual dependency on the tonal cue against the VOT cue influenced their perceptual accuracy when discriminating between lenis and aspirated stops with f0 as the only cue available.

The odds ratio of correct responses, in correlation with f0 cue weight (the main effect), was computed by doubling the beta coefficient (β = 0.07), as the sum-coded Trial Type effect encompassed two levels ('same' and 'different'). The results demonstrated that participants' odds ratio of correct responses significantly increased by a factor of 1.15 for each unit increase in the relative f0 weight (β = 0.07, SE = 0.03, z = 2.18, p = .029), suggesting that participants who assigned greater weight to onset f0 in the identification of aspirated vs. lenis stops demonstrated increased accuracy when discerning between lenis and aspirated stops in the AX discrimination task. Figure 3.9 visually represents this correlation between f0 weight and the predicted accuracy of discrimination in different trials, based on each participant's random effects and corresponding fixed effects. The figure exhibits a positive slope for discrimination probabilities in relation to the f0 weight, averaged for Group. This implies that the relative f0 cue weight, estimated by the categorization of aspirated stops over lenis stops in the 3AFC task, had a positive impact on their perceptual accuracy in discriminating between lenis and aspirated stops by increasing the estimated likelihood of accurate response.

To evaluate the statistical power of the relatively small beta coefficient (β = 0.068) of the fixed interaction factor of Trial Type with f0 weight, a simulation-based power analysis was performed using the 'simr' package (Green & MacLeod, 2016) in R, simulating the model 1,000 times. The results of the power analysis confirmed that the interaction predictor of Trial Type*f0 weight surpassed the robust power threshold of 80%, reaching 99.2% (98.43% – 99.65% with a 95% CI).

Figure 3.9. Scatter plot with a regression line showing the relationship between F0 weight and predicted log-odds of correct discrimination in different-word trials, averaged for Group.

Each data point represents an individual participant's predicted log-odds from the mixed-effects logistic regression model, adjusted for their respective group.



Reaction time (RT)

Further insight into the effect of CLI on the perceptual discrimination of the lenis-aspirated stop contrast was sought through a separate mixed-effects linear regression analysis of the log-transformed RT as the dependent variable. This analysis was limited to 'correct' responses within 'different' trials—a total of 8,902 observations. The model included group (L1-i: reference level, HS, and LTI), f0 weight, and their interactions as fixed effects while including by-subject and by-item random intercepts. As shown in Figure 3.10, which represents the estimated RTs for each group and participants, the HS group demonstrated longer RTs in comparison to the L1-i group, on average by 201 ms (0.25 on the log scale) ($\beta = 0.25$, SE = 0.063, $t(78.00) = 4.02$, $p < .001$). This suggests that HSs, when relying solely on the onset f0, required more time to accurately discriminate between lenis and aspirated stops than L1-i speakers. In contrast, the RT of the LTI group was not significantly different from that of the L1-i group ($\beta = 0.08$, $p = .25$). A subsequent post-hoc analysis, using the Tukey multiplicity adjustment method, revealed that the HS group was also significantly slower than the LTI group, by 151 ms on average (0.18 on the log scale) in discriminating the lenis-aspirated stop contrast ($\beta = 0.18$, SE = 0.07, $t(78) = 2.73$, $p = .02$). These patterns were also visually confirmed that LTI participants were on par with L1-i speakers, as shown by the similar distribution of their RTs in Figure 3.10. Each data point in this figure represents the mean estimated RT for an individual participant's 'correct' discrimination of the lenis-aspirated stop contrast, computed by combining the fixed effects with the by-subject random effect estimates.

In order to examine whether HSs' longer RTs are attributable to the effect of lexical frequency, a supplementary mixed-effects linear regression analysis was conducted, with Frequency balance ('balanced' and 'unbalanced' lexical frequency within the stimuli pair) as a fixed effect and its interaction with group, following the equivalent model structure as the original model. The results showed no significant effect of Frequency balance ($p > .05$).

There was no effect of f0 weight on L1-i speakers' RTs ($p = .9$) and no significant interaction with group (HS: $p = .2$, LTI: $p = .7$). Table 3.3 provides a summary of the mixed-effects linear regression analysis of RTs in the AX discrimination task.

Figure 3.10. Model-predicted, log-transformed RTs, adjusted for corresponding group fixed effects across participants.

Each data point represents the mean estimated RT for an individual participant's correct discrimination of the stop contrast. The dashed line at y = 6.6 indicates the model's estimated baseline RT (L1-i speakers), a reference for interpreting the relative position of individual predictions. The significance is based on the mixed-effects linear regression model analysis and the subsequent post-hoc analysis.

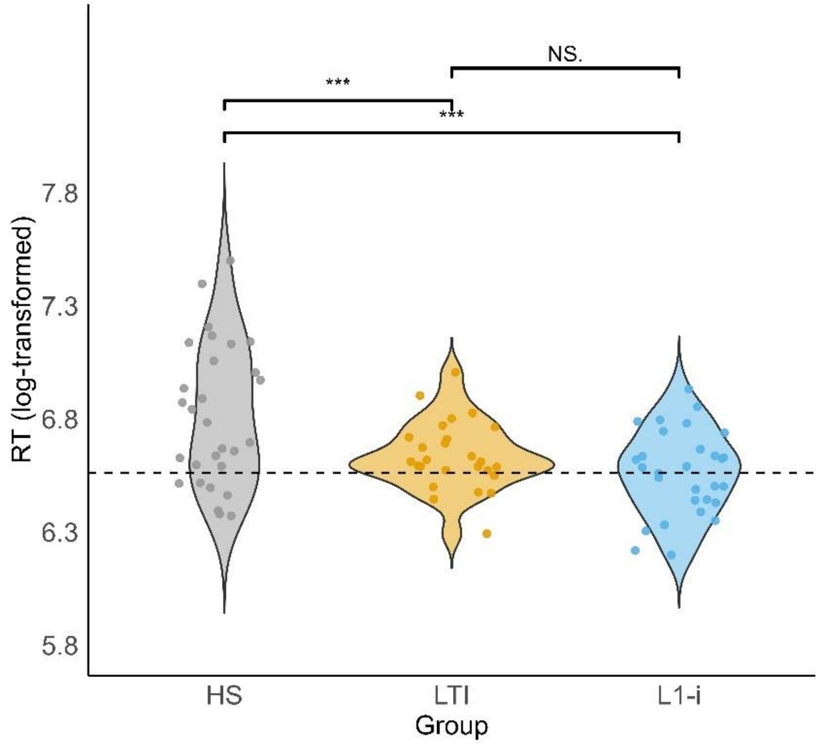


Table 3.4. A summary of the mixed-effects linear regression model examining RTs in the AX discrimination task.

	Estimate (SE)	*t*	*p*
Intercept	6.563 (0.044)	147.496	< .001
GroupHS	**0.253 (0.063)**	**4.018**	**< .001**
GroupLTI	0.075 (0.065)	1.156	.251
F0Weight	0.003 (0.030)	0.113	.910
GroupHS:F0Weight	−0.056 (0.041)	−1.369	.175
GroupLTI:F0Weight	−0.015 (0.042)	−0.350	.727

3.3 Interim discussion

The current study established three primary lines of inquiry that probed the L2 (English) CLI on L1 (Korean) perception, utilizing 3AFC and AX discrimination tasks. First, the study examined variability across groups in the perceptual cue weighting of onset f0 and VOT when identifying Korean stops. Second, it delved into the ability of HSs and LTIs to perceptually differentiate between lenis and aspirated stops when the contrast was signaled solely by onset f0. Third, the investigation sought to ascertain the potential correlation between an individual's sensitivity to the tonal cue as opposed to the VOT cue in L1 categorical perception and their capacity to discriminate a tonal-based L1 stop contrast. Addressing these research questions, the current study conducted a detailed examination of both categorical and discriminatory perception of L1 Korean stops among HSs and LTIs residing in the USA. These Korean–English bilinguals, characterized by their diverse language experiences, were contrasted with Korean-dominant native speakers in South Korea, L1-i speakers. The following sections discuss how the findings of the present study address each research question.

3.3.1 L2 CLI on L1 cue weighting

The examination of L1-i speakers' perceptual cue weighting demonstrated that onset f0 indeed plays a primary role with VOT being a secondary cue in the perception of lenis-aspirated and lenis-fortis stop contrasts, which is compatible with the results of the previous studies that examined L1-i speakers' perception of Korean stops (Kim, 2023; Kim et al., 2002; Kong, 2012; Kong et al., 2011, 2022; Kong & Kang, 2023; Kong & Yoon, 2013; Lee et al., 2013; Schertz et al., 2015). The results of the analysis of HSs and LTIs revealed that both bilingual groups differed from L1-i participants in their perceptual treatment of Korean laryngeal contrasts, but not in the same direction. In the identification of the lenis vs. aspirated stops, as predicted, HSs displayed a reduced reliance on onset f0 compared to L1-i speakers, while the LTI group unexpectedly demonstrated a greater sensitivity to the f0 cue compared to the L1-i group. This pattern was repeated in the cue weighting of lenis vs. fortis stop pairs, where the HS group showed a decreased dependence on the onset f0 cue compared to the L1-i group, while the LTI group exhibited increased reliance on onset f0 compared to the L1-i group. HSs were further characterized by their distinctive use of the VOT cue, as they tended to identify stimuli with long VOT as fortis stops, in contrast to LTI and L1-i speakers. For fortis and aspirated stop pairs, the HS group also displayed a unique cue-weighting pattern by attending to a tonal distinction between the two types of stops, while the LTI and L1-i groups ignored f0 differences for this contrast. The overall pattern that emerged from these observations is that HSs often employed distinctly non-Korean-like cue-weighting strategies in identifying Korean stops, while LTIs either matched or surpassed L1-i participants in their adherence to prototypical Korean cue weighting.

Therefore, the hypothesis that both groups of bilinguals would demonstrate a decreased reliance on onset f0 and an enhanced reliance on VOT as a result of assimilatory CLI from English was supported only partially. While HSs did in fact rely on f0 less than L1-i participants in two out of three pairwise contrasts, the LTI group was more sensitive to onset f0 than L1-i participants. The increased reliance on VOT did not consistently manifest itself in either of the two groups.

Nevertheless, there was a clear divergence in the perceptual behavior of the bilingual groups from that of the L1-i participants. The pattern of divergence is most compatible with the assumption of the assimilatory L2 English on L1 Korean CLI in HSs and the dissimilatory CLI in LTIs. As a result of assimilation to English, HSs began using f0 less in identifying Korean laryngeal categories. As a result of dissimilation from English, LTIs began relying on f0 more, applying an 'overly' Korean strategy in listening to Korean stops. It is interesting that the locus of CLI in both groups was the use of onset f0 rather than VOT, possibly because f0 is the cue that is used divergently in the two languages, while VOT plays a similarly important role in cuing laryngeal contrasts.

The divergent behavior of the two bilingual groups is a striking demonstration of the fact that patterns of L1 input can play a decisive role in determining the interaction between the sound systems of L1 and L2. In SLM-r terms, despite being on average more Korean-dominant and less proficient in English than HSs, LTIs must have received sufficient L1 and L2 input to create distinct stop categories for both languages, so much so that a dissimilatory CLI became possible. However, the present study acknowledges that without testing participants' English, this conclusion is tentative. Nevertheless, SLM-r posits that at an advanced stage of bilingual language acquisition speakers separate L1 and L2 categories, which can result in greater phonetic distance between L1 and L2 categories (Flege & Bohn, 2021). Given that LTI participants in the present study were more balanced than HSs in terms of language dominance, this finding is also congruent with earlier research indicating that more balanced bilingual speakers were better able to separate

L1 and L2 phonetic systems (Barlow et al., 2013; Guion, 2003; Kang & Guion, 2006; MacLeod et al., 2009; Sundara et al., 2006).

In contrast, HSs, who must have had at least as much or even more authentic L2 input throughout their lives as LTIs, were potentially less likely to have been provided with similarly ample and diverse L1 input as LTIs. As a result, HSs' L1 sound categories could, hypothetically, be less robust. The development of distinct L2 and L1 categories could be hindered in this case, creating conditions for assimilatory CLI. Thus, despite being L2-dominant and highly L2-proficient, HSs were predicted to be less able to separate the two languages and were more likely to assimilate them in the perceptual domain. Additional analysis also demonstrated that greater L1 (Korean) dominance (which for HSs meant more balanced bilingualism) was associated with greater reliance on onset f0 in categorizing lenis vs. aspirated stops among HSs and LTIs. Specifically, the analysis revealed that the weight of onset f0, which is the primary cue to the lenis-aspirated stop in L1 Korean, increased in the function of L1 dominance. This finding provides evidence for the crucial role of the quantity and quality of L1 input in determining the degree of L2 CLI on L1 speech categorization among bilingual speakers.

To summarize, somewhat unexpectedly, the results of the present investigation uncovered evidence in favor of dissimilatory perceptual CLI in LTIs—a population for which numerous production studies indicated assimilatory changes while evidence of dissimilation remained scarce (Flege & Eefting, 1988; Flege et al., 2003; Mack, 1990; Yusa et al., 2010). In parallel to dissimilation in production, perceptual dissimilation can be thought of as an expression of the need to keep the two sound systems distinct. For example, in production, the f0 properties of Korean stops could be hyperarticulated in order to emphasize the difference from English stops. For instance, Kang and Nagy (2016) demonstrated that Korean HSs in Canada made an enhanced onset f0 distinction between lenis and aspirated stops, compared to homeland speakers in South Korea. In parallel, in the perception of Korean stops, the reliance on f0 can be exaggerated as a way to maintain a perceptual distinction from English stops.

The results for HSs, in contrast, suggested the effect of assimilatory CLI from English, as the reliance on f0 was, on the whole, less pronounced for this group. However, there were two specific findings that did not fit into this general picture. First, HSs unexpectedly exhibited a small but appreciable degree of reliance on f0 in the identification of fortis vs. aspirated stops, the cue which under normal circumstances is not informative for this contrast in Korean and was, in accordance, completely unattended to by LTI and L1-i participants. This pattern underlines the trend of a distinctly non-Korean-like perceptual treatment of laryngeal contrasts by HSs. While it is not clear what prompted this unusual strategy, one possible explanation is that the overall pattern of laryngeal cueing in English served as a basis for this tendency. In Korean, onset f0 is a strongly predictive cue for some laryngeal contrasts (e.g., lenis vs. aspirated) and a completely uninformative one for others (e.g., fortis vs. aspirated); hence, listeners learn to turn their use of this cue on and off. In English, f0 is consistently present as a secondary but moderately useful cue to word-initial voicing distinctions (Abramson & Lisker, 1985; Dmitrieva et al., 2015; House & Fairbanks, 1953; Llanos et al., 2013; Ohde, 1984; Shultz et al., 2012; Whalen et al., 1993). Thus, an appropriate strategy for English would be a constant, albeit limited reliance on f0. This is, to some extent, exactly the pattern observed for HSs.

The second unexpected pattern is the relatively strong effect of VOT among HSs in the identification of lenis vs. fortis stops, combined with a tendency to classify some long VOT stops as fortis—a pattern conspicuously absent from LTI and L1-i participants' responses. The present study does not currently have a compelling explanation for this approach, beyond the possibility that HSs, who were presumably more reliant on English-like strategies of laryngeal identification, experienced some confusion with respect to the covariation between f0 and VOT in Korean. In English, longer VOT reliably covaries with higher onset f0—an association that is violated by both lenis (long VOT but lower f0) and fortis (short VOT but higher f0) stops. Recently, Perrachione et al. (2023) suggested that this unusual relationship between VOT and f0 could be a source of difficulty for English native speakers in acquiring Korean stops. This mismatched pattern of f0-VOT covariation could have had a similar effect on HSs who were English dominant.

3.3.2 L2 CLI in L1 perceptual discrimination

Another goal of the present investigation was to examine the bilinguals' use of onset f0 in discriminating Korean stops for which f0 acts as a primary distinguishing cue—lenis and aspirated stops. The results of this experiment, where the available useful cues were artificially reduced to f0 only, showed that LTI outperformed L1-i participants in terms of accuracy, while HSs were less accurate than the L1-i group, partially confirming the second hypothesis. Moreover, individual strength of reliance on onset f0 in the identification test was predictive of discriminatory performance, confirming the third hypothesis. Finally, analysis of RTs confirmed that HSs found the discrimination task relatively challenging, as evidenced by this group's longest RTs. In contrast, LTI's RTs were not significantly different from those of L1-i participants, suggesting that their optimal accuracy performance did not come at the expense of greater effort.

These results, first of all, demonstrate a link between participants' cue-weighting patterns and their discriminatory abilities, both at the group and at the individual levels, aligning with previous studies indicating that auditory acuity is related to phonetic categorization or precision (Franken et al., 2017; Hazan et al., 2010; Kachlicka et al., 2019; Perkell et al., 2004; Shultz et al., 2012). Those groups and individual participants who assigned lower weight to f0 as a cue to lenis-aspirated distinction in the identification experiment were also less successful in using this cue to discriminate between lenis and aspirated stops. The crucial implication of this finding is that CLI on L1 cue-weighting patterns has far-reaching consequences, affecting different aspects of L1 speech perception. Specifically, HSs, who were shown to be overall less reliant on f0 in identifying Korean stops, were also less able to use f0 in discriminating among Korean stops, especially if other cues, such as VOT, were unavailable.

Moreover, these results shed additional light on some previous findings in research on HSs' L1 perception, particularly explicating the apparent lack of significant differences between L1-i populations and HSs. A key distinction between the present study and relevant previous research lies in the stimuli design. For instance, previous discrimination studies (Oh et al., 2003, 2010; Seo et al., 2022) employed different-word stimuli that contrasted not only in terms of onset f0, but also in the naturally co-varying secondary cue, VOT. HSs performance that was on par with L1-i participants in these studies could be attributed to their skillful use of the secondary VOT cue, especially given HSs' dominance in English and VOT's primary role in cueing English laryngeal contrast. The results of the present investigation suggest that in the absence of this cue, HSs do not necessarily keep up with L1-i listeners in terms of discrimination accuracy, despite the fact that HSs' overall performance was highly accurate.

While the conditions of the present experiment were somewhat artificial, given that the naturally occurring VOT cue was removed, in natural settings, noise and other distractions may mask certain cues more than others, potentially creating conditions similar to those used in the discrimination experiment. The findings suggest that such a perceptual scenario could present a challenge for HSs whose dominant language is English. While HSs were able to accurately distinguish between lenis and aspirated stops based on onset f0, this accuracy came at the expense of longer response time. The increase in HSs' response time could potentially become even more significant in more casual, fast-paced communication scenarios, especially when contextual discourse cannot resolve the ambiguity, potentially affecting the quality of speech communication. Nevertheless, the current study acknowledges the limitation of the AX discrimination task and the 3AFC paradigm designed to assess discrete speech perception. Research has shown that listeners are sensitive to fine-grained acoustic details in phonetic category perception and that category perception operates on a gradient rather than a discrete scale (Kapnoula & McMurray, 2021; Kapnoula et al., 2017; Massaro & Cohen, 1983; McMurray et al., 2002). For instance, Kapnoula et al. (2017) showed that listeners relied on both onset f0 and VOT cues in a continuous manner in the perception of English voicing (Kapnoula et al., 2017). Therefore, future research can benefit from the use of a VAS task, which adequately captures the gradient nature of speech perception (Kutlu et al., 2022).

3.3.3 Alternative explanation for HSs' results

Findings of the present study have been interpreted as evidence of CLI; however, there is a potential alternative explanation that merits consideration. The divergence of HSs from LTI and L1-i participants could result from distinct L1 experience. Specifically, there is a possibility that HSs' perceptual behavior is a result of having learned a different 'variety' of Korean—a variety that did not undergo the sound change as in the Seoul-Gyeonggi variety. As a specific hypothetical case, it is plausible that some of their parents may speak the Gyeongsang variety of Korean in which VOT is a primary cue to the lenis-aspirated stop contrast in perception (Kong et al., 2022; Lee et al., 2013). Although the study took care to enroll participants whose parents were born no more than a decade before the purported onset of the sound change, it is nevertheless possible that they were not exposed to the sound change sufficiently in order to acquire it and transmit it to their children. If this were the case, HSs in the present study would be expected to rely primarily on VOT for all Korean laryngeal distinctions. For practical reasons, the investigator could not collect sound samples from participants' parents, and, therefore, the study cannot resolve this issue conclusively. Nonetheless, it is believed that the complete pattern of the findings is not fully compatible with this possibility. First, HSs relied on f0 to a great extent, although less so than L1-i listeners, therefore not supporting the possibility that f0 was not an important cue for them. Second, HSs outperformed L1-i listeners in using f0 as a cue to the fortis-aspirated distinction—a distinct behavior that nevertheless does not fit into the pattern of the overall unimportance of f0 as a laryngeal cue.

To conclude, although this possible alternative explanation is not perfectly predictive of the specific pattern of results obtained in the present study, we cannot rule out the possibility that distinct L1 experience has contributed to some extent, in conjunction with the CLI from English.

4. CROSS-LINGUISTIC INFLUENCE IN L1 PRODUCTION

4.1 Methods of Experiment 3: Controlled reading paradigm

4.1.1 Participants

The same three groups of participants who took part in the perception experiments also participated in the production task. For analysis, 7 recordings were excluded due to technical issues and volunteer withdrawal, resulting in a total number of 79 participants (HS: n = 27, LIT: n = 23, L1-i: n = 29).

4.1.2 Materials

A controlled reading paradigm was conducted to elicit participants speech production of Korean stops. This task involved 54 unique words as stimuli, which consisted of 18 triplets potentially differing in the word-initial stop (lenis-aspirated-fortis) (e.g., /kɨda/, 'to line,' /k^{h}ɨda/, 'big,' /k*ɨda/, 'to turn off'). Target triplets included monosyllables, disyllables, and trisyllables. The stops in the stimuli were balanced across three different places of articulation (bilabial-alveolar-velar) and were followed by one of the six different vowels (/a/, /i/, /u/, /e/, /o/, /ɨ/). All lenis and aspirated stop tokens were real words, while 14 out of 18 fortis stop tokens were nonce words due to lexical gaps. The stimuli in Korean orthography, along with IPA transcriptions and English translations, are listed in Appendix B.

4.1.3 Procedures

The recordings were conducted in the university speech labs (at three different locations: in Seoul, Illinois, and Indiana) or in a quiet location chosen by the participants, and in accordance with the safety guidelines related to the COVID-19 Pandemic. While this decision led to the use of different equipment for the recording process, all the equipment used was of professional quality[3]. Recorded audio files were digitized at a sampling rate of 44.1 kHz, quantized at 16 bits.

[3] Shure KSM32 Embossed Single-Diaphragm Cardioid Condenser Microphone, Shure Head-Mounted SM10A Microphone, Audio-Technica Handheld Cardioid Dynamic Microphone, ART Tube MP Project Series preamplifier, Marantz Professional Solid State recorders (PMD 661 MKII, PMD660)

Before performing the task, participants received both oral and written study instructions, mostly in Korean, although English was also used for English dominant participants, and signed the consent form. The task was administered using PsychoPy3 (Peirce et al., 2019). Participants sat in front of a computer screen where each stimulus was presented in the following carrier phrase "/taɨm tanʌnɨn _____ imnida/" 'The next word is ______." Participants were instructed to read each sentence aloud into the microphone positioned at a distance of approximately 7.8 inches (20cm). The recording began with five practice trials as a familiarization phase. Upon completing a sentence, participants were able to move to the next trial by pressing a spacebar on a keyboard. The keyboard did not function for the first two seconds of each trial to prevent participants from advancing to the next trial by mistake. The 54 unique stimuli were repeated three times, resulting in a total of 162 trials. The recording took approximately ten minutes to complete.

4.1.4 Analyses

The current study measured VOT using a machine learning-based automatic speech processing tool, Dr. VOT (Shrem et al., 2019). This system is built on a bidirectional recurrent neural network function, which has been proven effective in measuring temporal elements in speech recognition and natural language processing (see Graves et al., 2013; Mikolov et al., 2010 for more information). Dr. VOT automatically marked and annotated both the release of stops and the onset of the following vowels. All resulting data processed with Dr. VOT were manually checked for errors and adjusted to ensure measurement accuracy. Onset F0 was measured at the onset of the following vowel following a word-initial consonant at the first zero-crossing point where pitch was detected, using a custom script in Praat (Boersma & Weenink, 2022). Measured onset F0 was converted from Hz measure to the semitone scale in order to transform the measurements to relative difference of each data point from the mean of each participant's onset F0 across all stimuli on a log-scale (Kang, 2014; Lee & Jongman, 2019). The conversion formula used is as follows: $12 \times \log_2 \frac{x}{m}$ where x and m refer to each data point and individual mean, respectively (Dmitrieva et al., 2015; Shultz et al., 2012). Among the total number of 12,798 data points, 258 and 21 audio files were excluded from the VOT and onset F0 models, respectively, due to technical issues in the recordings.

For statistical analysis of the processed variables, onset F0 (on a semitone scale) and VOT (in ms), the current study implemented four Bayesian mixed-effects linear regression models, with onset F0 and VOT as the dependent variables. The first two models examined the effects of participant group and stop type by including Group (L1-i: reference, HS, LTI), Type (aspirated: reference, fortis, lenis), and their interaction as fixed effects while incorporating by-subject random intercepts and slopes for Stop Type and by-word intercepts as random effects. Therefore, not only did the models predict group-averaged effects of stop type on VOT and onset F0, but they also allowed for predictions of individual variability by stop type in each group in their implementation of these acoustic correlates. The current study adopted weakly informative priors for both models, using normal distributions for both intercepts and fixed effects, such that estimated coefficients were set to approach 0 unless the models found sufficiently large effects, with priors having little influence on the posterior distribution (Albert & Hu, 2019; McElreath, 2020). Since the intercepts of the models equaled the mean of aspirated stops produced by L1-i speakers (reference level), priors were determined based on a recent production study of corpus data (Bang et al., 2018). Specifically, the priors for the intercepts in the models had the means of 1 (semitone) and 70 ms (VOT) with the *SD* of 3 (semitone) and 70 ms (VOT), respectively. The priors for the means and *SDs* of the beta coefficients were 0 and 3 (semitone) and 0 ms and 70 ms (VOT), respectively.

To assess the influence of bilingual dominance on the production of Korean stop consonants (onset F0 and VOT) among bilingual speakers, two other Bayesian mixed-effects linear regression models were constructed for participants in the HS and LTI groups. These models, with semitone and VOT as the respective dependent variables, employed identical priors and sampling methods as the first two models. The study measured participants' language dominance, using MINT Sprint (Garcia & Gollan, 2021) in which participants named 80 items on a single screen in each language. MINT Sprint scores for each language were standardized to *z*-scores to calculate a language-independent measure of proficiency. The differential proficiency and dominance between English and Korean were quantified by subtracting English *z*-scores from Korean *z*-scores. Consequently, a negative value of the difference in MINT scores between the two languages (MINT_Diff here and after) indicates high English dominance, whereas a positive value indicates high Korean dominance. Scores near zero represent a state of balanced bilingualism. Each model incorporated Stop Type and MINT_Diff, along with their interaction, as fixed effects. In addition, random effects included by-subject intercepts and slopes for Stop Type and by-item intercepts.

Consequently, these models enable a comprehensive examination of how bilingual dominance influence the acoustic correlates of VOT and onset F0, while accounting for individual variability.

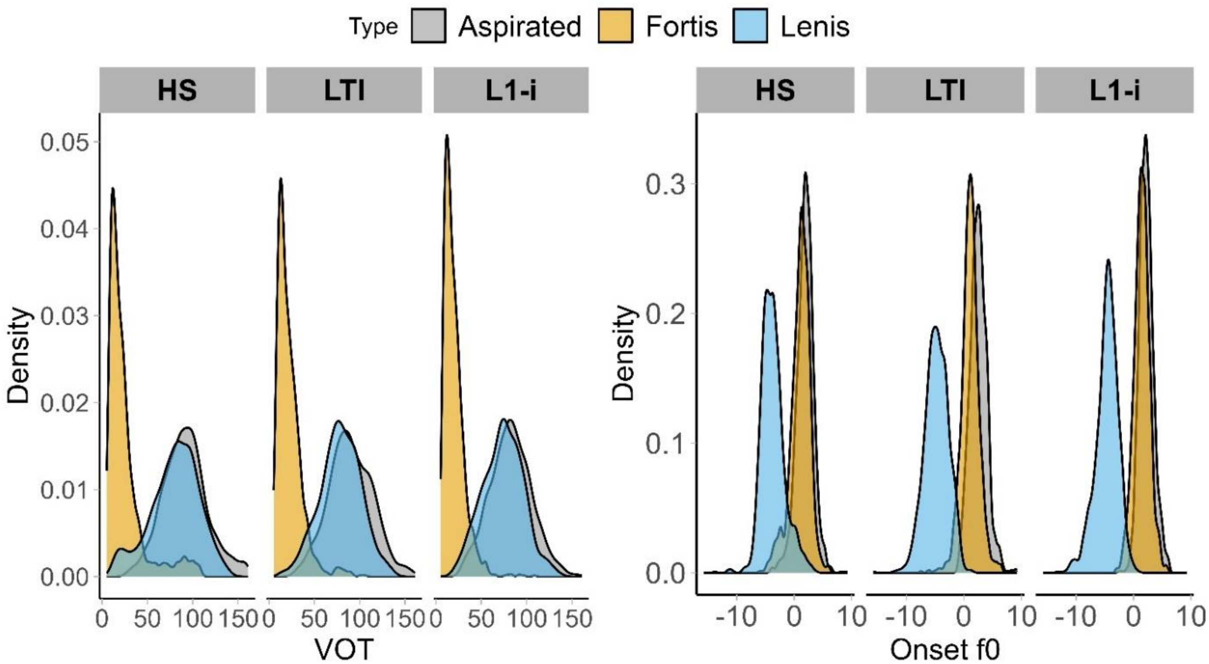
All statistical analyses were conducted, using R version 4.2.1 (R Core Team, 2022), and the Bayesian models were fitted by the Hamiltonian Markov Chain Monte Carlo (MCMC) sampler function in Stan (Carpenter et al., 2017), using the 'brm' package (Bürkner, 2017). To obtain posterior samples, 4 MCMC chains were used to draw 4,000 samples with 1,000 warm-up iterations in each model. The current study reports posterior distributions with 95% credible intervals (CI) for each estimate. The study reports the credibility of estimated effects through the index of Probability of Direction (pd), also referred to as Maximum Probability of Effect. pd is the probability ranging between 50% and 100%, which represents the probability of direction (positive or negative) of a given parameter. A pd of 50% indicates a posterior of zero, suggesting no evidence for an effect on the dependent variable while a pd close to 100% supports evidence for an effect with either positively or negatively skewed distributions. For instance, if 95% of the posterior distribution of a certain effect is skewed from 0 to either the negative or positive direction, it is inferred that the effect indeed exists with a probability of 95%. The current study views pd of 95% and above with the CI of 95% as convincing evidence for estimated effects. Although the choice of index to be used for estimating credibility of estimated effects remains disputed due to the complex influence of priors on the posterior distribution, pd is known to be an easily interpretable measure, intuitively comparable to the *p*-value in the frequentist framework, helping readers who are more familiar to the frequentist framework to readily interpret the effects in the Bayesian models (see also Makowski et al., 2019a, 2019b for more discussion).

4.2 Results

4.2.1 Acoustic characteristics of Korean stops

Figure 4.1 illustrates the distributions of VOT and onset F0 across groups and stop types, based on the raw data. Across all groups, a clear VOT distinction is noted, contrasting fortis stops from both aspirated and lenis stops, both of which display longer VOTs. Furthermore, almost completely overlapping VOT distributions for lenis and aspirated stops in all speaker groups reflect an ongoing VOT merger between these stop types. Regarding this stop contrast, each group indeed exhibits an onset F0-based contrast, with lenis stops produced with lower onset F0 values compared to aspirated and fortis stops. Notably, the LTI group shows a greater reliance on onset F0 in the production of aspirated stops, indicated by a more notable offset in its distributions of the parameter compared to lenis stops. The following sections report the relevant statistical results in more detail.

Figure 4.1. Density plots depicting the distributions of VOT and onset F0 (semitone) by group and stop type, based on the raw data.



4.2.2 Variability in VOT by group and stop type

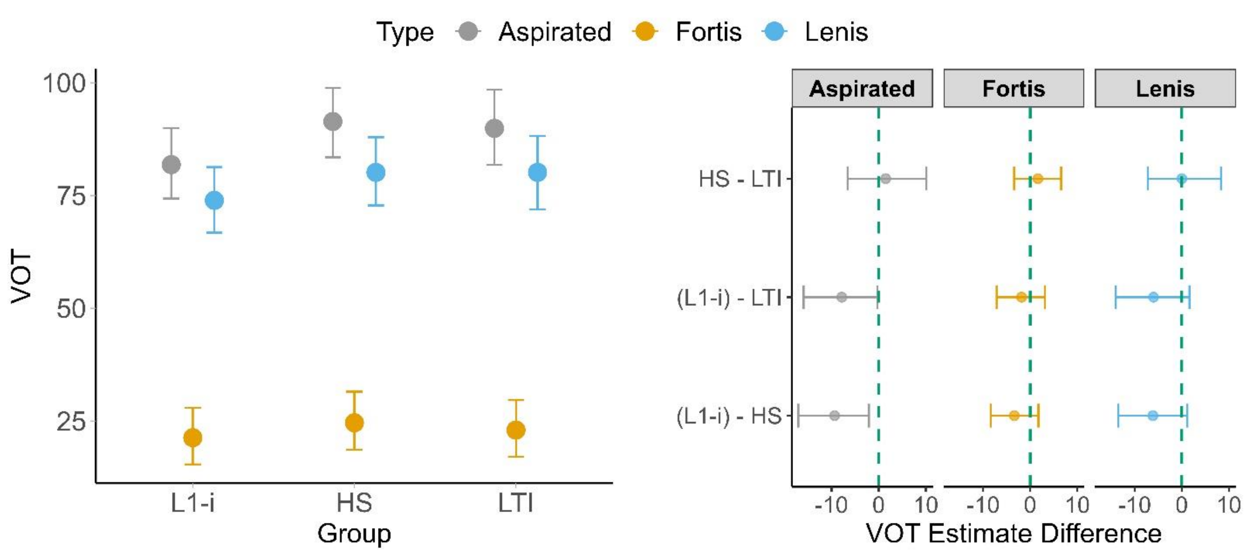
Table 4.1 provides the results of the Bayesian mixed-effects linear regression model examining VOT. The left panel of Figure 4.2 represents the estimated conditional effects, showing that both HS and LTI groups produce aspirated stops with longer VOT durations compared to the L1-i group, which has an average VOT duration of 81.97 ms. Specifically, the aspirated VOT of the HS group is extended by 9.37 ms (β = 9.37, 95%CI = [1.57, 16.71], pd = 98.95), and that of the LTI group by 7.92ms (β = 7.92, 95%CI = [0.04,15.81], pd =97.52). The model shows a credible effect for fortis stops within the L1-i group, with a decrease in VOT of 60.6 ms relative to aspirated stops, aligning with the characteristic short-lag VOT of fortis stops among L1-i speakers (β = −60.6, 95%CI = [−70.53, −51.20], pd = 100.0). A credible yet small effect is observed in the L1-i group for lenis stops, decreasing the VOT by 7.94 ms, compared to aspirated stops (β = −7.94, 95%CI = [−16.75, 0.86], pd = 100.0).

Table 4.1. A summary of the Bayesian mixed-effects linear regression model examining VOT.

	Estimate	Est.Error	L-95% CI	U-95% CI	Rhat[4]	pd
Intercept	81.97	3.95	74.32	89.92	1.01	100%
GroupHS	9.37	3.95	1.57	16.71	1.01	98.95%
GroupLTI	7.92	4.01	0.04	15.81	1.00	97.52%
TypeFortis	−60.60	4.96	−70.53	−51.20	1.01	100%
TypeLenis	−7.94	4.51	−16.75	0.86	1.01	95.95%
GroupHS:TypeFortis	−5.97	4.34	−13.88	2.77	1.01	90.83%
GroupLTI:TypeFortis	−6.11	4.40	−14.58	2.40	1.00	91.72%
GroupHS:TypeLenis	−3.24	3.47	−9.93	3.70	1.01	82.73%
GroupLTI:TypeLenis	−1.85	3.45	−8.86	4.82	1.01	70.20%

Figure 4.2. Conditional effects on VOT (left panel) and pairwise estimate differences (right panel), predicted by the Bayesian mixed-effects linear regression model.

The vertical and horizontal lines represent the 95% highest posterior density (HPD) intervals for each estimate.



[4] Rhat is a convergence diagnostic statistic measure reported in the brms output. Rhat is used to assess whether the chains in a MCMC simulation have converged to the same distribution. An Rhat value close to 1 (e.g., less than 1.1, preferably at 1.01 or less) indicates that the chains have converged, while larger values suggest that more samples are needed for convergence.

However, the statistical model did not detect any credible interaction effects between Group and Stop Type, with all pd values falling below the 95% threshold. This lack of credible interaction suggests that, except for aspirated stops, there are no group-specific differences in VOT for fortis and lenis stops, suggesting that the VOT differences between stop categories did not differ across groups. As shown in the left panel of Figure 4.2, both the HS and LTI groups follow a similar VOT pattern to that of the L1-i group, with aspirated stops having the longest VOT, followed by lenis with a small difference and by fortis stops with a great difference.

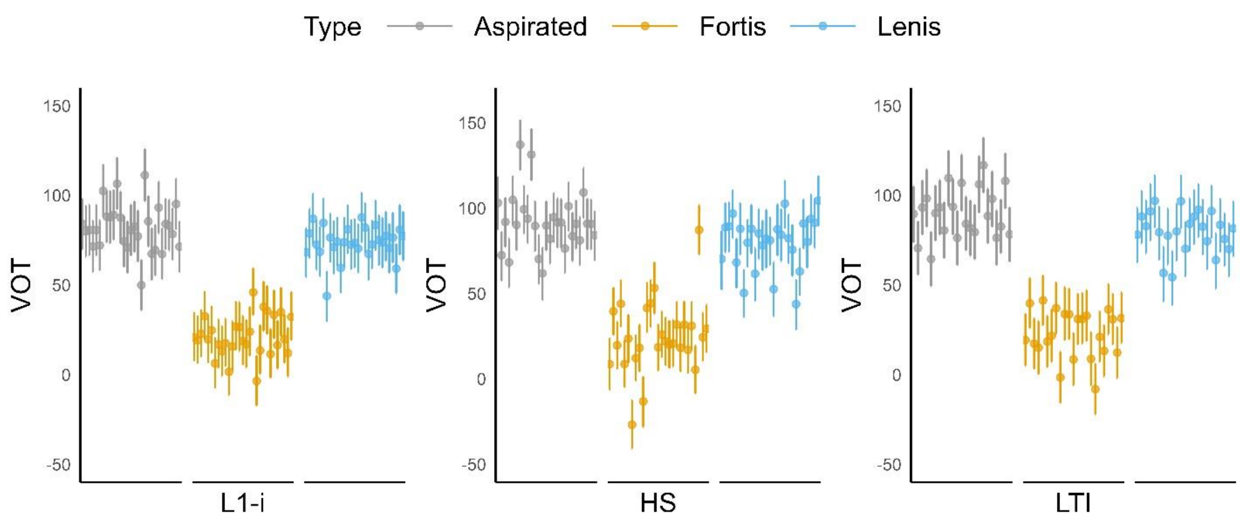
To further examine the results of the categorical factors of the model (especially to compare between the HS and LTI groups), a post-hoc pairwise analysis was conducted with a HPD interval probability of 95%. The right panel of Figure 4.2 represents the pairwise comparisons based on the estimates of the model and HPD intervals. The figure visually confirms the statistical result that the group variance of HS and LTI groups in VOT in comparison to the L1-i group is most pronounced in the production of aspirated stops. Furthermore, it is suggested that there were no notable differences in the implementation of VOT between the HS and LTI groups across stop types, also corroborated by the pds of the pairwise comparisons between the two groups (Aspirated: pd = 64.45%; Fortis: pd = 74.05%; Lenis: pd = 50.15%).

Figure 4.3, a caterpillar plot, shows the individual VOT estimates for Korean stops as produced by speakers across the three participant groups. The vertical lines on each data point represent the within-individual variability, visualizing the 95% HPD intervals for each participant, as predicted by the Bayesian model. This individual analysis reveals more details about the observed group differences. For instance, L1-i participants exhibited a relatively compact clustering of VOT estimates for lenis stops, compared to the HS and LTI group participants. In addition, striking individual variability is visible in the VOT estimates for fortis stops among HSs, with one of the participants producing fortis stops with almost the same length of VOTs as in aspirated and lenis stops. Regarding aspirated stops, it is visually discernable that LTI participants produced relatively longer VOTs compared to the L1-i group participants and that there are two HSs who were predicted by the statistical model to produce aspirated stops with considerably prolonged VOTs.

In summary, the examination of VOT using a Bayesian mixed-effects linear regression model revealed a credible group variability only in aspirated stops. Specifically, the model predicted that the HS and LTI groups produce aspirated stops with longer VOT durations than the

L1-I group. No credible interactions between group and stop type were found, indicating that the effect of stop type on VOT is consistent across groups, except for aspirated stops. All groups showed a similar pattern in their implementation of VOT for stops, with aspirated stops having the longest VOT, followed by lenis and fortis stops. Notably, the model detected individual variabilities in VOT estimates, particularly in the HS group.

Figure 4.3. Caterpillar plot representing the individual VOT estimates in production for Korean stops across groups, based on the Bayesian model prediction.
The vertical lines represent the within-individual variability with 95% HPD intervals.



4.2.3 Variability in onset f0 by group and stop type

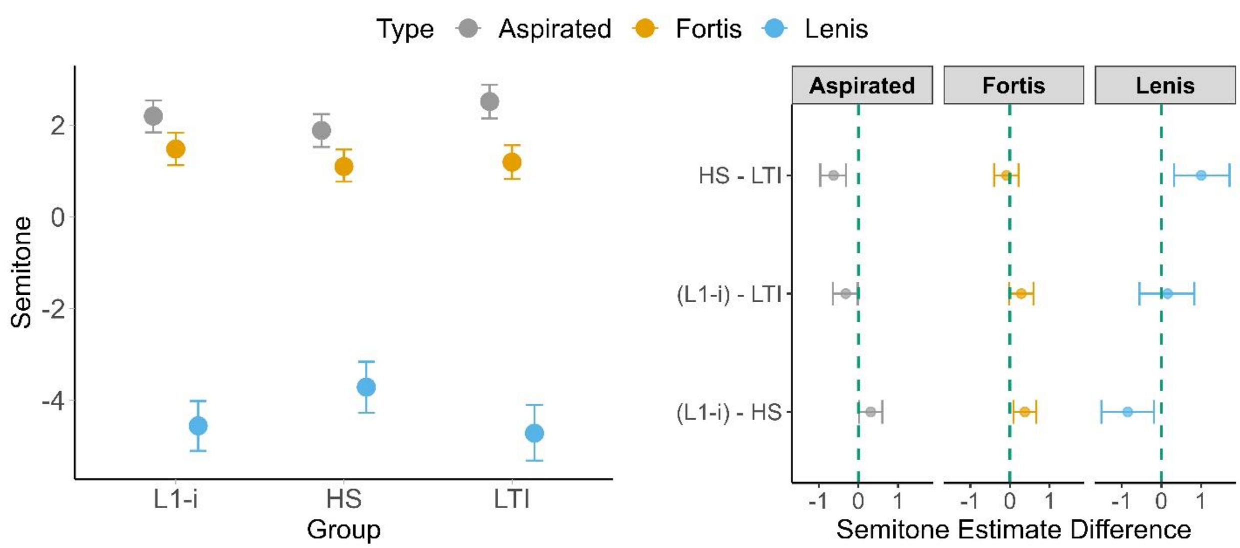
Table 4.2 provides the results from the statistical model examining the effects of Group and Stop Type on onset F0 converted to semitone. Regarding aspirated stops, the model indicated a credible decrease in semitone for the HS group by 0.31 compared to the estimate for the L1-i group with the average of 2.20 ($\beta = -0.31$, 95%CI = [−0.61, −0.01], pd = 97.90). In contrast, the LTI group showed an increase of 0.32 semitone for aspirated stops relative to the L1-i group, with the effect being credible ($\beta = 0.32$, 95%CI = [0.01, 0.64], pd = 97.82). The effect for fortis stops was found to be credible, showing a decrease in semitone by 0.72 compared to aspirated stops among the L1-i speakers ($\beta = -0.72$, 95%CI = [−1.19, −0.25], pd = 99.70). A substantial and credible decrease was also observed for lenis stops among L1-i speakers, with a decrease of 6.76

in semitone compared to aspirated stops ($\beta = -6.76$, 95%CI = [−7.51, −6.00], pd = 100.0). These results confirm the three-way tonal contrast in Korean stops among the L1-i speakers.

Table 4.2. A summary of the Bayesian mixed-effects linear regression model examining semitone.

	Estimate	Est.Error	L-95% CI	U-95% CI	Rhat	pd
Intercept	2.20	0.18	1.84	2.54	1.00	100%
GroupHS	−0.31	0.15	−0.61	−0.01	1.00	97.90%
GroupLTI	0.32	0.16	0.01	0.64	1.00	97.82%
TypeFortis	−0.72	0.24	−1.19	−0.25	1.00	99.70%
TypeLenis	−6.76	0.38	−7.51	−6.00	1.01	100%
GroupHS:TypeFortis	−0.07	0.18	−0.42	0.29	1.00	65.28%
GroupLTI:TypeFortis	−0.61	0.19	−0.98	−0.22	1.00	99.85%
GroupHS:TypeLenis	1.16	0.46	0.24	2.06	1.00	99.12%
GroupLTI:TypeLenis	−0.47	0.49	−1.46	0.47	1.00	82.67%

Figure 4.4. Conditional effects on onset F0 in semitone (left panel) and pairwise estimate differences (right panel), predicted by the Bayesian mixed-effects linear regression model.



Examining interaction effects, a notable discrepancy between aspirated and fortis stops was observed within the LTI group but not in the HS group (HS:Fortis: $\beta = -0.07$, 95%CI = [−0.42, 0.29], pd = 65.28; LTI:Fortis: $\beta = -0.61$, 95%CI = [−0.98, −0.22], pd = 99.85). This indicates that the difference in onset F0 between aspirated and fortis stops is greater in the LTI group than in the

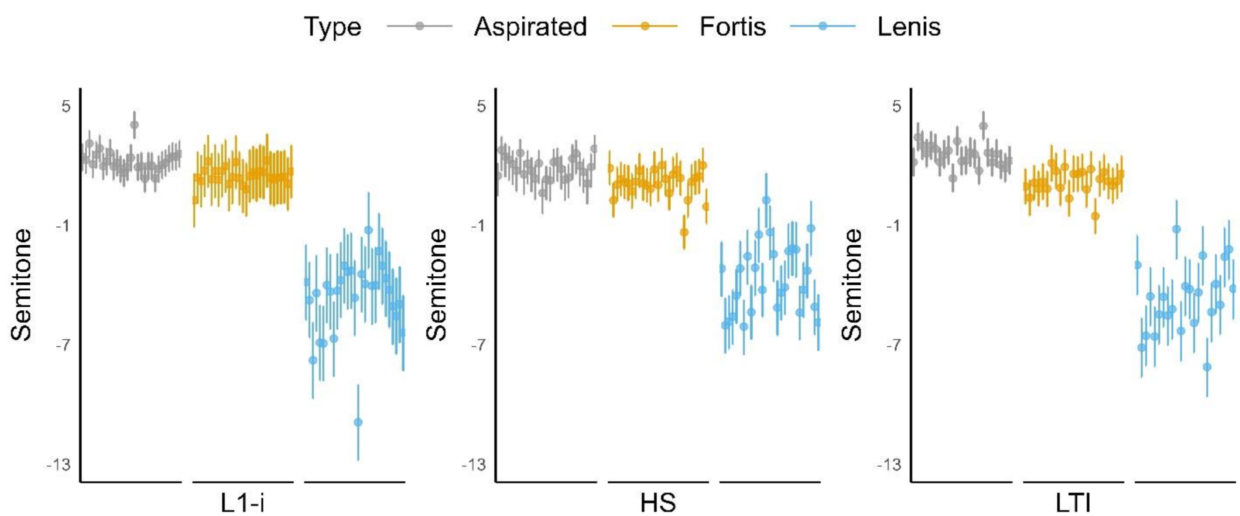
L1-i group, suggesting dissimilation between the two stop types. In contrast, the HS group's pattern of onset F0 for aspirated and fortis stops closely mirrored that of the L1-i group. The interaction effect between group and lenis stops revealed that the distinction via onset F0 between aspirated and lenis stops was less pronounced in the HS group than in the L1-i group, suggesting assimilation between the two stop types (HS:Lenis: $\beta = 1.16$, 95%CI = [0.24, 2.06], pd = 99.12). However, no such effect was observed in the LTI group (LTI:Fortis: $\beta = -0.47$, 95%CI = [−1.46, 0.47], pd = 82.67), patterning with the L1-i group. These statistical results are visually illustrated in the left panel of Figure 4.4.

The right panel of Figure 4.4 shows the results of post-hoc pairwise comparisons, based on the model estimates. The subsequent analysis revealed a credible decrease in semitone for aspirated stops in the HS group compared to the LTI group by 0.63 ($\beta = -0.63$, 95%CI = [−0.96, −0.32], pd = 99.98). Regarding fortis stops, there was no credible difference in onset F0 between the two groups ($\beta = -0.09$, 95%CI = [−0.40, 0.22], pd = 71.88). In contrast, the HS group exhibited an increased semitone for lenis stops relative to the LTI group ($\beta = 1.01$, 95%CI = [0.32, 1.72], pd = 99.75).

The caterpillar plot in Figure 4.5, which presents the individual analyses based on the HPD intervals, confirms the observed group variations, with increased onset F0 values in lenis stops among HSs comparable to aspirated and fortis stops within the group, compared to LTI and L1-i groups. More importantly, this figure presents insights into individual variability within each group. Notably, the L1-i group participants exhibit more densely clustered data points for both aspirated and fortis stops, as compared to the HS and LTI group participants, suggesting less individual variability in the use of onset F0 as a cue to the aspirated and fortis stop contrast. In contrast, the individual analysis revealed greater individual variability in onset F0 estimates among both HSs and LTIs, as shown in their widespread data points on the y-axis for all stop types, compared to L1-i speakers. The figure also visualizes a distinct pattern in the use of onset F0 for the aspirated-fortis stop contrast within the LTI group, with the majority of LTIs making a visible contrast via onset f0 between the two stop types, which is not observed in the L1-i group, suggesting dissimilation between aspirated and fortis stops.

In summary, marked differences in the implementation of onset F0 for stop production were observed in both the HS and LTI groups, as compared to the L1-i group. Importantly, this group variability was particularly manifest in the lenis-aspirated stop contrast. The HS group produced aspirated stops with reduced onset F0 while articulating lenis stops with increased onset F0, leading to a diminished gap in onset F0 between lenis and aspirated stops, compared to that of the L1-i group. In contrast, the LTI group demonstrated an increase in onset F0 for aspirated stops, extending the onset-F0 distance between aspirated and lenis stops beyond the range observed in the L1-i group, showing a greater onset F0-based contrast between the two types of stops than the L1-i group. Furthermore, the LTI group participants showed a distinct pattern in the aspirated and fortis stop contrast via onset f0, which was not observed in the L1-i group as well as the HS group.

Figure 4.5. Caterpillar plot representing the individual semitone estimates in production for Korean stop consonants across groups, based on the Bayesian model prediction.



4.2.4 Effects of bilingual profiles on acoustic correlates

The results of the VOT model, examining the impact of bilingual dominance as indexed by MINT_Diff, confirm the three-way distinction in Korean stops among LTI and HS group participants. Specifically, the model estimated the mean VOT for aspirated stops among bilingual participants at 90.87 ms (β_0 = 90.87, 95% CI = [83.70, 97.69], pd = 100.0). For fortis stops, the predicted mean VOT showed a reduction of 66.39 ms compared to aspirated stops, categorizing them within the short-lag VOT range (β = −66.39, 95% CI = [−75.32, −56.72], pd = 100.0). Lenis stops were produced with a shorter VOT compared to aspirated stops, by 11.15 ms (β = −11.15, 95% CI = [−20.20, −2.94], pd = 99.67). Despite extended VOTs across all stop types relative to

L1-i speakers, the bilingual groups' VOT patterns in Korean stop production closely paralleled those of L1-i speakers, as illustrated in Figure 4.2.

The statistical analysis revealed that the extended VOT values observed among bilingual participants are not attributable to their bilingual dominance, as measured by MINT_Diff. Specifically, MINT_Diff showed no credible influence on VOT for any stop type (Aspirated: pd = 53.60; Fortis: pd = 80.65; Lenis: pd = 51.10). These findings suggest that the longer VOT observed across Korean stops in bilingual participants is linked more to inherent characteristics of the speakers in the bilingual groups rather than to aspects of their bilingual profiles. This pattern is visually depicted in the left panel of Figure 4.6, where the nearly horizontal lines on the y-axis extending towards the x-axis represent this trend. Summaries of the statistical results of both the VOT and onset F0 models are provided in Table 4.3.

Figure 4.6. Conditional effects of bilingual dominance, measured as MINT_Diff, on the estimates of VOT (left panel) and onset F0 (right panel) in Korean stops.

Each data point represents the estimated mean VOT and semitone values for different stop types, calculated for each participant as a function of their MINT_Diff score. The shaded areas indicate the 95% CI.

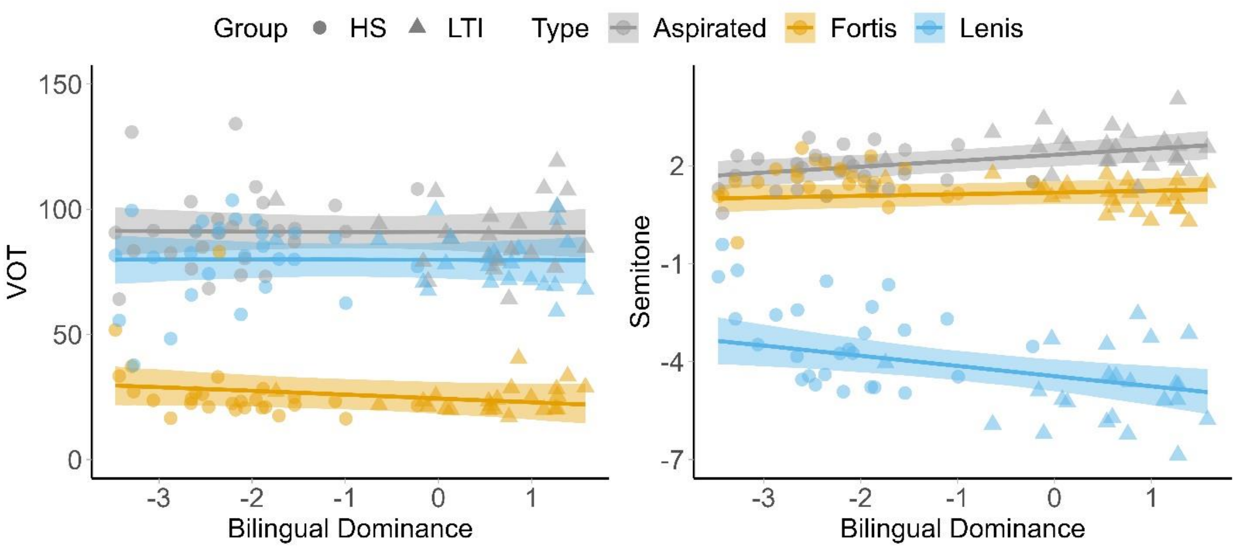


Table 4.3. Summaries of the Bayesian mixed-effects linear regression models examining the effects of bilingual dominance on VOT and semitone.

	VOT Model					
	Estimate	Est.Error	L-95% CI	U-95% CI	Rhat	pd
Intercept	90.87	3.58	83.70	97.69	1.01	100%
TypeFortis	−66.39	4.71	−75.32	−56.72	1.01	100%
TypeLenis	−11.15	4.39	−20.20	−2.94	1.01	99.67%
Z_MINT_Diff	−0.10	1.42	−2.83	2.74	1.01	53.60%
TypeFortis:Z_MINT_Diff	−1.39	1.62	−4.56	1.79	1.00	80.65%
TypeLenis:Z_MINT_Diff	0.04	1.26	−2.41	2.57	1.00	51.10%

	Onset F0 Model					
	Estimate	Est.Error	L-95% CI	U-95% CI	Rhat	pd
Intercept	2.34	0.18	1.98	2.70	1.01	100%
TypeFortis	−1.16	0.25	−1.64	−0.69	1.00	100%
TypeLenis	−6.77	0.35	−7.46	-6.07	1.01	100%
Z_MINT_Diff	0.18	0.06	0.07	0.29	1.00	99.90%
TypeFortis:Z_MINT_Diff	−0.13	0.07	−0.26	0.01	1.00	96.67%
TypeLenis:Z_MINT_Diff	−0.49	0.15	−0.79	−0.19	1.00	99.92%

The Bayesian mixed-effects model assessing the influence of bilingual profiles on the onset F0 (semitone) of Korean stops also substantiates the three-way stop distinction among the bilingual speakers. The estimated mean of onset F0 was highest for aspirated stops at 2.34 (β_0 = 2.34, 95% CI = [1.98, 2.70], pd = 100.0). In comparison, fortis stops exhibited a modest reduction in semitone value by 1.16 relative to aspirated stops (β = −1.16, 95% CI = [−1.64, −0.69], pd = 100.0). More notably, lenis stops showed a substantial decrease in semitone value by 6.77 compared to aspirated stops (β = −6.77, 95% CI = [−7.46, −6.07], pd = 100.0), indicating the tonal distinction between lenis and aspirated stops. This pattern suggests that bilingual speakers' use of onset F0 in the production of Korean stops on the whole also parallels that observed in L1-i speakers.

Contrasting with VOT, however, the onset F0 in Korean stops among bilingual speakers showed a significant correlation with their bilingual profiles. Specifically, an increase in MINT_Diff by one unit was associated with a rise in the estimated mean semitone for aspirated stops by 0.18 (β = 0.18, 95% CI = [0.07, 0.29], pd = 99.90). In contrast, for fortis stops, the estimated mean decreased by 0.13 per unit increase in MINT_Diff relative to the increase for aspirated stops ($\beta = -0.13$, 95% CI = [−0.26, 0.01], pd = 96.67). Lenis stops exhibited a more pronounced decrease in semitone, reducing by 0.49 per unit increase in MINT_Diff compared to aspirated stops ($\beta = -0.49$, 95% CI = [−0.79, −0.19], pd = 99.92). These results indicate a widening gap in onset F0 between aspirated stops and fortis/lenis stops as bilinguals' dominance in Korean increase. This trend is visually confirmed in the right panel of Figure 4.6, which illustrates the increasing semitone values of fortis and aspirated stops in relation to increasing MINT_Diff scores, with lenis stops showing a reduction in semitone values as MINT_Diff increases.

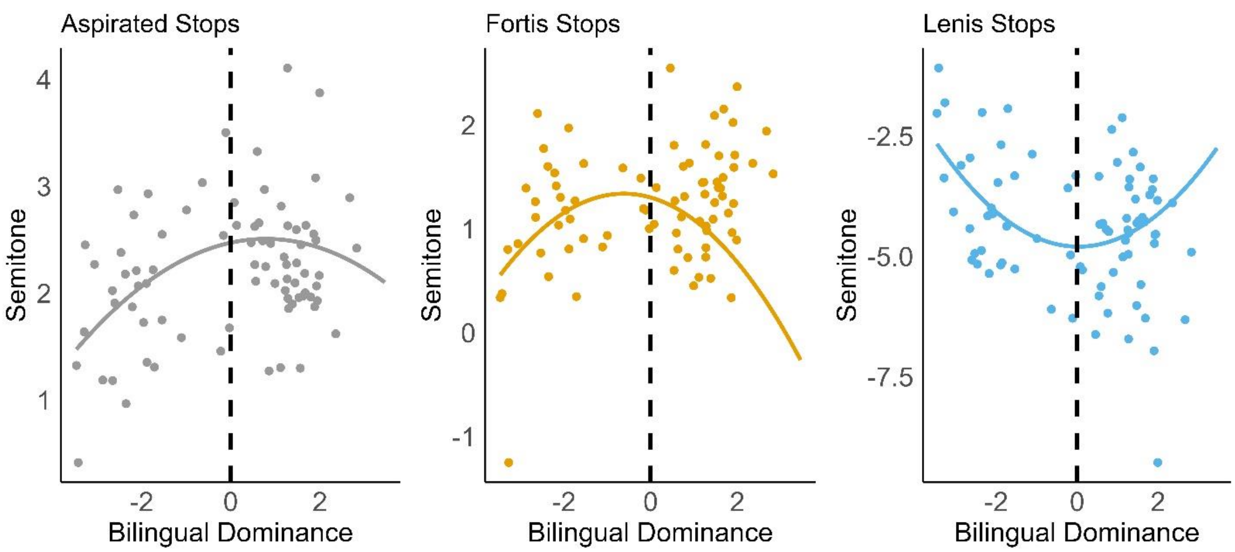
In order to validate the third hypothesis and further investigate the nature of the relationship between bilingual balance and the contrast in onset f0, it was crucial to determine if the observed increase in contrast towards bilingual balance would diminish again as the balance continues shifting towards strong Korean dominance. To address this, the study performed three separate quadratic regression analyses for each stop type within the bilingual speakers' dataset (Table 4.4). Each model specifically examined the relationship between semitone and bilingual balance as measured by MINT_Diff. Figure 4.7 represents the polynomial quadratic analysis for each stop type, with each data point corresponding to the mean value of each stop type per participant. As shown in the figure, the non-monotonic quadratic curves, estimated by each regression model, highlight a peak or bottom (vertex) in onset f0 contrast around a MINT_Diff score of zero. Specifically, for aspirated stops, the onset f0 value escalated until reaching a MINT_Diff score of 0.79, beyond which it began to decrease as the balance shifted towards Korean dominance. In the case of fortis stops, the apex of the onset f0 value was observed at a MINT_Diff score of −0.62. For lenis stops, the model predicted the nadir of the onset f0 value at a MINT_Diff score of 0.03. These patterns demonstrate that bilingual participants with a balanced dominance in both languages exhibit a more pronounced contrast in onset f0 than their less balanced peers, suggesting category dissimilation, irrespective of the direction of dominance. The rationale for including the MINT_Diff instead of the BLP score was for the desire to obtain a more objective measure of dominance in the investigation. Specifically, BLP scores can vary significantly depending on cultural backgrounds of the participants, which is the case in the present study with HSs being immersed in the American culture, L1-i speakers in the Korean culture, and the LTIs presumably situated between the two cultures(for more discussion, see Olson, 2023).

Table 4.4. Summaries of the quadratic regression models examining the relationship between bilingual balance and semitone.

Aspirated Stops				
	Estimate	Std.Error	*t*	*p*
Intercept	2.47	0.04	59.35	< .001
MINT_Diff	0.09	0.03	2.80	.005
MINT_Diff_Squared	−0.06	0.02	−3.77	.001
Fortis Stops				
	Estimate	Std.Error	*t*	*p*
Intercept	1.30	0.05	27.50	< .001
MINT_Diff	−0.12	0.04	−3.24	.001
MINT_Diff_Squared	−0.10	0.02	−5.59	< .001
Lenis Stops				
	Estimate	Std.Error	*t*	*p*
Intercept	−4.80	0.06	−83.30	< .001
MINT_Diff	−0.01	0.04	−0.26	0.796
MINT_Diff_Squared	0.18	0.02	8.35	< .001

Figure 4.7. Quadratic curves (parabolas) based on quadratic regression analysis for each stop type and semitone among bilingual participants.

Each data point corresponds to the mean value for each stop type per participant.



In summary, the Bayesian mixed-effects linear regression models investigating the impact of bilingual dominance confirm the canonical three-way distinction in Korean stops among bilingual speakers. These models illustrate that both VOT and onset F0 patterns among bilingual speakers in general align with those observed in L1-i speakers. Furthermore, the findings underscore a distinct influence of bilingual profiles on these acoustic features. The VOT model suggests that the extended VOT durations across stop types among bilingual speakers, relative to L1-i speakers, is an intrinsic characteristic of these speakers not related to their bilingual dominance. Interestingly, in contrast, the onset F0 model reveals a significant effect of bilingual dominance on the use of onset F0. The findings indicate that as MINT_Diff increases, there is a notable increase in the disparity of onset F0 between aspirated and fortis/lenis stops among bilingual speakers. Furthermore, subsequent quadratic regression analyses revealed that category contrast in onset f0 is most pronounced among balanced bilinguals. This points to a phonetic dissimilation among the stop categories, influenced by the bilinguals' language dominance.

4.3 Interim discussion

The current study investigated L2 (English) CLI on the acoustic parameters of VOT and onset F0 in the production of L1 (Korean) stop categories among bilingual speakers, focusing on HSs and LTIs in comparison to L1-i speakers. Crucially, the study aimed to elucidate the mechanisms of CLI in bilingual speakers as proposed by SLM-r (Flege & Bohn, 2021). This involved examining how language dominance, quantified by the MINT Sprint (Garcia & Gollan, 2021), impact L1 phonetic categories. The sections below discuss the findings from the statistical analyses in relation to the study hypotheses, along with alternative interpretations and implications of these results.

4.3.1 L2 CLI on acoustic correlates of Korean stops

Concerning the L1-i speakers, the results from the statistical analyses aligned with previous research, confirming the three-way distinction in Korean stops via both VOT and onset F0 (Bang et al., 2018; Cho et al., 2002; Han & Weitzman, 1970; Hardcastle, 1973; Kagaya, 1974; Kang, 2014; Kang & Guion, 2006, 2008; Kang & Nagy, 2016; Kim, 1965; Kong et al., 2011; Lisker & Abramson, 1964; Silva, 2006). In line with prior studies, the findings highlight the primary role of onset F0 in differentiating lenis and aspirated stops, with a minimal VOT difference as short as 8 ms between the two stop types. Interestingly, while the difference in VOT between lenis and aspirated stops was statistically significant, the magnitude of this temporal difference—below the just noticeable difference (JND) threshold of 23 ms for VOT, according to Hazan et al. (2009)—may mean that it does not contribute significantly to the perceived contrast. This finding supplements and expands upon previous literature documenting the sound change in the lenis-aspirated stop contrast in Korean, where VOT is becoming less salient as an acoustic cue, with onset F0 rising as a new primary cue (Bang et al., 2018; Choi et al., 2020; Kang, 2014; Kang & Guion, 2006, 2008; Kim & Jongman, 2022; Kong et al., 2011; Lee et al., 2020; Lee & Jongman, 2012, 2019; Schertz et al., 2015; Silva, 2006).

The analyses of the HS and LTI groups suggested differences in comparison to the L1-i group. Both groups maintained the canonical three-way stop distinction characteristic of Korean, mirroring the pattern observed in the L1-i group. However, HS and LTI groups exhibited extended VOT durations across stop types compared to L1-i speakers. Despite this increase, their implementation of VOT for the three-way stop contrast did not significantly deviate from that of the L1-i group, except for aspirated stops. This finding contrasts with the first hypothesis that the bilingual groups would demonstrate a wider VOT gap among Korean stop categories, as VOT is a primary cue in the voicing contrast of English stops, and these groups are proficient and experienced in English.

This pattern in the bilingual groups suggests group-specific phonetic characteristics inherent to speakers in the Korean–English group rather than a result of bilingual dominance. Prior research has documented an enhancement in VOT across stop categories among Korean–English bilinguals (Chang & Mandock, 2019; Cheng, 2017; Kang & Guion, 2006; Kang & Nagy, 2016; Lee & Iverson, 2012; Oh, 2019; Oh & Daland, 2011; Yoon, 2015). However, these studies present mixed results regarding the VOT gap among stop categories, particularly between lenis and aspirated stops. Specifically, some studies reported extended VOT for both lenis and aspirated stops without significant differences in the role or weight of VOT for this contrast compared to L1-i speakers (Chang & Mandock, 2019; Cheng, 2017; Kang & Guion, 2006; Lee & Iverson, 2012; Yoon, 2015). Chang and Mandock (2019), for instance, interpreted this VOT enhancement as a manifestation of a clear speech effect. While VOT across different stops in Korean may be prolonged due to a clear speech effect, considering the consistency in findings across multiple studies, it appears plausible to posit that lengthened VOT in L1 Korean stops among bilinguals may be CLI from L2 English. On the other hand, other studies reported an increased VOT gap between lenis and aspirated stops in HS groups, although they noted a reduction in the VOT difference with younger age and among female speakers, which characterizes the pioneering group that led the lenis-aspirated stop sound change in 1960's (Kang & Nagy, 2016; Oh, 2019; Oh & Daland, 2011). This variation in the findings of previous studies implies that while longer VOT is a common phonetic feature among Korean–English bilinguals, the degree of differentiation in VOT among Korean stop categories may be subject to factors such as age and gender. Moreover, the observed pattern of enhanced VOT without a corresponding increase in the distinction between stop types challenges the establishment of a straightforward relationship between language backgrounds, such as the timing of L2 acquisition (HSs: early childhood vs. LTIs: adulthood) affecting language dominance, and phonetic manifestation of contrast in L1.

The analysis of onset F0 among the HS and LTI groups yielded results that partially aligned with the second hypothesis, which posited a less pronounced distinction in onset F0 among these groups compared to L1-i speakers. This hypothesis was grounded in the assumption that onset F0 plays a secondary role in the voicing contrast of English stops, thereby potentially diminishing its prominence in the Korean stops produced by Korean–English bilinguals. Contrary to this expectation, the findings revealed an intriguing pattern, supporting the hypothesis only partially: while the HS group showed a reduced distinction in onset F0, particularly, between lenis and aspirated stops, compared to L1-i speakers, the LTI group exhibited an opposing pattern. Specifically, the HS group demonstrated an elevation in the onset F0 of lenis stops, yet a reduction for aspirated and fortis stops, leading to a shorter distance between lenis and aspirated stops. This decrease in the onset F0 distance among Korean stops is a pattern comparable to L2 learners of Korean whose L1 is English (Chang & Mandock, 2019) and second-generation Korean HSs as in the participants of the present study (Cheng, 2017; Kang & Nagy, 2016).

In a striking contrast, the LTI group showed an augmented distinction in onset F0, especially between lenis and aspirated stops, with an increased onset F0 in aspirated stops. This divergence suggests that the LTI group more closely adheres to the contemporary Korean norm, where onset F0 is primary cue to the lenis-aspirated stop contrast, than the HS group. This distinguished pattern of the LTI group from the HS group evidence that L1 inputs and language dominance play a crucial role in determining the degree of L2 CLI on L1 phonetics. The LTI group had ample opportunity to acquire and master their L1 (Korean) prior to immersion in an L2 (English)-dominant environment, whereas the HS group began immersion in their L2 during early childhood. Moreover, while the LTI group was less proficient in English than the HS group, they were more balanced, as indicated by their MINT_Diff score close to zero, than the HS group with the score in the range of English dominance. Consequently, LTIs may maintain a greater capacity for differentiating Korean stop categories using the mechanisms akin and even superior to L1-i speakers. However, the more pronounced distinction in onset F0 observed in the LTI group, particularly in the lenis-aspirated contrast, compared to the L1-i group, contradicts the hypothesis that onset F0 would play a diminished role for LTI bilinguals.

The observed patterns in both the LTI and HS groups suggest that age of acquisition of L2 may not fully explain the variations in the use of onset F0 among Korean–English bilinguals. For instance, Cheng (2017) found that adult immigrant Korean–English bilinguals employed onset F0 more robustly, as evidenced by the difference in onset F0 among Korean stops, in speech production compared to second-generation HSs, akin to those in the current study. Conversely, other research found no significant differences in the use of onset F0 in the production of Korean stops between L1-i speakers and early Korean–English bilinguals (e.g., Kang & Guion, 2006; Oh & Daland, 2011). Due to such mixed results regarding the effects of age, found across different topics in bilingual phonetics and phonology, SLM-r postulates that quality and quantity of L2 input are crucial factors determining CLI during bilingual category formation (Flege & Bohn, 2021). According to SLM-r, bilinguals may reach a stage where they can separate L1 and similar L2 categories, potentially leading to acoustic category dissimilation in production. Supporting this argument, several studies have demonstrated that more balanced bilingual speakers are more adept at differentiating between L1 and L2 phonetic systems (Barlow et al., 2013; Guion, 2003; Kang & Guion, 2006; MacLeod et al., 2009; Sundara et al., 2006). Therefore, the pronounced distinction in onset F0 among LTIs, in contrast to L1-i speakers, may reflect an advanced stage of bilingual category formation. In this stage, L1 categories diverge from L2 categories acoustically in ways that enhance within-language phonological contrasts, indicative of a dissimilation process between the L1 (Korean) and L2 (English).

4.3.2 Bilingual balance as a factor of CLI in L1 categories

To validate the SLM-r predictions on bilingual category formation and assess whether reliance on onset F0 in the production of L1 (Korean) stop categories is influenced by the stage of category formation posited by the model, it is crucial to investigate the impact of language dominance on L2 CLI on L1 phonetics. The present study approximated the state of bilingual development by measuring the difference in MINT Sprint scores between Korean and English (MINT_Diff) and explored its effects on the acoustic parameters of VOT and onset F0 among HSs and LTIs. If balanced bilinguals are more adept at separating L1 and L2 phonetic systems, more balanced bilinguals would exhibit a narrower VOT distinction among Korean stop categories, especially between lenis and aspirated stops, as VOT is a secondary or not even an informative cue in this contrast. In contrast, these bilinguals would demonstrate an increased reliance on onset F0, the primary cue to the contrast, resulting in more pronounced distinctions among Korean stop categories than those with less balanced dominance.

The analyses of the statistical models revealed no credible effect of bilingual profiles on VOT among bilingual speakers. That is, bilingual participants including both HSs and LTIs produced Korean stops with longer VOT than L1-i speakers, irrespective of their bilingual dominance. This finding lends support to the claim that longer VOT in Korean stops produced by Korean–English bilinguals may be a group-inherent characteristic, rather than an outcome from individual differences in bilingual balance. This finding can account for why multiple studies have consistently observed extended VOT in Korean stops among several groups Korean–English bilinguals varying in linguistic backgrounds (Chang & Mandock, 2019; Cheng, 2017; Kang & Guion, 2006; Kang & Nagy, 2016; Lee & Iverson, 2012; Oh, 2019; Oh & Daland, 2011; Yoon, 2015).

Regarding onset F0, a credible effect of bilingual profiles on the degree of L2 CLI was observed. More specifically, balanced bilingual speakers exhibited a more pronounced distinction in onset F0 between lenis and aspirated stops by increasing the onset F0 of aspirated stops along with reduced onset F0 in lenis stops. Crucially, subsequent quadratic regression analyses substantiated that the contrast in Korean stop categories, as defined by onset f0, progressively intensifies from a state of English dominance towards a balanced bilingual state, and subsequently diminishes as the dominance shifts towards Korean. These findings elucidate the mechanisms of L2 CLI on L1 phonetic categories. It suggests that a balanced level of bilingualism may lead to an enhanced sensitivity to or use of specific phonetic cues in both languages, in accordance with language-specific phonetic priorities (e.g., VOT for English stops). Moreover, the observed increase in the onset F0 distance between lenis and aspirated stops as the function of bilingual balance provides concrete evidence that balanced bilingualism as the advanced stage of bilingual category formation can lead to phonetic dissimilation between L1 and L2, aligning conceptually with the SLM-r predictions (Flege & Bohn, 2021).

While the present study offers insights into the variability in the acoustic correlates of Korean stops in Korean–English bilinguals, alternative interpretations merit consideration. One possibility is that the observed lengthening of VOT in the HS and LTI groups might not be an intrinsic characteristic of Korean–English bilinguals. For instance, this VOT lengthening across various Korean stops could be attributed to a clear speech effect, as suggested by Chang and Mandock (2019). However, it cannot explain why L1-i participants alone did not show a clear speech effect. Moreover, the VOT durations for Korean stops produced by the HS and LTI groups were significantly longer than those by the L1-i group. This pattern aligns with previous studies that consistently reported longer VOT in Korean stops among Korean–English bilinguals, suggesting that clear speech alone might not fully explain this observation in Korean–English bilinguals (Chang & Mandock, 2019; Cheng, 2017; Kang & Guion, 2006; Kang & Nagy, 2016; Lee & Iverson, 2012; Oh, 2019; Oh & Daland, 2011; Yoon, 2015).

In addition, the present study interprets the observed increase in the onset F0 distance among Korean stops, which is category dissimilation within L1 categories, also as indicative of dissimilation between L1 (Korean) and L2 (English) systems of contrasts, for bilinguals to differentiate the two phonetic systems effectively. However, it is important to note that this study did not include a direct examination of English stops, which could provide more direct evidence to support these findings. Despite this limitation, the interpretation of an increased reliance on onset F0 for Korean stops aligns with prior research examining both VOT and onset F0 in Korean and English stops (Lee & Iverson, 2012; Oh, 2019). These studies, based on the previous version of SLM (Flege, 1987, 1995, 2003), report category dissimilation between L1 and L2 stop categories, underscoring bilinguals' ability and need to differentiate between the phonetic systems of their two languages. Therefore, the study contributes additional empirical support to the relatively limited body of evidence for category dissimilation (Flege & Eefting, 1988; Flege et al., 2003; Mack, 1990; Yusa et al., 2010), further corroborating that category dissimilation is a potential outcome observable at advanced stages of bilingual category formation, as postulated by SLM-r (Flege & Bohn, 2021).

5. DISCUSSION

Portions of the current chapter have been published in *Journal of Phonetics:*

Seo, Y., & Dmitrieva, O. (2024). L2 cross-linguistic influence on L1 perception: Evidence from heritage speakers and long-term immigrants. *Journal of Phonetics*, *104*, 101314. DOI: https://doi.org/10.1016/j.wocn.2024.101314.

5.1 Summary of the results

5.1.1 Experiment 1: 3AFC identification task

The primary goal of the 3AFC identification task was to investigate L2 (English) CLI on perceptual cue weighting of VOT and onset f0, the acoustic parameters for L1 (Korean) stop categories. In each trial, participants listened to a word that varied in the word-initial stop, comprising eight-by-eight VOT and onset f0 continua. These could potentially be perceived as one of the three stop categories: lenis (long-lag VOT and low onset f0), aspirated (long-lag VOT and high onset f0), and fortis stops (short-lag VOT and high onset f0). The study conducted both between-group and within-group analyses to examine group and individual variabilities in their perceptual cue weighting to Korean stops. By adopting the cue-weighting approach that can detect more fine phonetic differences, compared to traditional measures (e.g., a simple discrimination task), the analysis of the 3AFC task data revealed CLI in L1 perception among the bilingual groups.

The results from three mixed-effects logistic regression models revealed interesting group patterns across all three stop contrasts. First, HSs showed decreased sensitivity to the onset f0 cue in the perception of lenis and aspirated stops, compared to both LTIs and L1-i speakers, while LTIs demonstrated the highest level of sensitivity to this cue among the three groups. Furthermore, the study found that the perceptual distance between lenis and aspirated stops via onset f0 was greatest in the LTI group, suggesting category dissimilation. In contrast, the HS group exhibited decreased distance between these two stop types compared to the L1-i group, indicating category assimilation.

Second, the analysis of lenis and fortis stops, contrasted via both onset f0 and VOT, revealed that onset f0 outweighs VOT in both the L1-i and LTI groups. Participants in these groups relied more on onset f0 than on VOT when perceiving the lenis-fortis stop contrast. However, while the HS group could use both cues, their perception of this contrast was significantly less categorical, evidenced by their distinct reliance on VOT by perceiving stimuli with long-lag VOT as fortis stops. Furthermore, their reliance on onset f0 in the perception of lenis and fortis stops significantly decreased compared to the other two groups.

Third, the analysis of aspirated and fortis stops, primarily contrasted via VOT alone, with aspirated stops having longer VOT than fortis stops, revealed a distinct use of onset f0 in the HS group. Specifically, HSs tended to perceive stimuli with higher onset f0 as aspirated stops, differentiating them from fortis stops when VOT could not cue the contrast between the two stop types. On the other hand, the LTI group patterned with the L1-i group in their implementation of the two acoustic cues.

These findings collectively provide evidence of L2 CLI in the L1 phonetic system in the perception domain. Crucially, the current study elucidates the mechanisms of L2 CLI in L1 cue weighting, attributable to the timing of L2 acquisition and the quantity and quality of L1 and L2 input. Specifically, CLI in HSs manifested as decreased sensitivity to onset f0 and assimilation via the use of onset f0, compared to L1-i group speakers. On the contrary, CLI in LTIs was evidenced by increased sensitivity to onset f0 and dissimilation in the use onset f0. Given that HSs experience a shift in their dominant language, with their L2 becoming dominant over their L1 due to early immersion in the L2, their decreased reliance on onset f0 is attributed to their reduced L1 experience. This linguistic experience with the decreased quality and quantity of L1 input leads to decreased sensitivity to onset f0, as VOT is the primary cue for the voicing contrast in English. In contrast, LTIs, who began their L2 (English) immersion in the US as adults, were bilinguals with a relative balance in language dominance and proficiency between Korean and English. CLI in LTIs led to increased sensitivity to language-specific cue primacy, presumably in order to enable efficient perception of both Korean and English speech.

Finally, HSs' unique reliance on onset f0 for fortis stops can be attributed to the difference in the covariation of acoustic parameters between Korean and English stops. In English laryngeal stops, VOT and onset f0 covary, with increased VOT corresponding to high onset f0. In contrast, fortis stops in Korean exhibit the opposite pattern, characterized by short-lag VOT and high onset f0. Therefore, HSs' distinctive reliance on onset f0 for identifying fortis stops may be due to this cross-linguistic difference in acoustic parameters; their English-based perceptual strategy led them to perceive fortis stops with short-lag VOT as having lower onset f0 compared to aspirated stops with long-lag VOT.

5.1.2 Experiment 2: AX discrimination task

The primary goal of the AX discrimination task was to investigate L2 CLI in L1 perceptual discrimination by examining the effects of group and trial type on participants' perceptual accuracy in distinguishing L1 stop categories, when forced to rely on the VOT cue. In each trial, participants listened to a pair of Korean words that could differ in the word-initial stop being either a lenis or aspirated stop. Crucially, the two types of stops were only contrasted via onset f0 by making the VOT the same duration at 70 ms across all stimuli.

The results from the mixed-effects logistic regression models revealed significant deviations in participants' discriminatory performance across groups. Interestingly, the order of perceptual accuracy among the participant groups followed the order of onset f0 sensitivity in the 3AFC task. Specifically, the HS group's discrimination performance lagged behind those of the LTI and L1-i groups, while the LTI group outperformed the L1-i group.

The study also found that participants' relative weighting of the onset f0 cue impacts their ability to discriminate lenis from aspirated stops, primarily contrasted via the onset f0 cue. That is, participants who assigned greater weight to onset f0 in the 3AFC identification task demonstrated increased accuracy in discriminating between these stop types in the AX discrimination task.

The analysis of RTs for accurate responses in different-word trials showed that HSs required more time for accurate discrimination between lenis and aspirated stops, suggesting increased cognitive load during the discrimination process. In contrast, LTIs were not significantly different in RTs from those of L1-i speakers, suggesting that their increased accuracy was not compensated by prolonged processing time.

These findings suggest that HSs are not immune to L2 CLI in the speech domain, as evidenced by their decreased perceptual accuracy compared to both LTI and L1-i group participants. Furthermore, the finding of the increased accuracy in the LTI group compared to the other groups suggests that CLI can have facilitatory effects in the perception domain. In line with the 3AFC task, this may be attributed to the need in LTIs to engage in both Korean and English communications in daily lives.

5.1.3 Experiment 3: Controlled reading paradigm task

The primary goal of the controlled reading paradigm task was to investigate L2 CLI on the two acoustic parameters, VOT and onset f0 in the speech production of L1 (Korean) stop categories. In this task, participants read aloud a list of stimuli in a carrier phrase consisting of Korean minimal triplets differing in the word-initial stop (e.g., /pul/ 'fire', /p^{h}ul/ 'grass', /p*ul/ 'horn'). An additional goal of this task was to investigate the potential impact of bilingual balance in language dominance between Korean and English on the implementation of the two acoustic parameters.

The results from Bayesian mixed-effects linear regression models showed findings in parallel with the 3AFC identification task. Regarding onset f0, HSs articulated lenis stops with increased onset f0 compared to L1-i speakers, while LTIs produced them with reduced onset f0. On the other hand, the onset f0 decreased in the production of both aspirated and fortis stops by HSs compared to L1-i speakers, suggesting that the distance among stop categories in HSs' was shorter than that of L1-i speakers. On the other hand, the LTIs produced aspirated and fortis stops with increased and decreased onset f0, respectively, compared to the L1-i speakers, which suggests category dissimilation. The analysis of VOT showed somewhat different patterns in both HS and LTI groups compared to the L1-i group. Specifically, both bilingual groups had a tendency to produce all types of Korean stops with prolonged VOT in relevance to the L1-i group. These findings suggest parameter-specific L2 CLI from English in L1 Korean speech, congruent with the findings from the 3AFC identification task.

In order to investigate the potential effects of bilingual dominance among bilingual speakers, as estimated by a measure of MINT Sprint, separate quadratic polynomial regression analyses were conducted for each stop category. The study found that bilingual balance has a significant effect on the parameter of onset f0, such that bilinguals with relative balance between L1 and L2 showed greater contrasts via onset f0 by either increasing or decreasing it, creating further distance between categories. Surprisingly, this pattern diminished as the balance measure shifted towards a dominance in either language, exhibiting a non-monotonic relation between bilingual balance and the use of onset f0. In alignment with the perception experiments, this finding may emphasize the need to separate two phonetic systems of the two different languages in balanced bilinguals.

5.2 Implications for bilingual phonetics and phonology

5.2.1 Perception experiments

The results of the investigation of L2 CLI in L1 perception are compatible with the general predictions of the bidirectional CLI made by SLM-r, including the effect of L2 on the L1 in speech perception of the diverse bilingual populations, such as HSs and LTIs. Moreover, the possibility of both the assimilatory and dissimilatory CLI, predicted by SLM-r, was supported by the findings. On the other hand, the study was not able to predict with precision the direction of CLI (assimilatory or dissimilatory). It is believed that three aspects of the theoretical framework the study relied upon contributed to this inability to make precise predictions. First, the fundamental conditions that determine the likelihood of L2 category creation, according to SLM-r, and, therefore, the likelihood and nature of CLI, such as the amount and quality of L2 and L1 input are difficult to measure with precision. Consequently, it is difficult to determine how much input is necessary for certain types of CLI to occur.

Second, these factors do not result in CLI directly, but via the mechanism of L2 category creation. Yet, there is no established, reliable, and agreed upon method for determining when a distinct L2 sound category has been created—a problem which undermines the viability of the explanatory mechanism for CLI in SLM-r. Approaching the issue from another direction, the lack of accepted tests for L2 category creation, which precludes the establishment of a clear link between this stage of L2 acquisition and CLI, compromises CLI as a diagnostic for category formation.

And third, factors other than the ones listed above are likely to contribute to the likelihood of assimilatory or dissimilatory CLI, in particular, sociolinguistic factors, such as language attitudes (Law et al., 2021; Law & Francis, 2015), societal context, and language status (Ahn & Chang, 2022). For instance, Ahn and Chang (2022) demonstrated this by showing that Korean heritage children in the US who only use their heritage language at home develop emotional words differently compared to those who use Korean in broader social contexts, with the latter group patterning similar to monolingual children. Similarly, the societal status of language (majority vs. minority language) can influence many aspects of bilingual language development. Future research may target these factors in investigating the mechanism behind CLI by incorporating bilingual populations with more diverse sociolinguistic profiles, such as Korean returnees or

Pilipino-Korean bilinguals in South Korea. Concurrently, future refinements or re-developments of theoretical approaches to L2 and bilingual speech need to address societal factors in order to generate more precise and testable predictions with respect to CLI in various bilingual populations.

5.2.2 Production experiment

The findings of the production experiment provide valuable benchmarks for understanding how bilingual experiences translate into certain types of CLI. Crucially, the inclusion of two distinct bilingual populations, LTIs and HSs, underscores the vital role of L1 experience in bilingualism and CLI. While previous research often focused on the amount of L2 input as a decisive factor in L2 category creation and CLI, the present findings suggest that L1 input and experience can be of equal importance. Although HSs had more L2 input and experience, LTIs were the ones whose performance suggested more independent L1 and L2 perceptual categories, which the study attributes to the more developed and established L1 sound system, including more robust L1 sound categories. This conclusion echoes the findings of Ahn et al. (2017) who showed that bilinguals' accuracy in the perception of L1 contrast correlated with the age of reduced contact with L1, highlighting the importance of L1 input in the maintenance of L1 perception (see also Chang & Ahn, 2023).

The findings of the present study have significant implications, both theoretically and practically. First, the study corroborated and confirmed the observed sound change in Korean lenis and aspirated stops among young speakers in the Seoul-Gyeonggi area. The diminishing role of VOT as a primary cue, with its temporal difference between stop types falling below the JND threshold and being replaced by onset F0 as the new primary cue, aligns with previous research documenting the sound change in Korean laryngeal stops (Bang et al., 2018; Cho et al., 2002; Choi et al., 2020; Han & Weitzman, 1970; Hardcastle, 1973; Kagaya, 1974; Kang, 2014; Kang & Guion, 2006, 2008; Kang & Nagy, 2016; Kim, 1965; Kim & Jongman, 2022; Kong et al., 2011; Lisker & Abramson, 1964; Silva, 2006). More crucially, the examination of VOT and onset F0 in Korean stops produced by HSs and LTIs in the Midwestern United States underscores that this sound change is not confined to Korea but is also progressing among Korean–English bilinguals in the diaspora. The minimal VOT difference between lenis and aspirated stops for HSs and LTIs, while statistically significant, was still below the JND threshold (11.6 ms and 11.1 ms, respectively), a modest elevation from the L1-i group's difference (8.4 ms). This finding indicates that although the VOT difference between these stop types is statistically different, it might not be perceptually significant enough to serve as a reliable cue for their contrast, necessitating a reliance on onset F0 as the primary cue. This pattern is consistent with prior studies that have documented a similar pattern in the lenis-aspirated stop contrast among Korean–English bilingual communities in North America (Chang & Mandock, 2019; Cheng, 2017; Kang & Nagy, 2016). This ongoing phonetic shift among Korean–English bilinguals in diverse geographical regions suggests the fluidity of a sound change in bilingual contexts, where L1 inputs from family, socialization, and media and distinct linguistic environments coalesce to shape and follow the ongoing sound change in homeland. This finding contributes to our understanding of sound change in bilingual populations, suggesting that such changes are not confined to their place of origin but can also occur in diaspora communities with unique linguistic and sociocultural contexts.

The present study makes a significant theoretical contribution by providing empirical support for the SLM-r predictions regarding CLI in diverse bilingual populations (Flege & Bohn, 2021). The SLM-r model proposes both cross-linguistic assimilation and dissimilation are predicted at different stages of bilingualism. In particular, advanced bilinguals are predicted to be capable of dissimilating phonetically between similar L1 and L2 categories. The study findings, particularly the pronounced distinction in use of onset F0 among the LTI group in comparison to both HS and L1-i groups, broadly align with the model's theoretical predictions. Moreover, the observed influence of bilingual balance on the onset F0 distinction between lenis and aspirated stops extends the assumptions of SLM-r. While SLM, developed primarily for and based on adult L2 learners, highlights the amount of L2 experience as the driving force behind bilingual category formation and resulting CLI, the present study suggests that it is the balance between the two languages that drives CLI. These results shed light on the interaction between language dominance and the phonetic realization of L1 categories in bilingual speakers. Specifically, the greater reliance on onset F0 among the LTI group for distinguishing Korean stop categories underscores a pivotal aspect of the SLM-r model: the ability of advanced bilinguals, especially those with balanced dominance, to navigate and separate overlapping phonetic spaces of their L1 and L2. Furthermore, the study extends the understanding of CLI by demonstrating how the balance of language dominance can modulate phonetic distinctions among L1 categories. The enhanced distinction in onset F0 among the LTI group, as opposed to the HS group and L1-i speakers, suggests a developmental trajectory where advanced bilinguals move beyond assimilation of L2 categories into their L1, reaching a stage of phonetic dissimilation, where L1 and L2 categories are maintained as distinct, a process that is tied to their bilingual experience and balance. This contributes to a broader understanding of bilingual speech production, illustrating how bilinguals not only manage two distinct linguistic systems but also how their interaction and profiles in these languages can shape their L1 phonetic categories.

Finally, the study poses several practical implications, including heritage language acquisition and education and language maintenance, with a specific focus on the experiences of Korean Americans and Korean communities in the US. This demographic represents a unique case in the landscape of heritage languages in North America. Data from the National Heritage Language Survey in the US (Carreira & Kagan, 2011) indicates that a vast majority (99%) of Korean Americans become predominantly English speakers by age 17, a pattern that markedly differs from that of other heritage language groups, such as Spanish or Russian speakers, many of whom maintain their heritage language more robustly through family and community interactions. Informal conversations with some Korean American participants, though not part of the formal experiment, shed light on their linguistic experiences and attitudes. Many of these speakers, particularly first-generation immigrants, expressed reluctance to pass on a Korean accent in English to their children. This apprehension stems from a concern that such an accent might become an additional challenge for their children, who are already navigating life in the US as a minority group. These testimonies underscore the importance of recognizing and addressing heritage language maintenance and identity in Korean American communities.

5.3 Heritage language acquisition

The present study examining CLI in HSs contributes to the current discussion on heritage language acquisition in a relevant body of literature. Some of the representative approaches to HSs' language acquisition over the past two decades include L1 attrition and incomplete acquisition. The L1 attrition approach stipulates that HSs' L1 competence deteriorates over time with increased L2 input and use (Cuza, 2008; Polinsky, 2011; Schmid, 2011). Within this view, HSs' decreased reliance on the primary cue of onset f0 for the lenis-aspirated stop contrast, as well as their decreased perceptual accuracy is explained by diminished L1 competence. However, the validation of the L1-attrition approach remains incomplete without the longitudinal data of HSs in the present study.

The incomplete acquisition approach posits that HSs' decreased performance in linguistic assessments of the L1 is attributable to incomplete acquisition of the L1 due to decreased L1 experience, compared to L1 monolinguals (Montrul, 2002, 2008, among others). In the case of the HSs in the present study, this approach presupposes that HSs were not able to learn and master the contrast between lenis and aspirated stops, which is primarily cued via onset f0. Therefore, it is predicted that the incomplete acquisition of Korean stops among HSs resulted in decreased reliance on the onset f0 cue and decreased perceptual accuracy for the lenis-aspirated stop contrast. However, previous research has demonstrated that Korean children acquire Korean stops by the age of three or even earlier (Kim & Stoel-Gammon, 2009; Kong et al., 2011). Therefore, the unique patterns observed among HSs in the present study, who were expected to have acquired Korean stops before being immersed in the English language, do not lend compelling support to the incomplete acquisition approach.

More recently, the Bilingual Alignment Framework (Sánchez, 2019) stands out among other approaches with its explicit predictions of bidirectional CLI. This framework suggests that the linguistic domains in L1 and L2 of HSs are not fixed but flexible and mutually affect each other, and the degree of influence is affected by various linguistic factors, such as the quantity and quality of input in each language. Notably, this framework extends its scope beyond the morphosyntactic domain to the phonological level. However, relevant existing studies have predominantly focused on the morphosyntactic domain of HSs (Cuza & Sánchez, 2022; Hur et al., 2020; Perez-Cortes et al., 2019; Sánchez, 2019; Shin et al., 2023). Within the view of this framework, HSs' decreased reliance on onset f0 for the lenis-aspirated stop contrast in both perception and production domains can be viewed as a result of the alignment with English. However, it should be acknowledged that the present study did not involve the investigation of participants' English speech, which should await future research for a comprehensive test of the framework.

5.4 Limitations and future directions

The present study is not without limitations that merit consideration for future research. First, the present study did not include a direct investigation of L1 CLI from Korean on L2 English. This decision was partly due to the extensive literature documenting L1 Korean CLI on the perception and/or production of L2 English stops, based on which the present study posited the bidirectionality of CLI in both Korean and English (Kang & Guion, 2006; Kim & Tremblay, 2021; Kong & Kang, 2023; Kong & Yoon, 2013; Lee & Iverson, 2012; Oh & Daland, 2011; Oh, 2019; Schmidt, 1996, 2007). Specifically, previous studies have shown a consensus that Korean-dominant Korean–English bilingual speakers are more likely to rely on onset f0 in the perception and production of English stops, as onset f0 is the primary cue to the lenis-aspirated and lenis-fortis stop contrasts in Korean. For instance, Kong and Kang (2023) investigated CLI in the perception of both L1 (Korean) and L2 (English) stop categories by examining the implementation of perceptual cues of onset f0 and VOT by Korean–English bilinguals. The study found bidirectional CLI in the perception of both Korean and English, which was accounted for by proficiency in L2 English. Specifically, the study demonstrated a negative correlation between English proficiency and onset f0 weight, such that bilinguals with higher proficiency in English were less sensitive to onset f0 when perceiving English stops, primarily contrasted via VOT, compared to those with lower proficiency. The present study also demonstrated the same patterns of CLI in both perception and production in the direction of Korean. A future study may investigate both L1 and L2 speech perception and production to provide a more comprehensive picture of the bidirectionality of CLI in the phonetic and phonological domains among bilingual speakers (e.g., Jung, 2023; Kang & Guion, 2006).

Second, while the present study administered an AX discrimination task to examine L2 CLI on the perceptual accuracy in discriminating L1 stop categories, future research may revisit the assessment of perceptual accuracy by implementing an additional perception experiment that can measure the gradient nature of the perceptual process of discriminatory decision-making, such as a VAS task. As suggested by anonymous journal reviewers, such a task will be able to provide a more detailed picture of CLI in discriminatory decision-making through a gradient scale (Kapnoula & McMurray, 2021; Kapnoula et al., 2017; Kong & Edwards, 2016; Kutlu et al., 2022; Massaro & Cohen, 1983; McMurray et al., 2002). Nevertheless, the results from the AX discrimination task implemented in the current project revealed that several linguistic backgrounds, such as the timing of L2 acquisition and immersion and the quantity and quality of L2 and L1 input, can impact the perceptual accuracy of L1 category discrimination. Furthermore, the present study demonstrated that individual differences in onset f0 cue weight is significantly associated with their discriminatory performance. A future study incorporating both a 3AFC identification task and a VAS task may unravel a more detailed mechanism of CLI in the perceptual process of identifying and discriminating speech sounds.

Third, LTI participants in the present study were younger on average than HSs' parents and thus do not necessarily represent the same generation. Consequently, their Korean may not be comparable to HSs' parental input. In future studies, an inclusion of an older group of LTIs who can be categorized as belonging to the HSs' parent generation would allow for an investigation of the effect of generation in accounting for HSs' distinct patterns in the perception and production of L1 speech, as they could serve as a control group.

6. CONCLUSION

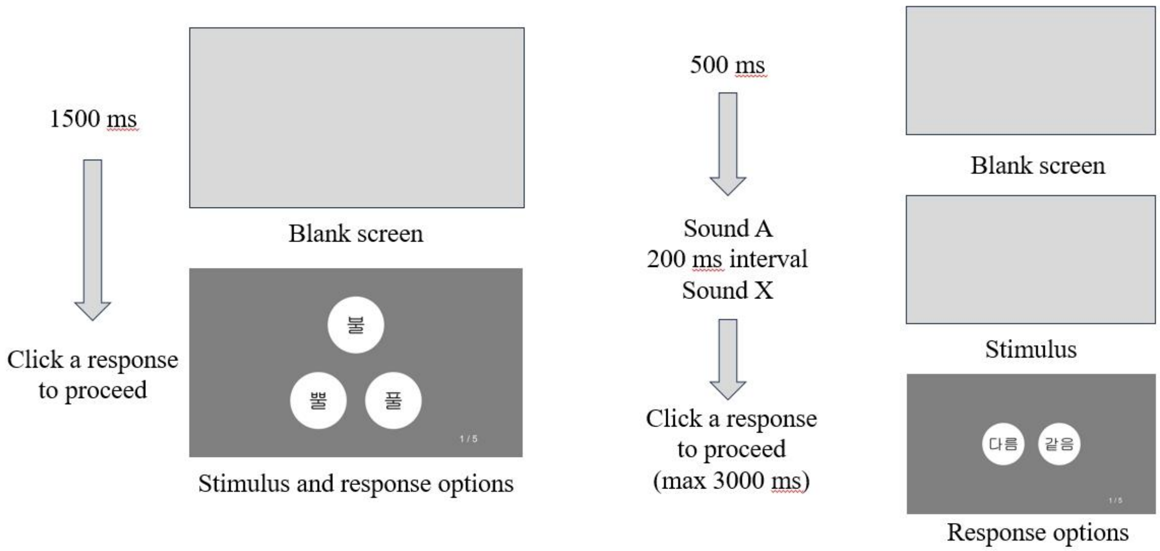
The present study, through a comparative analysis of perceptual cue weighting and discrimination among Korean HSs, LTIs, and L1-i speakers, provides original insights into the impact of L2 CLI on L1 perception and contributes to the broader understanding of bilingual phonetics. The findings underscore that L2 CLI has a significant impact on L1 in both perceptual categorization and discrimination. In the 3AFC task, HSs showed less reliance on the primary cue of onset f0 than L1-i speakers, suggesting an assimilatory effect of English on their Korean cue weighting. In contrast, LTI speakers exhibited a greater reliance on onset f0 than L1-i speakers, suggesting a dissimilatory effect of English on cue weighting in Korean. In the AX discrimination task, which necessitated the perceptual discrimination between lenis and aspirated stops based solely on onset f0 without the secondary cue of VOT, HSs showed a decreased perceptual accuracy compared to both LTI and L1-i speakers. Interestingly, the LTI group, which demonstrated an increased sensitivity to the f0 cue in the identification task, outperformed the L1-i group. Additionally, a link was found at the individual level between perceptual cue weighting and discrimination. Finally, the present study provides evidence that L1 input and experience play an important role in determining the direction and extent of CLI in L1 speech perception.

The current study also investigated the degree and mechanisms of L2 CLI on L1 phonetic categories by examining the variations in the acoustic parameters of Korean stops produced by Korean HSs and LTIs in the US, in comparison to L1-i speakers in South Korea. The research aimed to test the predictions of SLM-r regarding bilingual category formation, particularly focusing on the influence of bilingual balance as a gauge of bilingual development on the acoustic correlates of L1 speech categories. Results of the production experiment mirrored those from the perception experiment of the 3AFC task. The findings of the production task demonstrated the sound change in lenis and aspirated stops in Korean not only in the L1-i group but in the bilingual groups. Significantly, while both HS and LTI groups exhibited increased VOT across different stop categories, the LTI group, in particular, showed a more marked distinction via F0 compared to both the HS and L1-i groups. Furthermore, the study found that the distance in onset F0 among L1 stop categories becomes greater with the increase in bilingual balance.

In conclusion, the findings of the present study contribute to theoretical frameworks in bilingual speech, expanding our understanding of bilingual phonetics and phonology. These findings expand our knowledge of CLI in bilingual phonetics by lending empirical support to the theoretical postulates of SLM-r, which posits that bilingual speakers are subject to CLI, extending this prediction to speech perception and production among HSs. Future studies may support the findings of the present study further by testing the SLM-r predictions on bilingual category formation through the investigation of bidirectional CLI in L1 and L2. Future research can also expand our understanding of CLI in bilingual speech perception by examining the gradiency in the perception of phonetic categories. Furthermore, the results of the present study are compatible with the assumptions of the Bilingual Alignment Framework, thus contributing to extending its scope to the phonetic and phonological domains.

APPENDIX A. SAMPLE TASK SCREENS

Figure A.1. Sample task screens (left: 3AFC task, right: AX discrimination task)



APPENDIX B. STIMULI LISTS

Table B.1. Stimuli list for the AX discrimination task

Number	Type	Word	International Phonetic Alphabet	Gloss
1	Lenis	방관	[paŋgwan]	look on
2	Lenis	반대	[pande]	objection
3	Lenis	발단	[palt*an]	beginning
4	Lenis	달	[tal]	the moon
5	Lenis	달력	[talɾjʌk]	calendar
6	Lenis	단맛	[tanmat]	sweet taste
7	Lenis	가래	[kaɾe]	phlegm
8	Lenis	갈	[kal]	reed
9	Lenis	갈날	[kalnal]	day to go
10	Aspirated	판관	[pʰaŋgwan]	judge
11	Aspirated	판대	[pʰant*e]	board
12	Aspirated	팔단	[pʰalt*an]	eighth level
13	Aspirated	탈	[tʰal]	mask
14	Aspirated	탄력	[tʰalɾjʌk]	elasticity
15	Aspirated	탄맛	[tʰanmat]	burnt taste
16	Aspirated	카레	[kʰaɾe]	curry
17	Aspirated	칼	[kʰal]	knife
18	Aspirated	칼날	[kʰalnal]	blade

Table B.2. Stimuli list for the controlled reading paradigm task

Word	IPA	Gloss
불	[pul]	fire
발	[pal]	foot
비	[pi]	rain
부르다	[purɨda]	to call
발찌	[paltɕ*i]	anklet
비다	[pida]	empty
두자	[tudʑa]	let's place (something)
대양	[tejaŋ]	ocean
다수	[tasu]	many numbers
도끼	[tok*i]	exe
닫다	[tat*a]	to close
들다	[tɨlda]	to pick (something) up
그림	[kɨrim]	picture
가드	[kadɨ]	guard
기스	[kisɨ]	scratch
그다	[kɨda]	to line
공	[koŋ]	ball
간	[kan]	liver, taste (Noun)
풀	[pʰul]	grass
팔	[pʰal]	hand
피	[pʰi]	blood
푸르다	[pʰurɨda]	clear blue
팔찌	[pʰaltɕ*i]	bracelet
피다	[pʰida]	to bloom
투자	[tʰudʑa]	investment
태양	[tʰejaŋ]	the sun
타수	[tʰasu]	number of hits
토끼	[tʰok*i]	rabbit
탓다	[tʰat*a]	earned, burned
틀다	[tʰɨlda]	to turn on
크림	[kʰɨrim]	cream
카드	[kʰadɨ]	card
키스	[kʰisɨ]	kiss
크다	[kʰɨda]	big
콩	[kʰoŋ]	bean
칸	[kʰan]	room, space
뿔	[p*ul]	horn
빨	[p*al]	nonce word
삐	[p*i]	nonce word
뿌르다	[p*urɨda]	nonce word
빨찌	[p*altɕ*i]	nonce word
삐다	[p*ida]	to sprain
뚜자	[t*udʑa]	nonce word
때양	[t*ejaŋ]	nonce word
따수	[t*asu]	nonce word
또끼	[t*ok*i]	nonce word
땄다	[t*at*a]	won, picked
뜰다	[t*ɨlda]	nonce word
끄림	[k*ɨrim]	nonce word
까드	[k*adɨ]	nonce word
끼스	[k*isɨ]	nonce word
끄다	[k*ɨda]	to turn off
꽁	[k*oŋ]	nonce word
깐	[k*an]	nonce word

APPENDIX C. MINT SPRINT

Figure C.1. Adapted MINT SPRINT for Korean–English bilinguals.

APPENDIX D. BILINGUAL LANGUAGE PROFILE

Questionnaire D.1. Korean version

한국어 영어 이중언어 프로필

본 설문조사는 University of Texas at Austin 의 Center for Open Educational Resources and Language Learning 의 후원 하에 다양한 배경을 가지고 다양한 상황에서 한국어와 영어를 사용하는 이중 언어 사용자들에 대한 더 나은 이해를 돕고자 만들어졌습니다 . 여러분들의 언어 배경, 언어 사용, 언어에 대한 태도, 그리고 언어 능숙도에 관한 아래의 질문에 응답해주시면 됩니다. 본 설문지는 19 개의 질문들로 구성되어 있으며, 대략 10 분정도가 소요됩니다. 정해진 답은 없습니다. 모든 질문에 답해주시고, 솔직한 답변 부탁드립니다. 설문에 응해 주셔서 감사합니다.

I. 인적사항

이름 ____ 날짜 _____/_____/________

나이 _____ 남성 / 여성 / 기타 현재 거주지: 시/도__________ 국가 ____ _____

최고학력: 중졸 이하 고졸 전문대졸/대학교 수료

4 년제 대학교 졸업 석사수료 석사

박사 기타: ____________

II. 언어배경

언어적 배경에 대한 대답으로 알맞은 것을 고르시오.

1. 아래의 언어들을 몇 살때부터 배우기 시작하셨나요?

영어

출생이후 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20+

한국어

출생이후 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20+

2. 아래의 언어들을 몇 살때부터 사용하기 편안하다고 느끼셨나요?

영어

기억하고 있는 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20+ not yet

가장 빠른 나이

한국어

기억하고 있는 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20+ not yet

가장 빠른 나이

3. 몇 년 동안 아래의 언어로 된 수업 (문법, 역사, 수학 등등) 을 들으셨나요 (초등학교부터 대학교까지)?

영어

0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20+

한국어

0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20+

4. 아래의 언어를 말하는 나라에서 몇 년을 지내셨나요?

영어

0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20+

한국어

0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20+

5. 아래의 언어를 말하는 가족과 몇 년을 함께 보내셨나요?

영어

0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20+

한국어

0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20+

6. 아래의 언어를 말하는 직업환경에서 몇 년을 일하셨나요? 3

영어

0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20+

한국어

0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20+4

III. 언어 사용

여러분의 언어적 사용에 대한 대답으로 알맞은 것을 고르시오. 주어진 질문에 모든 언어의 전체 사용은 100%여야합니다.

7. 일상적인 한 주 동안, 친구들과 아래의 언어들을 사용하는 비율이 몇 퍼센트를 차지합니까?

영어

0% 10% 20% 30% 40% 50% 60% 70% 80% 90% 100%

한국어

0% 10% 20% 30% 40% 50% 60% 70% 80% 90% 100%

기타 언어

0% 10% 20% 30% 40% 50% 60% 70% 80% 90% 100%

8. 일상적인 한 주 동안, 가족들과 아래의 언어들을 사용하는 비율이 몇 퍼센트를 차지합니까?

영어

0% 10% 20% 30% 40% 50% 60% 70% 80% 90% 100%

한국어

0% 10% 20% 30% 40% 50% 60% 70% 80% 90% 100%

기타언어

0% 10% 20% 30% 40% 50% 60% 70% 80% 90% 100%

9. 일상적인 한 주 동안, 학교나 직장에서 아래의 언어들을 사용하는 비율이 몇 퍼센트를 차지합니까?

영어

0% 10% 20% 30% 40% 50% 60% 70% 80% 90% 100%

한국어

0% 10% 20% 30% 40% 50% 60% 70% 80% 90% 100%

기타언어

0% 10% 20% 30% 40% 50% 60% 70% 80% 90% 100%

10. 혼잣말을 할 때, 얼마나 자주 아래의 언어로 말합니까?

영어

0% 10% 20% 30% 40% 50% 60% 70% 80% 90% 100%

한국어

0% 10% 20% 30% 40% 50% 60% 70% 80% 90% 100%

기타언어

0% 10% 20% 30% 40% 50% 60% 70% 80% 90% 100%

11. 숫자를 셀 때, 얼마나 자주 아래의 언어들을 사용합니까?

영어

0% 10% 20% 30% 40% 50% 60% 70% 80% 90% 100%

한국어

0% 10% 20% 30% 40% 50% 60% 70% 80% 90% 100%

기타언어

0% 10% 20% 30% 40% 50% 60% 70% 80% 90% 100% 5

IV. 언어 능숙도

본 섹션에서는 여러분의 언어 능숙도를 0 (전혀 능숙하지 않음)에서 부터 6 (아주 능숙함) 까지를 사용해서 평가하시오

0=전혀 능숙하지 않음 6=아주 능숙함

12. a. 영어 말하기 능력이 어느정도 입니까? 0 1 2 3 4 5 6

b. 한국어 말하기 능력이 어느정도 입니까? 0 1 2 3 4 5 6

13. a. 영어 듣기 능력이 어느정도 입니까? 0 1 2 3 4 5 6

b. 한국어 듣기 능력이 어느정도 입니까? 0 1 2 3 4 5 6

14. a. 영어 읽기 능력이 어느정도 입니까? 0 1 2 3 4 5 6

b. 한국어 읽기 능력이 어느정도 입니까? 0 1 2 3 4 5 6

15. a. 영어 쓰기 능력이 어느정도 입니까? 0 1 2 3 4 5 6

b. 한국어 쓰기 능력이 어느정도 입니까? 0 1 2 3 4 5 6 6

V. 언어 태도

여러분의 언어태도에 관한 질문에 0 (동의하지 않음)에서 6 (동의함) 까지를 이용해 대답하시오.

0=동의하지 않음 6=동의함

16. a. **영어로** 말 할 때 나 자신답게 느껴진다. 0 1 2 3 4 5 6

b. **한국어로** 말 할 때 나 자신답게 느껴진다. 0 1 2 3 4 5 6

17. a. 나는 **영어문화권의** 일원인 것 같다. 0 1 2 3 4 5 6

b. 나는 **한국어문화권의** 일원인 것 같다. 0 1 2 3 4 5 6

18. a. 자신이 **영어** 모국어 화자처럼 영어를 사용 (또는 최종적으로 사용) 0 1 2 3 4 5 6

하는 것이 중요하다.

b. 자신이 **한국어** 모국어 화자처럼 한국어를 사용 (또는 최종적으로 사용) 0 1 2 3 4 5 6

하는 것이 중요하다.

19. a. 나는 다른 사람들이 내가 **영어** 모국어화자라고 생각하길 원한다 0 1 2 3 4 5 6

b. 나는 다름 사람들이 내가 **한국어** 모국어화자라고 생각하길 원한다. 0 1 2 3 4 5 6

<u>Questionnaire D.2. English version</u>

Bilingual Language Profile: English-Korean

We would like to ask you to help us by answering the following questions concerning your language history, use, attitudes, and proficiency. This survey was created with support from the Center for Open Educational Resources and Language Learning at the University of Texas at Austin to better understand the profiles of bilingual speakers in diverse settings with diverse backgrounds. The survey consists of 19 questions and will take less than 10 minutes to complete. This is not a test, so there are no right or wrong answers. Please answer every question and give your answers sincerely. Thank you very much for your help.

I. Biographical Information

Name ____ Today's Date _____/_____/________

Age_____ Male / Female / Other Current place of residence: City/state:__________ Country:_________

Highest level of formal education: Less than high school High school Some college

College (B.A., B.S.) Some graduate school Masters

PhD/MD/JD Other: ____________

II. Language history

In this section, we would like you to answer some factual questions about your language history by placing a check in the appropriate box.

1. At what age did you **start learning** the following languages?

English

Since birth 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20+

Korean

Since birth 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20+

2. At what age did you **start to feel comfortable** using the following languages?

English

As early as I 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20+ not yet

can remember

Korean

As early as I 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20+ not yet

can remember

3. How many years of **classes (grammar, history, math, etc.)** have you had in the following languages (primary school through university)?

English

0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20+

Korean

0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20+

4. How many years have you spent in a **country/region** where the following languages are spoken?

English

0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20+

Korean

0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20+

5. How many years have you spent in a **family** where the following languages are spoken?

English

0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20+

Korean

0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20+

6. How many years have you spent in a **work environment** where the following languages are spoken?

English

0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20+

Korean

0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20+

III. Language use

In this section, we would like you to answer some questions about your language use by placing a check in the appropriate box. Total use for
all languages in a given question should equal 100%.

7. In an average week, what percentage of the time do you use the following languages **with friends**?

English

0% 10% 20% 30% 40% 50% 60% 70% 80% 90% 100%

Korean

0% 10% 20% 30% 40% 50% 60% 70% 80% 90% 100%

Other languages

0% 10% 20% 30% 40% 50% 60% 70% 80% 90% 100%

8. In an average week, what percentage of the time do you use the following languages **with family**?

English

0% 10% 20% 30% 40% 50% 60% 70% 80% 90% 100%

Korean

0% 10% 20% 30% 40% 50% 60% 70% 80% 90% 100%

Other languages

0% 10% 20% 30% 40% 50% 60% 70% 80% 90% 100%

9. In an average week, what percentage of the time do you use the following languages **at school/work**?

English

0% 10% 20% 30% 40% 50% 60% 70% 80% 90% 100%

Korean

0% 10% 20% 30% 40% 50% 60% 70% 80% 90% 100%

Other languages

0% 10% 20% 30% 40% 50% 60% 70% 80% 90% 100%

10. When you talk to yourself, how often do you **talk to yourself** in the following languages?

English

0% 10% 20% 30% 40% 50% 60% 70% 80% 90% 100%

Korean

0% 10% 20% 30% 40% 50% 60% 70% 80% 90% 100%

Other languages

0% 10% 20% 30% 40% 50% 60% 70% 80% 90% 100%

11. When you count, how often do you **count** in the following languages?

English

0% 10% 20% 30% 40% 50% 60% 70% 80% 90% 100%

Korean

0% 10% 20% 30% 40% 50% 60% 70% 80% 90% 100%

Other languages

0% 10% 20% 30% 40% 50% 60% 70% 80% 90% 100%

IV. Language proficiency

In this section, we would like you to rate your language proficiency by giving marks from 0 to 6.

0=not well at all 6=very well

12. a. How well do you speak **English**? 0 1 2 3 4 5 6

b. How well do you speak **Korean**? 0 1 2 3 4 5 6

13. a. How well do you understand **English**? 0 1 2 3 4 5 6

b. How well do you understand **Korean**? 0 1 2 3 4 5 6

14. a. How well do you read **English**? 0 1 2 3 4 5 6

b. How well do you read **Korean**? 0 1 2 3 4 5 6

15. a. How well do you write **English**? 0 1 2 3 4 5 6

b. How well do you write **Korean**? 0 1 2 3 4 5 6

5

V. Language attitudes

In this section, we would like you to respond to statements about language attitudes by giving marks from 0-6.

0=disagree 6=agree

16. a. I feel like myself when I speak **English**. 0 1 2 3 4 5 6

b. I feel like myself when I speak **Korean**. 0 1 2 3 4 5 6

17. a. I identify with an **English-speaking** culture. 0 1 2 3 4 5 6

b. I identify with an **Korean-speaking** culture. 0 1 2 3 4 5 6

18. a. It is important to me to use (or eventually use) **English** like a native speaker. 0 1 2 3 4 5 6

b. It is important to me to use (or eventually use) **Korean** like a native speaker. 0 1 2 3 4 5 6

19. a. I want others to think I am a native speaker of **English**. 0 1 2 3 4 5 6

b. I want others to think I am a native speaker of **Korean**. 0 1 2 3 4 5 6

REFERENCES

Abrahamsson, N. (1999). Vowel epenthesis of/sC (C)/onsets in Spanish/Swedish interphonology: A longitudinal case study. *Language Learning*, *49*(3), 473–508.

Abramson, A. S., & Lisker, L. (1985). Relative power of cues: F0 shift versus voice timing. *Phonetic Linguistics: Essays in Honor of Peter Ladefoged*, 25–33.

Ahn, S., & Chang, C. B. (2022). Emotion word development in bilingual children living in majority and minority contexts. *Applied Linguistics*, *43*(5), 845–866.

Ahn, S., Chang, C. B., DeKeyser, R., & Lee-Ellis, S. (2017). Age effects in first language attrition: Speech perception by Korean-English bilinguals. *Language Learning*, *67*(3), 694–733.

Albert, J., & Hu, J. (2019). *Probability and Bayesian modeling*. CRC Press.

Baker, W., & Trofimovich, P. (2005). Interaction of native-and second-language vowel system (s) in early and late bilinguals. *Language and Speech*, *48*(1), 1–27.

Bang, H.-Y., Sonderegger, M., Kang, Y., Clayards, M., & Yoon, T.-J. (2018). The emergence, progress, and impact of sound change in progress in Seoul Korean: Implications for mechanisms of tonogenesis. *Journal of Phonetics*, *66*, 120–144.

Barlow, J. A., Branson, P. E., & Nip, I. S. (2013). Phonetic equivalence in the acquisition of/l/by Spanish–English bilingual children. *Bilingualism: Language and Cognition*, *16*(1), 68–85.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*, 1–48.

Bergmann, C., Nota, A., Sprenger, S. A., & Schmid, M. S. (2016). L2 immersion causes non-native-like L1 pronunciation in German attriters. *Journal of Phonetics*, *58*, 71–86.

Best, C. T. (1994). The emergence of native-language phonological influences in infants: A perceptual assimilation model. In J. Goodman & H. Nusbaum (Eds.), *The development of speech perception: The transition from speech sounds to spoken words*, *167*(224), 233–277. Cambridge, MA: MIT Press.

Best, C. T. (1995). A direct realist perspective on cross-language speech perception. In W. Strange (Ed.), *Speech Perception and Linguistic Experience: Issues in Cross-language Research*, 107–126, Baltimore, MD: York Press.

Best, C. T., & Tyler, M. D. (2007). Nonnative and second-language speech perception: Commonalities and complementarities. In O.-S. Bohn & M. J. Munro (Eds.), *Language experience in second language speech learning*, 13–34. John Benjamins Publishing Company. DOI: https://doi.org/10.1075/lllt.17.07bes

Birdsong, D., Gertken, L. M., & Amengual, M. (2012). Bilingual language profile: An easy-to-use instrument to assess bilingualism. *COERLL, University of Texas at Austin*.

Boersma, P., & Weenink, D. (2022). *Praat: Doing phonetics by computer [Computer program]* (Version 6.3.03) [Computer software].

Bridges, D., Pitiot, A., MacAskill, M. R., & Peirce, J. W. (2020). The timing mega-study: Comparing a range of experiment generators, both lab-based and online. *PeerJ*, *8*, e9414.

Broselow, E., Chen, S.-I., & Wang, C. (1998). The emergence of the unmarked in second language phonology. *Studies in Second Language Acquisition*, *20*(2), 261–280.

Buhusi, C. V., & Meck, W. H. (2005). What makes us tick? Functional and neural mechanisms of interval timing. *Nature Reviews Neuroscience*, *6*(10), 755–765.

Bürkner, P.-C. (2017). brms: An R package for Bayesian multilevel models using Stan. *Journal of Statistical Software*, *80*, 1–28.

Caramazza, A., Yeni-Komshian, G. H., Zurif, E. B., & Carbone, E. (1973). The acquisition of a new phonological contrast: The case of stop consonants in French-English bilinguals. *The Journal of the Acoustical Society of America*, *54*(2), 421–428.

Carpenter, B., Gelman, A., Hoffman, M. D., Lee, D., Goodrich, B., Betancourt, M., Brubaker, M., Guo, J., Li, P., & Riddell, A. (2017). Stan: A probabilistic programming language. *Journal of Statistical Software*, *76*(1).

Carreira, M., & Kagan, O. (2011). The results of the National Heritage Language Survey: Implications for teaching, curriculum design, and professional development. *Foreign Language Annals*, *44*(1), 40–64.

Chang, C. B. (2012). Rapid and multifaceted effects of second-language learning on first-language speech production. *Journal of Phonetics*, *40*(2), 249–268.

Chang, C. B. (2016). Bilingual perceptual benefits of experience with a heritage language. *Bilingualism: Language and Cognition*, *19*(4), 791–809.

Chang, C. B. (2021). Phonetics and phonology of heritage languages. In S. Montrul and M. Polinsky (Eds.), *The Cambridge handbook of heritage languages and linguistics*, 581–612, Cambridge University Press.

Chang, C. B., & Ahn, S. (2023). Examining the role of phoneme frequency in first language perceptual attrition. *Languages*, *8*(1), 53.

Chang, S.-E., & Mandock, K. (2019). A phonetic study of Korean heritage learners' production of Korean word-initial stops. *Heritage Language Journal*, *16*(3), 273–295.

Chang, S.-E., & Weiss-Cowie, S. (2021). Hyper-articulation effects in Korean glides by heritage language learners. *International Journal of Bilingualism*, *25*(1), 3–20.

Cheng, A. (2017). VOT merger and f0 contrast in Heritage Korean in California. *UC Berkeley PhonLab Annual Report, 13*(1).

Cho, T., Jun, S.-A., & Ladefoged, P. (2002). Acoustic and aerodynamic correlates of Korean stops and fricatives. *Journal of Phonetics, 30*(2), 193–228.

Choi, J., Kim, S., & Cho, T. (2020). An apparent-time study of an ongoing sound change in Seoul Korean: A prosodic account. *PLoS ONE 15*(10): e0240682, 1–29.

Cuza, A. (2008). *The L2 acquisition and L1 attrition of the interpretation and use of aspectual properties in Spanish among English-speaking L2 learners and long-term Spanish immigrants*. Ph.D. thesis, University of Toronto, Canada.

Cuza, A., & Sánchez, L. (2022). The acquisition of grammatical gender in child and adult heritage speakers of Spanish. *The Acquisition of Gender: Crosslinguistic Perspectives, 63*, 71–94

Dmitrieva, O. (2019). Transferring perceptual cue-weighting from second language into first language: Cues to voicing in Russian speakers of English. *Journal of Phonetics*, *73*, 128–143.

Dmitrieva, O., Jongman, A., & Sereno, J. (2010). Phonological neutralization by native and non-native speakers: The case of Russian final devoicing. *Journal of Phonetics*, *38*(3), 483–492.

Dmitrieva, O., Llanos, F., Shultz, A. A., & Francis, A. L. (2015). Phonological status, not voice onset time, determines the acoustic realization of onset f0 as a secondary voicing cue in Spanish and English. *Journal of Phonetics, 49*, 77–95.

Eckman, F. R. (1987). The reduction of word-final consonant clusters in interlanguage. *Sound Patterns in Second Language Acquisition*, 143–162.

Elman, J. L., Diehl, R. L., & Buchwald, S. E. (1977). Perceptual switching in bilinguals. *The Journal of the Acoustical Society of America, 62*(4), 971–974.

Escudero, P. (2005). *Linguistic perception and second language acquisition: Explaining the attachment of optimal phonological categorization*. Ph.D. thesis, Utrecht University.

Escudero, P. (2009). The linguistic perception of similar L2 sounds. In P. Boersma & S. Hamann (Eds.), *Phonology in Perception*, *15*, 151–190. De Gruyter.

Escudero, P., Benders, T., & Lipski, S. C. (2009). Native, non-native and L2 perceptual cue weighting for Dutch vowels: The case of Dutch, German, and Spanish listeners. *Journal of Phonetics, 37*(4), 452–465.

Flege, J. E. (1987). The production of "new" and "similar" phones in a foreign language: Evidence for the effect of equivalence classification. *Journal of Phonetics*, *15*(1), 47–65.

Flege, J. E. (1995). Second language speech learning: Theory, findings, and problems. *Speech Perception and Linguistic Experience: Issues in Cross-Language Research*, *92*, 233–277.

Flege, J. E. (2003). Assessing constraints on second-language segmental production and perception. *Phonetics and Phonology in Language Comprehension and Production: Differences and Similarities*, *6*, 319–355.

Flege, J. E., & Bohn, O. S. (2021). The revised speech learning model (SLM-r). *Second language speech learning: Theoretical and empirical progress*, 3–83.

Flege, J. E., Bohn, O.-S., & Jang, S. (1997). Effects of experience on non-native speakers' production and perception of English vowels. *Journal of Phonetics*, *25*(4), 437–470.

Flege, J. E., & Eefting, W. (1987a). Cross-language switching in stop consonant perception and production by Dutch speakers of English. *Speech Communication*, *6*(3), 185–202.

Flege, J. E., & Eefting, W. (1987b). Production and perception of English stops by native Spanish speakers. *Journal of Phonetics*, *15*(1), 67–83.

Flege, J. E., & Eefting, W. (1988). Imitation of a VOT continuum by native speakers of English and Spanish: Evidence for phonetic category formation. *The Journal of the Acoustical Society of America*, *83*(2), 729–740.

Flege, J. E., Schirru, C., & MacKay, I. R. (2003). Interaction between the native and second language phonetic subsystems. *Speech Communication*, *40*(4), 467–491.

Francis, A. L., Baldwin, K., & Nusbaum, H. C. (2000). Effects of training on attention to acoustic cues. *Perception & Psychophysics*, *62*(8), 1668–1680. https://doi.org/10.3758/BF03212164

Francis, A. L., & Nusbaum, H. C. (2002). Selective attention and the acquisition of new phonetic categories. *Journal of Experimental Psychology: Human Perception and Performance*, *28*(2), 349.

Franken, M. K., Acheson, D. J., McQueen, J. M., Eisner, F., & Hagoort, P. (2017). Individual variability as a window on production-perception interactions in speech motor control. *The Journal of the Acoustical Society of America*, *142*(4), 2007–2018.

Garcia, D. L., & Gollan, T. H. (2021). The MINT Sprint: Exploring a Fast Administration Procedure with an Expanded Multilingual Naming Test. *Journal of the International Neuropsychological Society*, 1–17.

Garcia-Sierra, A., Diehl, R. L., & Champlin, C. (2009). Testing the double phonemic boundary in bilinguals. *Speech Communication*, *51*(4), 369–378.

Gerrits, E., & Schouten, M. E. (2004). Categorical perception depends on the discrimination task. *Perception & Psychophysics*, *66*(3), 363–376.

Gibbon, J. (1977). Scalar expectancy theory and Weber's law in animal timing. *Psychological Review*, *84*(3), 279.

Gibson, E. J. (1969). *Principles of perceptual learning and development.* Appleton-Century-Crofts.

Goldstone, R. L. (1998). PERCEPTUAL LEARNING. *Annual Review of Psychology*, *49*(1), 585–612. https://doi.org/10.1146/annurev.psych.49.1.585

Graves, A., Mohamed, A., & Hinton, G. (2013). Speech recognition with deep recurrent neural networks. In *Proceedings of the 2013 IEEE International Conference on Acoustics, Speech and Signal Processing*, 6645–6649.

Green, P., & MacLeod, C. J. (2016). SIMR: An R package for power analysis of generalized linear mixed models by simulation. *Methods in Ecology and Evolution*, *7*(4), 493–498.

Guion, S. G. (2003). The vowel systems of Quichua-Spanish bilinguals. *Phonetica*, *60*(2), 98–128.

Han, M. S., & Weitzman, R. S. (1970). Acoustic features of Korean/P, T, K/,/p, t, k/and/ph, th, kh. *Phonetica*, *22*(2), 112–128.

Hancin-Bhatt, B., & Bhatt, R. M. (1997). Optimal L2 syllables: Interactions of transfer and developmental effects. *Studies in Second Language Acquisition*, 331–378.

Harada, T. (2003). L2 influence on L1 speech in the production of VOT. In *Proceedings of the 15th International Congress of Phonetic Sciences (ICPhS)*, 1085–1088.

Hardcastle, W. J. (1973). Some observations on the tense-lax distinction in initial stops in Korean. *Journal of Phonetics*, *1*(3), 263–272.

Hazan, V., & Boulakia, G. (1993). Perception and production of a voicing contrast by French-English bilinguals. *Language and Speech*, *36*(1), 17–38.

Hazan, V., Kim, J., & Chen, Y. (2010). Audiovisual perception in adverse conditions: Language, speaker and listener effects. *Speech Communication*, *52*(11–12), 996–1009.

Hazan, V., Messaoud-Galusi, S., Rosen, S., Nouwens, S., & Shakespeare, B. (2009). Speech perception abilities of adults with dyslexia: Is there any evidence for a true deficit? *Journal of Speech, Language, and Hearing Research*, *52*(6), 1510–1529.

Holt, L. L., & Lotto, A. J. (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.

House, A. S., & Fairbanks, G. (1953). The influence of consonant environment upon the secondary acoustical characteristics of vowels. *The Journal of the Acoustical Society of America*, *25*(1), 105–113.

Hur, E., Lopez Otero, J. C., & Sánchez, L. (2020). Gender agreement and assignment in Spanish heritage speakers: Does frequency matter? *Languages*, *5*(4), 48.

Idemaru, K., & Holt, L. L. (2011). Word recognition reflects dimension-based statistical learning. *Journal of Experimental Psychology: Human Perception and Performance*, *37*(6), 1939.

Johnson, K., & Babel, M. (2010). On the perceptual basis of distinctive features: Evidence from the perception of fricatives by Dutch and English speakers. *Journal of Phonetics*, *38*(1), 127–136.

Jung, Y.-J. (2023). *Production and perception of Korean and Korean-accented English clear speech*. Ph.D. thesis, Purdue University.

Kachlicka, M., Saito, K., & Tierney, A. (2019). Successful second language learning is tied to robust domain-general auditory processing and stable neural representation of sound. *Brain and language*, *192*, 15–24.

Kagaya, R. (1974). A fiberscopic and acoustic study of the Korean stops, affricates and fricatives. *Journal of Phonetics*, *2*(2), 161–180.

Kajouj, F., & Kager, R. (2019). Effects of bilingualism on cue weighting: How do bilingual children perceive the Dutch [ɑ]-[a:] contrast? *International Journal of Bilingualism*, *23*(2), 509–524.

Kang, K.-H., & Guion, S. G. (2006). Phonological systems in bilinguals: Age of learning effects on the stop consonant systems of Korean-English bilinguals. *The Journal of the Acoustical Society of America*, *119*(3), 1672–1683.

Kang, K.-H., & Guion, S. G. (2008). Clear speech production of Korean stops: Changing phonetic targets and enhancement strategies. *The Journal of the Acoustical Society of America*, *124*(6), 3909–3917.

Kang, Y. (2014). Voice Onset Time merger and development of tonal contrast in Seoul Korean stops: A corpus study. *Journal of Phonetics*, 45, 76–90.

Kang, Y., & Nagy, N. (2016). VOT merger in Heritage Korean in Toronto. *Language Variation and Change*, *28*(2), 249–272.

Kapnoula, E. C., & McMurray, B. (2021). Idiosyncratic use of bottom-up and top-down information leads to differences in speech perception flexibility: Converging evidence from ERPs and eye-tracking. *Brain and Language*, *223*, 105031.

Kapnoula, E. C., Winn, M. B., Kong, E. J., Edwards, J., & McMurray, B. (2017). Evaluating the sources and functions of gradiency in phoneme categorization: An individual differences approach. *Journal of Experimental Psychology: Human Perception and Performance*, *43*(9), 1594.

Kim, C.-W. (1965). On the autonomy of the tensity feature in stop classification (with special reference to Korean stops). *Word*, *21*(3), 339–359.

Kim, H. (2023). *Tracking the Time Course of Phonetic Cue Integration in the Perception of Korean Stop Contrasts by Korean and English Listeners*. Ph.D. thesis, University of Kansas.

Kim, H., & Jongman, A. (2022). The influence of inter-dialect contact on the Korean three-way laryngeal distinction: An acoustic comparison among Seoul Korean speakers and Gyeongsang speakers with limited and extended residence in Seoul. *Language and Speech*, *65*(3), 531–553.

Kim, H., & Tremblay, A. (2021). Korean listeners' processing of suprasegmental lexical contrasts in Korean and English: A cue-based transfer approach. *Journal of Phonetics*, *87*, 101059.

Kim, M., & Stoel-Gammon, C. (2009). The acquisition of Korean word-initial stops. *The Journal of the Acoustical Society of America*, *125*(6), 3950–3961.

Kim, M.-R., Beddor, P. S., & Horrocks, J. (2002). The contribution of consonantal and vocalic information to the perception of Korean initial stops. *Journal of Phonetics, 30*(1), 77–100.

Kondaurova, M. V., & Francis, A. L. (2010). The role of selective attention in the acquisition of English tense and lax vowels by native Spanish listeners: Comparison of three training methods. *Journal of Phonetics*, *38*(4), 569–587.

Kong, E. J. (2012). Perception of Korean stops with a three-way laryngeal contrast. *Phonetics and Speech Sciences*, *4*(1), 13–20.

Kong, E. J., Beckman, M. E., & Edwards, J. (2011). Why are Korean tense stops acquired so early?: The role of acoustic properties. *Journal of Phonetics*, *39*(2), 196–211.

Kong, E. J., & Edwards, J. (2016). Individual differences in categorical perception of speech: Cue weighting and executive function. *Journal of Phonetics*, *59*, 40–57.

Kong, E. J., Holliday, J. J., & Lee, H. (2022). Post-adolescent changes in the perception of regional sub-phonemic variation. *Journal of Phonetics*, *90*, 101114.

Kong, E. J., & Kang, S. (2023). Individual differences in categorical judgment of L2 stops: A link to proficiency and acoustic cue-weighting. *Language and Speech*, 00238309221108647.

Kong, E. J., & Yoon, I. H. (2013). L2 proficiency effect on the acoustic cue-weighting pattern by Korean L2 learners of English: Production and perception of English stops. *Phonetics and Speech Sciences*, *5*(4), 81–90.

Korea Ministry of Education. (2015). General guidelines for elementary, middle, and high school curriculums. Retrieved from https://www.moe.go.kr/boardCnts/viewRenew.do?boardID=141&lev=0&statusYN=C&s=moe&m=0404&opType=N&boardSeq=60747

Kutlu, E., Chiu, S., & McMurray, B. (2022). Moving away from deficiency models: Gradiency in bilingual speech categorization. *Frontiers in Psychology*, 7428.

Lang, B., & Davidson, L. (2019). Effects of exposure and vowel space distribution on phonetic drift: Evidence from American English learners of French. *Language and Speech*, *62*(1), 30–60.

Law, W. L., Dmitrieva, O., & Francis, A. L. (2021). Language attitudes modulate phonetic interactions between languages in bilingual speakers in diglossic settings. *Linguistic Approaches to Bilingualism*, *11*(3), 289–322.

Law, W. L., & Francis, A. L. (2015). Phonetic divergence in bilingual speakers is modulated by language attitude. *The Journal of the Acoustical Society of America*, *138*(3), 1945–1945.

Lee, H., Holliday, J. J., & Kong, E. J. (2020). Diachronic change and synchronic variation in the Korean stop laryngeal contrast. *Language and Linguistics Compass*, *14*(7), e12374.

Lee, H., & Jongman, A. (2012). Effects of tone on the three-way laryngeal distinction in Korean: An acoustic and aerodynamic comparison of the Seoul and South Kyungsang dialects. *Journal of the International Phonetic Association*, *42*(2), 145–169.

Lee, H., & Jongman, A. (2019). Effects of sound change on the weighting of acoustic cues to the three-way laryngeal stop contrast in Korean: Diachronic and dialectal comparisons. *Language and Speech*, *62*(3), 509–530.

Lee, H., Politzer-Ahles, S., & Jongman, A. (2013). Speakers of tonal and non-tonal Korean dialects use different cue weightings in the perception of the three-way laryngeal stop contrast. *Journal of Phonetics*, *41*(2), 117–132.

Lee, S., Potamianos, A., & Narayanan, S. (1999). Acoustics of children's speech: Developmental changes of temporal and spectral parameters. *The Journal of the Acoustical Society of America*, *105*(3), 1455–1468.

Lee, S. A. S., & Iverson, G. K. (2012). Stop consonant productions of Korean–English bilingual children. *Bilingualism: Language and Cognition*, *15*(2), 275–287.

Lee-Ellis, S. (2012). *Looking into Bilingualism through the heritage speaker's mind*. Ph.D. thesis, University of Maryland, College Park.

Lisker, L., & Abramson, A. S. (1964). A cross-language study of voicing in initial stops: Acoustical measurements. *Word*, *20*(3), 384–422.

Llanos, F., Dmitrieva, O., Shultz, A., & Francis, A. L. (2013). Auditory enhancement and second language experience in Spanish and English weighting of secondary voicing cues. *The Journal of the Acoustical Society of America*, *134*(3), 2213–2224.

Lord, G. (2008). Second language acquisition and first language phonological modification. In *Selected Proceedings of the 10th Hispanic Linguistics Symposium*, 184–193.

Mack, M. (1990). Phonetic transfer in a French-English bilingual child. *Language Attitudes and Language Conflict*, 107–124.

MacKain, K. S., Best, C. T., & Strange, W. (1981). Categorical perception of English/r/and/l/by Japanese bilinguals. *Applied Psycholinguistics*, *2*(4), 369–390.

MacKay, I. R., Flege, J. E., Piske, T., & Schirru, C. (2001). Category restructuring during second-language speech acquisition. *The Journal of the Acoustical Society of America*, *110*(1), 516–528.

MacLeod, A. A., Stoel-Gammon, C., & Wassink, A. B. (2009). Production of high vowels in Canadian English and Canadian French: A comparison of early bilingual and monolingual speakers. *Journal of Phonetics*, *37*(4), 374–387.

Major, R. C. (1992). Losing English as a first language. *The Modern Language Journal*, *76*(2), 190–208.

Major, R. C. (1996). L2 acquisition, L1 loss, and the critical period hypothesis. *Second-Language Speech: Structure and Process*, 147–159.

Makowski, D., Ben-Shachar, M. S., Chen, S. A., & Lüdecke, D. (2019a). Indices of effect existence and significance in the Bayesian framework. *Frontiers in Psychology*, *10*, 2767.

Makowski, D., Ben-Shachar, M. S., & Lüdecke, D. (2019b). bayestestR: Describing effects and their uncertainty, existence and significance within the Bayesian framework. *Journal of Open Source Software*, *4*(40), 1541.

Massaro, D. W., & Cohen, M. M. (1983). Categorical or continuous speech perception: A new test. *Speech Communication*, *2*(1), 15–35.

McElreath, R. (2020). *Statistical rethinking: A Bayesian course with examples in R and Stan*. Chapman and Hall/CRC.

McMurray, B., Tanenhaus, M. K., & Aslin, R. N. (2002). Gradient effects of within-category phonetic variation on lexical access. *Cognition*, *86*(2), B33–B42.

Mikolov, T., Karafiát, M., Burget, L., Cernockỳ, J., & Khudanpur, S. (2010). Recurrent neural network based language model. *Interspeech*, 2(3), 1045–1048.

Montrul, S. (2002). Incomplete acquisition and attrition of Spanish tense/aspect distinctions in adult bilinguals. *Bilingualism: Language and cognition*, *5*(1), 39–68.

Montrul, S. (2008). *Incomplete acquisition in bilingualism: Re-examining the age factor*. *Amsterdam*: John Benjamins. DOI: 10.1075/ sibil.39.

Oh, E. (2019). Korean-English bilingual children's production of stop contrasts. *Phonetics and Speech Sciences*, *11*(3), 1–7.

Oh, J. S., Au, T. K.-F., & Jun, S.-A. (2010). Early childhood language memory in the speech perception of international adoptees. *Journal of Child Language*, *37*(5), 1123–1132.

Oh, J. S., Jun, S.-A., Knightly, L. M., & Au, T. K. (2003). Holding on to childhood language memory. *Cognition*, *86*(3), B53–B64.

Oh, M., & Daland, R. (2011). Stops and Phrasing in Korean and English Monolinguals and Bilinguals. In *Proceedings of the 17th International Congress of Phonetic Sciences (ICPhS)*, 1530–1533.

Ohde, R. N. (1984). Fundamental frequency as an acoustic correlate of stop consonant voicing. *The Journal of the Acoustical Society of America*, *75*(1), 224–230.

Olson, D. J. (2023). Measuring bilingual language dominance: An examination of the reliability of the Bilingual Language Profile. *Language Testing*, 02655322221139162.

Park, H., & de Jong, K. J. (2008). Perceptual category mapping between English and Korean prevocalic obstruents: Evidence from mapping effects in second language identification skills. *Journal of Phonetics*, *36*(4), 704–723.

Peirce, J., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., Kastman, E., & Lindeløv, J. K. (2019). PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods*, *51*(1), 195–203.

Perez-Cortes, S., Putnam, M. T., & Sánchez, L. (2019). Differential access: Asymmetries in accessing features and building representations in heritage language grammars. *Languages*, *4*(4), 81.

Perkell, J. S., Guenther, F. H., Lane, H., Matthies, M. L., Stockmann, E., Tiede, M., & Zandipour, M. (2004). The distinctness of speakers' productions of vowel contrasts is related to their discrimination of the contrasts. *The Journal of the Acoustical Society of America*, *116*(4), 2338–2344.

Perrachione, T. K., Gabrieli, J. D., & Finn, A. S. (2023). Learning to distinguish the three-way laryngeal contrast of Korean plosives by native English speakers. In *Proceedings of the 20th International Congress of Phonetic Sciences (ICPhS)*, 2686–2690.

Picard, M. (2002). The differential substitution of English/θ ð/in French: The case against underspecification in L2 phonology. *Lingvisticæ Investigationes*, *25*(1), 87–96.

Polinsky, M. (2011). Reanalysis in adult heritage language: New evidence in support of attrition. *Studies in Second Language Acquisition*, *33*(2), 305–328.

Polinsky, M. (2018). *Heritage languages and their speakers* (Vol. 159). Cambridge University Press.

R Core Team. (2022). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Sánchez, L. (2019). Bilingual alignments. *Languages*, *4*(4), 82.

Schertz, J., Cho, T., Lotto, A., & Warner, N. (2015). Individual differences in phonetic cue use in production and perception of a non-native sound contrast. *Journal of Phonetics*, *52*, 183–204.

Schmid, M. S. (2011). *Language attrition*. Cambridge University Press.

Schmidt, A. M. (1996). Cross-language identification of consonants. Part 1. Korean perception of English. *The Journal of the Acoustical Society of America*, *99*(5), 3201–3211.

Schmidt, A. M. (2007). Cross-language consonant identification. *Language Experience in Second Language Speech Learning*, 185200.

Schwartz, B. D., & Sprouse, R. A. (1996). L2 cognitive states and the full transfer/full access model. *Second Language Research*, *12*(1), 40–72. DOI: https://doi.org/10.1177/026765839601200103

Seo, Y., & Dmitrieva, O. (2024). L2 cross-linguistic influence on L1 perception: Evidence from heritage speakers and long-term immigrants. *Journal of Phonetics*, *104*, 101314. DOI: https://doi.org/10.1016/j.wocn.2024.101314

Seo, Y., Dmitrieva, O., & Cuza, A. (2022). Crosslinguistic influence in the discrimination of Korean stop contrast by heritage speakers and second language learners. *Languages*, *7*(1), 6.

Seo, Y., & Olson, J. D. (in press). Phonetic shifts in bilingual vowels: Evidence from intersentential and intrasentential code-switching. *International Journal of Bilingualism*.

Shin, N., Cuza, A., & Sánchez, L. (2023). Structured variation, language experience, and crosslinguistic influence shape child heritage speakers' Spanish direct objects. *Bilingualism: Language and Cognition*, *26*(2), 317–329.

Shrem, Y., Goldrick, M., & Keshet, J. (2019). Dr. VOT: Measuring positive and negative voice onset time in the wild. ArXiv Preprint ArXiv:1910.13255

Shultz, A. A., Francis, A. L., & Llanos, F. (2012). Differential cue weighting in perception and production of consonant voicing. *The Journal of the Acoustical Society of America*, *132*(2), EL95–EL101.

Silva, D. J. (2006). Acoustic evidence for the emergence of tonal contrast in contemporary Korean. *Phonology*, *23*(2), 287–308.

Simonet, M. (2014). Phonetic consequences of dynamic cross-linguistic interference in proficient bilinguals. *Journal of Phonetics*, *43*, 26–37.

Stevenson, P. W. (1973). Reaction time measurements in speech discrimination tasks—An automated system with closed response sets. *Journal of Phonetics*, *1*(4), 347–367.

Sundara, M., Polka, L., & Baum, S. (2006). Production of coronal stops by simultaneous bilingual adults. *Bilingualism: Language and Cognition*, *9*(1), 97–114.

Tremblay, A., Broersma, M., Zeng, Y., Kim, H., Lee, J., & Shin, S. (2021). Dutch listeners' perception of English lexical stress: A cue-weighting approach. *The Journal of the Acoustical Society of America*, *149*(6), 3703–3714.

Trofimovich, P., Gatbonton, E., & Segalowitz, N. (2007). A dynamic look at L2 phonological learning: Seeking processing explanations for implicational phenomena. *Studies in Second Language Acquisition*, *29*(3), 407–448.

Valdés, G. (2005). Bilingualism, heritage language learners, and SLA research: Opportunities lost or seized? *The Modern Language Journal*, *89*(3), 410–426.

Whalen, D. H., Abramson, A. S., Lisker, L., & Mody, M. (1993). F 0 gives voicing information even with unambiguous voice onset times. *The Journal of the Acoustical Society of America*, *93*(4), 2152–2159.

Williams, L. (1977). The perception of stop consonant voicing by Spanish-English bilinguals. *Perception & Psychophysics*, *21*(4), 289–297.

Winn, M. B. (2020). Manipulation of voice onset time in speech stimuli: A tutorial and flexible Praat script. *The Journal of the Acoustical Society of America*, *147*(2), 852–866.

Wrembel, M., Marecka, M., & Kopečková, R. (2019). Extending perceptual assimilation model to L3 phonological acquisition. *International Journal of Multilingualism*, *16*(4), 513–533.

Yamada, R. A., & Tohkura, Y. (1992). Perception of American English/r/and/l/by native speakers of Japanese. *Speech Perception, Production, and Linguistic Structure*, 155–174.

Yoon, S.-Y. (2015). Acoustic properties of Korean stops as L1 produced by L2 learners of the English language. *Communication Sciences & Disorders*, *20*(2), 178–188.

Yusa, N., Nasukawa, K., Koizumi, M., Kim, J., Kimura, N., & Emura, K. (2010). Unexpected effects of the second language on the first. *K. Dziubalska-Ko Laczyk, M. Wrembel, and M. Kul (Eds.), New Sounds 2010: Proceedings of the 6th International Symposium on the Acquisition of Second Language Speech, Poznan, Poland*, 580–584.