

Article

Crosslinguistic Influence in the Discrimination of Korean Stop Contrast by Heritage Speakers and Second Language Learners

Yuhyeon Seo ^{1,*}, Olga Dmitrieva ² and Alejandro Cuza ²¹ Department of Linguistics, Purdue University, West Lafayette, IN 47907, USA² School of Languages and Cultures, Purdue University, West Lafayette, IN 47907, USA; odmitrie@purdue.edu (O.D.); acuza@purdue.edu (A.C.)

* Correspondence: seo86@purdue.edu

Abstract: The present study examines the extent of crosslinguistic influence from English as a dominant language in the perception of the Korean lenis–aspirated contrast among Korean heritage speakers in the United States (N = 20) and English-speaking learners of Korean as a second language (N = 20), as compared to native speakers of Korean immersed in the first language environment (N = 20), by using an AX discrimination task. In addition, we sought to determine whether significant dependencies could be observed between participants’ linguistic background and experiences and their perceptual accuracy in the discrimination task. Results of a mixed-effects logistic regression model demonstrated that heritage speakers outperformed second language learners with 85% vs. 63% accurate discrimination, while no significant difference was detected between heritage speakers and first language-immersed native speakers (85% vs. 88% correct). Furthermore, higher verbal fluency was significantly predictive of greater perceptual accuracy for the heritage speakers. The results are compatible with the interpretation that the influence of English on the discrimination of the Korean laryngeal contrast was stronger for second language learners of Korean than for heritage speakers, while heritage speakers were not apparently affected by dominance in English in their discrimination of Korean lenis and aspirated stops.

Keywords: Korean; English; heritage; L2; discrimination; lenis; aspirated

Citation: Seo, Yuhyeon, Olga Dmitrieva, and Alejandro Cuza. 2022. Crosslinguistic Influence in the Discrimination of Korean Stop Contrast by Heritage Speakers and Second Language Learners.

Languages 7: 6. <https://doi.org/10.3390/languages7010006>

Academic Editors: Christine Shea and Ji Young Kim

Received: 8 October 2021

Accepted: 20 December 2021

Published: 4 January 2022

Publisher’s Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Heritage speakers (HS) are defined as individuals raised in homes where a minority language is spoken, while a different language is dominantly spoken in their country of residence (Valdés 2000, 2005). According to the U.S. Census Bureau (2018), there are 1.1 million Korean HSs in the US who are, by definition, bilingual speakers (Korean and English) to some extent. Although they represent a distinct, and in some ways unique, category of bilinguals, the existing models of additional language acquisition can be extended to account for heritage language development. Specifically, the Speech Learning Model (SLM/SLM-r Flege 1995, 2003; Flege and Bohn 2020), developed primarily with adult second language learners in mind, postulates that coexisting sound systems share a joint phonetic space, where first (L1) and second language (L2) sound categories jostle for position. To our knowledge, SLM is the only model that explicitly predicts that acquisition of additional languages can influence learners’ L1. A number of mechanisms of SLM predict various types of crosslinguistic influence at different stages of L2 acquisition, including mergers or assimilation, as well as dissimilation, but the critical prediction of the theory is that crosslinguistic influence is expected between the sound categories of L1 and L2. The model assumes that crosslinguistic influence is primarily imposed on the sound categories which are phonetically and phonologically similar but not acoustically identical across the two languages (Flege 1987). A crosslinguistic link is established among such categories, which results in mutual attraction (or, at times, repulsion). For the purposes of

this study, we define crosslinguistic influence as phonetic changes in a sound category of one language which make it more similar to a related sound category of another language. The goal of this study was to determine to what extent such crosslinguistic influence from English on Korean laryngeal categories in stop consonants can affect the discrimination of Korean lenis and aspirated stops by HSs and L2 learners of Korean.

Crosslinguistic influence in bilingual L1 speech production is well established (Baker and Trofimovich 2005; Bergmann et al. 2016; Caramazza et al. 1973; Chang 2012; Dmitrieva et al. 2010; Flege 1987; Flege and Eefting 1987a, 1987b; Fowler et al. 2008; Guion 2003; Harada 2003; Lang and Davidson 2019; Major 1992; Peng 1993; Sancier and Fowler 1997; Stoehr et al. 2017). For example, Dmitrieva et al. (2010) showed that Russian L1 immigrants residing in the US pronounced Russian word-final obstruents in a more English-like fashion than monolingual Russian speakers. Similarly, Chang (2012) showed that English-speaking learners of Korean drifted towards Korean norms in their pronunciation of English vowels and word-initial stops during an intensive language course in South Korea.

In perception, evidence is less abundant. Nevertheless, recent work demonstrates that experience with additional languages can affect the way native speech is perceived. For example, in Dmitrieva (2019), Russian (L1)–English (L2) bilinguals applied English-like perceptual strategies in identifying Russian (L1) sounds (see also Antoniou et al. 2012; Garcia-Sierra et al. 2009; Hazan and Boulakia 1993). This evidence suggests that crosslinguistic influence between sound categories that are comparable across languages can result in the restructuring of L1 production targets and L1 perceptual representations.

The present work examines the extent of crosslinguistic influence from English on the perception of Korean stop consonants in Korean HSs and English-speaking L2 learners. Korean exemplifies a three-way laryngeal distinction in stop consonants: fortis (also known as tense), lenis (also known as lax or plain), and aspirated, e.g., [t* – t – t^h]. The Korean stop contrasts are typologically unusual and perceptually challenging, but at the same time sufficiently similar to the English laryngeal categories to trigger crosslinguistic interactions (Ahn et al. 2017). In fact, previous research contains evidence suggestive of such crosslinguistic influence between English and Korean laryngeal categories in HSs and other types of bilinguals, both in production and perception (Chang and Mandock 2019; Cheon and Lee 2013; Cheng 2017). In the current study, we focused on the perception of word-initial lenis–aspirated distinction in Korean stops—a contrast that has already received plentiful attention in the literature but remains an attractive topic due to its uncommon phonetic implementation. To our knowledge, this is the first study specifically examining the perception of lenis–aspirated contrast by HSs using the methodological approach outlined in the later portion of this section.

In the remainder of this introduction, we review phonetic properties of Korean laryngeal stop contrasts as well as developmental patterns in their acquisition and provide a comparison to the phonetics of English voicing distinctions in stops. We also review crosslinguistic assimilation patterns between Korean and English series of stops and zoom in on the Korean laryngeal categories that are the most challenging for both L1 and L2 learners. We then formulate the predicted directions of crosslinguistic interactions between Korean and English laryngeal categories and review existing findings supporting these predictions. Finally, we propose our hypotheses and provide more details about the design of the study.

Two phonetic parameters are majorly involved in distinguishing the three Korean stop contrasts in production and cueing their identification in perception: voice onset time or VOT (time elapsed between the release of a stop and the onset of vocal fold vibration) and onset F0 (fundamental frequency at the onset of the vowel following the stop consonant). Fortis stops have the shortest VOT (about 20 ms, according to Kang and Guion 2008), followed by lenis and then aspirated stops, both of which have relatively long VOTs (about 70 ms for both, from Kang and Guion 2008). Lenis stops are distinguished from fortis and aspirated ones via the lowest onset F0 (Cho et al. 2002; Han and Weitzman 1970; Kagaya 1974; Kim et al. 2002; Kong et al. 2011; Lee et al. 2013, 2020; Lee and Jongman 2019). Moreover,

the word-initial lenis–aspirated stop contrast is believed to be undergoing tonogenesis—the emergence of a tonal distinction on the basis of existing consonantal laryngeal contrasts and the ultimate replacement of the laryngeal categories by tone [Kang 2014](#). Specifically, an ongoing sound change in the Seoul-Gyeonggi dialect of Korean is merging the VOTs of lenis and aspirated stops, especially among younger speakers ([Bang et al. 2018](#); [Chang and Mandock 2019](#); [Kang 2014](#); [Kang and Guion 2008](#); [Kong and Yoon 2013](#); [Silva 2006](#)), with onset F0 becoming the dominant cue to this distinction ([Kang et al. 2010](#); [Kim et al. 2002](#); [Kong et al. 2011](#); [Lee et al. 2013, 2020](#); [Lee and Jongman 2019](#)).

The English language, in contrast, has a two-category contrast between voiced and voiceless stops. In the word-initial position, voiced stops are characterized by shorter VOTs while voiceless stops have longer VOTs ([Abramson and Lisker 1985](#)). Onset F0 also correlates with phonological voicing in English ([Dmitrieva et al. 2015](#); [Ohde 1984](#); [House and Fairbanks 1953](#); [Lehiste and Peterson 1961](#)) but plays a decisively secondary role in perception ([Idemaru et al. 2012](#); [Llanos et al. 2013](#); [Whalen et al. 1993](#)).

Developmental evidence indicates that it takes longer for Korean-speaking children to master native laryngeal categories (not before the age of 4) than for English-speaking children to acquire English voicing categories (fully developed by the age of 2 or 3, depending on the source) ([Bernthal et al. 2013](#); [Jun 2007](#); [Kim and Stoel-Gammon 2009](#); [Lowenstein and Nitttrouer 2008](#)). Furthermore, fortis consonants are the earliest to be acquired in Korean (by 17 months of age), while lenis ones, and their contrast with aspirated stops, in particular, are the last to be mastered ([Choi et al. 2019](#); [Jun 2007](#); [Kim and Stoel-Gammon 2009](#)). [Kong et al. \(2011\)](#) suggested that early acquisition of fortis stops is due to the fact that only VOT needs to be mastered in order to distinguish them from the other categories. In contrast, the necessity to use F0 in order to identify lenis stops makes them acquisitionally challenging, both for L1 and L2 learners ([Chang and Mandock 2019](#); [Cheon and Lee 2013](#); [Ko 2018](#); [Oh et al. 2010](#)). With respect to our study, relatively late acquisition of Korean laryngeal categories, especially the lenis–aspirated contrast, and its perceptual complexity, may make this contrast more susceptible to crosslinguistic influence from the dominant language in HSs and L2 learners.

In terms of the perceived similarities between English and Korean stop categories, both English-immersed non-heritage and HSs of Korean readily assimilate Korean aspirated stops to English voiceless stops ([Cheon and Lee 2013](#); [Schmidt 1996](#)). Crosslinguistic assimilation patterns are less clear-cut for the rest of the Korean categories. Nevertheless, [Cheon and Lee \(2013\)](#) showed that non-heritage and HSs of Korean perceived Korean lenis stops as better exemplars of the English voiced than voiceless category. These findings suggest that the Korean lenis–aspirated pair can be assimilated to the English voiced and voiceless stops, respectively, by Korean–English bilinguals. Given their acoustic and perceptual similarity, SLM predicts that Korean and English laryngeal categories can influence each other in the speech of bilinguals.

Previous research indeed demonstrated patterns compatible with the crosslinguistic influence between Korean and English laryngeal categories in bilingual speakers. For example, [Kang and Guion \(2006\)](#) showed that Korean–English bilinguals produced a greater distinction between English voiced and voiceless stops in terms of onset F0 than English monolinguals (see also [Kim 2012](#); [Kong and Yoon 2013](#)). Reliance on onset F0 in the perception of voicing in English was also shown to be greater for Korean speakers than for English speakers ([Kim 1994](#); [Kong and Yoon 2013](#); [Kim 1994](#)). Given the fact that Korean laryngeal contrasts are cued heavily via onset F0 while English voiced–voiceless contrast relies almost exclusively on VOT, these findings strongly suggest the effect of Korean on bilinguals' English.

On the flip side, as a result of English affecting Korean, several studies reported that bilinguals (HSs in particular)' Korean productions of the lenis–aspirated contrast exhibited a greater reliance on VOT and a lesser reliance on F0, in contrast with monolingual speakers of Korean ([Cheng 2017](#); [Kang and Nagy 2016](#); [Oh and Daland 2011](#); see also ([Lee and Iverson 2012](#); [Oh 2019](#); [Yoon 2015](#)), for evidence from bilingual children). Unfortunately,

to our knowledge, no previous work has examined cue-weighting in perception of Korean stops among Korean–English bilinguals in order to determine whether their perceptual reliance on VOT was also greater than in Korean monolinguals. Kong (2012) reported that while F0 dominated the perception of lenis stops by native Korean speakers living in the USA, VOT also contributed to the categorization of Korean lenis stops. This result could be interpreted as evidence of English affecting the perception of Korean, but in the absence of a monolingual control group, this interpretation remains tentative.

Thus, previous research provides evidence of mutual crosslinguistic influence between Korean and English laryngeal categories in the bilingual speech production and perception. Building on this evidence, we hypothesize that Korean HSs in the US will perceive the Korean lenis–aspirated contrast differently from L1-immersed (L1-i) speakers, due to the influence of English. We assume that the difference will be due to the non-native weighting of perceptual cues to the contrast (a lesser reliance on F0 and a greater reliance on VOT, compared to L1-i speakers). Research in L2 acquisition demonstrates that listeners who do not sufficiently utilize in perception an acoustic dimension that is highly involved in implementing the contrast in production, may underperform on discrimination or in the identification of relevant categories (Cebrian 2006; Yamada and Tohkura 1990, 1992).

The present study represents the first step in the investigation of the hypothesis outlined above. Our main research question concerns the extent to which HSs' discrimination of the Korean lenis–aspirated stop contrast differs from that of L2 learners and L1-i speakers of Korean. We make the additional comparison between HSs and L2 learners of Korean (English L1) in order to gain a better understanding of the roles of the age of onset of acquisition, and the roles of experiential factors, such as language use and exposure. Since L2 learners typically begin acquiring their L2 significantly later than HSs, and often do not have equivalent opportunities for language exposure and practice, we hypothesized that the HSs will outperform the L2 learners. As part of our hypothesis, we also expect both groups to behave differently from the L1-i speakers due to contact with English. Finally, we investigate whether aspects of participants' linguistic background and experiences such as age of acquisition, proficiency, or frequency of language use and exposure are predictive of HSs' and L2 learners' performance on the discrimination task.

2. Methods

2.1. Participants

Three groups of participants took part in the study: 20 HSs of Korean, 20 English-speaking L2 learners of Korean, and 20 Korean L1-i speakers. All participants in the experimental groups were recruited online either via Prolific, an online participant recruitment platform for social science research, or using a snowball technique. Before performing the task, all participants completed a consent form and filled in a language background questionnaire. All participants were compensated for their participation and completed the modified version of Language Experience and Proficiency Questionnaire (LEAP-Q, Marian et al. 2007). The following section reports details about the language backgrounds of the participants. No participant reported any disability or difficulty in speaking, hearing, and vision, which could hinder participation in the AX discrimination task that involves responding to both visual and auditory stimuli.

The HS group consisted of 20 speakers of Korean born and raised in the United States and whose parents were first-generation immigrants and native speakers of Korean. The group included 11 females and 9 males, ranging between 19 and 42 years old ($M = 25.5$). At the moment of the experiment, they reported being exposed to Korean, as opposed to English, about 39% of the time, on average (although this number varied from 10% to 100%). In terms of choosing to speak Korean, when both Korean and English were available options, participants reported that they opted for Korean only 30% of the time, on average (again, with a wide range from 0% to 100%). HSs attributed their knowledge of Korean primarily to interactions with family members, watching TV, and reading in Korean. The majority of the participants have never resided in Korea, although five of them reported

stays between 1–3 years in duration, and one lived in Korea for 10 years. The reported age of onset of English acquisition in this group, on average, was at the age of 2 (ranging from 1 to 5). Their average self-reported English speaking and comprehension proficiency was 6.8 and 6.7, respectively (on a 7-point scale), with the vast majority indicating excellent proficiency. Their average self-reported Korean speaking, and comprehension proficiency was somewhat lower, at 4.7 and 5.1 (on a 7-point scale). In addition to self-reported proficiency estimates, we employed a more objective measure of Korean proficiency by evaluating participants’ verbal fluency in a narrative task via calculating the articulation rate. The average articulation rate for HSs was 4.2 syllables per second.

The L2 group consisted of 20 L2 learners (14 females and 6 males) with an age range of 18 to 42 years old ($M = 26.3$). All of them were native speakers of American English, learning Korean as an L2. All but three of these participants have not spent any time in Korea; the duration of stay for these three was from 1 to 2 years. Their age of onset of acquisition was around 21 years of age, qualifying them as adult L2 learners. The participants reported considerably more exposure to English (83%) than to Korean (17%). Most of the exposure came from watching TV and reading in Korean. Their average self-reported English speaking and comprehension proficiency was a perfect 7 on both counts, while their average Korean speaking and comprehension proficiency was at 2.7 and 3 (on a 7-point scale). Their average articulation rate in the narrative task was 3.4 syllables per second.

The L1-i group consisted of 20 native speakers residing in Korea (14 females and 6 males; age range, 20 and 34 years old ($M = 24.2$). All of them were born and raised in South Korea, and their residences at the time of the experiment were in the so-called Seoul-Gyeonggi area in which a standard variety of Korean is spoken. These participants were either students or alumni of a university located in Seoul. As English is ubiquitous in the South Korean educational system, all L1-i participants had some exposure to English. These participants’ self-reported start of English acquisition was higher than that of HSs, on average ($M = 7.5$) and on the individual basis (for the most part, the earliest reported AOA of 4-5 years of age in this group was the latest AOA for the heritage group). Their average self-reported English speaking and comprehension abilities were only 3 and 4, respectively (on a 7-point scale), while their self-reported Korean proficiency was at 7 on both counts. Their average articulation rate in the narrative task was 4.2 syllables per second. Table 1 reports the summary of the language backgrounds of the participants in all groups.

Table 1. Language backgrounds of L2 learners and HSs (means and SDs).

Question	HS	L2	L1-i
Age	25.5 (5.9)	26.3 (7.2)	24.2 (3.3)
AOA (HS, L1-i: English, L2: Korean)	2.1 (1.6)	22.3 (9.5)	7.5 (2.9)
Current exposure to Korean ^a	39% (20.4%)	17% (13%)	76.1% (27.7%)
Percentage of choosing to speak Korean over English ^b	30.4% (27.9%)	17.1% (19.2%)	69.1% (35.2%)
Korean speaking proficiency ^c	4.7 (1.6)	2.7 (1.8)	6.9 (0.3)
Korean comprehension	5.1 (1.3)	3 (1.8)	6.9 (0.3)
English speaking proficiency	6.8 (0.7)	7 (0)	3.3 (1.4)
English comprehension	6.7 (1.3)	7 (0)	4.2 (1.4)

^a: a 100% scale is used for exposure (0%: English only–100%: Korean only); ^b: a 100% scale is used for preferred language (0%: English only–100%: Korean only); ^c: Speaking and comprehension proficiency in both languages are represented on a 7-point Likert-scale (1: very poor–7: native-like).

2.2. Materials

2.2.1. Stimuli for the AX Discrimination Task

Seventy-two monosyllabic CV words, adapted from Schmidt (2007), were used as stimuli. Four female native speakers, three of whom were originally from Seoul and one from the Chung Cheong area, recorded the stimuli. These speakers had been in the US for no more than three years at the time of the recording and reported being exposed to English about 30% of the time or less, on average (see Schmidt 2007 for more detail). From the

original set of stimuli, which consisted of 684 CV items, 72 were selected (18 unique CV combinations recorded by four different speakers). These words all began with a lenis or aspirated stop consonant of three different places of articulation (/p/, /t/, /k/) and contained three different vowels (/a/, /i/, /u/), resulting in a total of 18 unique CV words.

To ensure the quality of the audio files, a Korean investigator in the current study evaluated and selected the stimuli. In addition, we performed acoustic analysis along with a visual inspection of the spectrograms of all stimuli, using Praat 6.1.10 (Boersma and Weenink 2021) to ensure that lenis stops and aspirated stops were differentiated by F0 and had long-lag VOT values. The results showed that the contrast between the two types of stops was implemented via both F0 ($t = 10.65$, $p < 0.001$) and VOT ($t = 5.62$, $p < 0.001$). However, although the VOT values were significantly different between lenis and aspirated stops, a considerable overlap in the VOT values between the two kinds of stops is observed, which is especially conspicuous when compared to F0 values as shown in Figure 1. This overlap in VOT values may contribute to the perceptual difficulty in its categorization. This analysis strongly suggests that the use of onset F0 is important for the discrimination of these stimuli because VOT is likely to be a somewhat unreliable cue on its own.

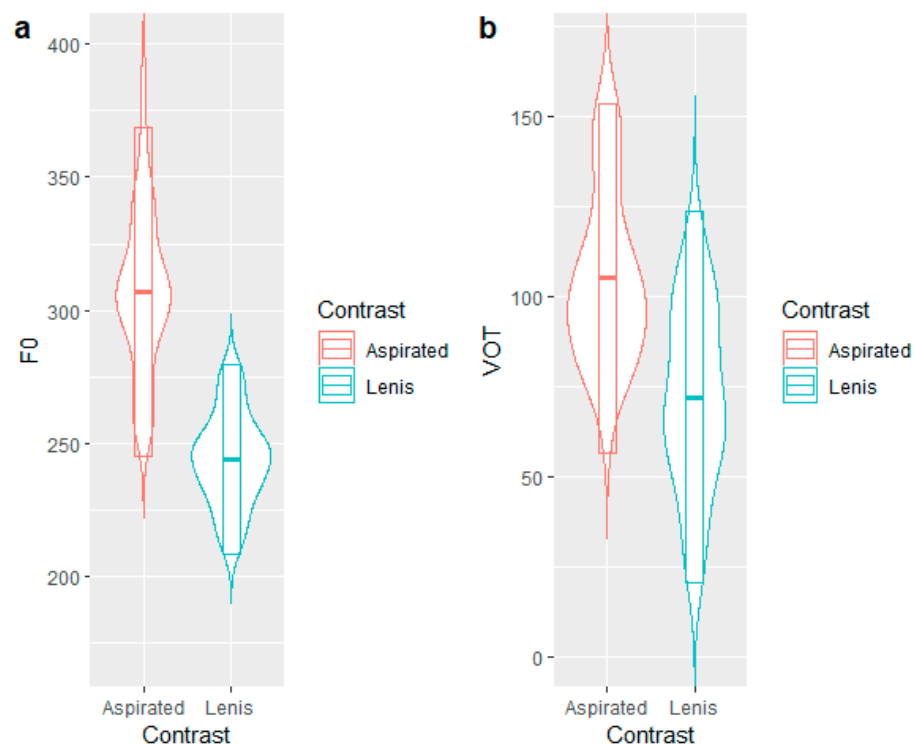


Figure 1. Violin plots of the lenis–aspirated stop contrast in the stimuli. (a) shows the comparison of F0, and (b) shows the comparison of VOT.

2.2.2. Verbal Narrative for Proficiency Estimation

A picture book of *Little Red Riding Hood* consisting of a sequence of wordless pictures (nine pictures total) was used for a picture description narrative task, which provided verbal narratives to be used for estimating participants' speaking proficiency in the Korean language. Narrative elicitation is a standard measure in the literature to examine grammatical growth (Cuza 2010; Montrul 2002; Rojas and Iglesias 2013; Sebastián and Slobin 1994). Participants took about five minutes on average to narrate the story.

2.3. Procedures

2.3.1. AX Discrimination Task

All groups performed an AX discrimination task which is a simple but widely used task for group comparison in terms of perceptual performance (Lee-Ellis 2012), using Gorilla interface (Anwyl-Irvine et al. 2020; an online research platform for behavioral scientists). Using the headphone check function in Gorilla, we ensured that all participants were wearing headphones or earbuds at the start of the experiment. During the AX task, participants listened to pairs of Korean monosyllabic words that potentially differed in the word-initial stop (lenis/aspirated) only and judged whether the two word-initial consonants were the same or different. On each trial, words A and X were played with an interval of 200 ms between the two. Right after the playback, two color-coded 'buttons' were displayed on the screen, a red button corresponding to 'different' and a blue button corresponding to 'the same' response. Participants registered their decision by clicking on one of the two buttons. Trials were separated by a period of 500 ms of a blank screen. Stimuli were randomized for each presentation. The same stimuli were used for all groups. An equal number of different and same pairs was used. The members of each pair were also presented an equal number of times in each order (AX and XA) and each unique pair was presented twice to each participant. The total number of trials amounted to 288 ((3 places of articulation (/p/, /t/, /k/) * 3 types of vowels (/a/, /i/, /u/) * 2 orders (AX and XA) * 2 repetitions * 4 different speakers) + an equal number of 'same' pairs).

Importantly, in the 'same' pairs the two words were not acoustically identical, but they did begin with the same type of consonant (in terms of place of articulation and laryngeal specification). In other words, for each 'same' pair, two different tokens of the same word recorded by the same speaker were used. This decision was made to render this admittedly very simple task somewhat more complex. As a result, participants had to determine that the two initial stops were phonologically the same, rather than simply acoustically identical, to make a 'same' decision. Additionally, phonetic information from the same-pair words could not help in making this decision since the two of the 'same' words were actually two different recordings of the same consonant-vowel combination.

Before the experimental task began, written instructions (in English for L2 learners and HSs and in Korean for L1-i speakers) were displayed on the screen and a short practice block of five trials was provided in order to familiarize participants with the task. This task took approximately 12 min for each participant to complete.

2.3.2. Picture Description Narrative Task

The purpose of the picture description task was to obtain verbal narratives which could be used to estimate participants' verbal proficiency via obtaining a measure of their speech fluency. All participants who had completed the AX discrimination task also performed a picture description narrative task on the same platform, Gorilla. Participants had an opportunity to take an optional five-minute break after completing the AX task and before moving on to the picture description task. After the break, written instructions were displayed on the screen, which requested that the recording be performed in a quiet place to minimize background noise in the acoustic signal. In this task, participants were asked to narrate a story of *Little Red Riding Hood* in Korean by describing a sequence of wordless pictures. Their production was recorded using the recording function in Gorilla, and participants could use any recording device of their preference for this task. The productions elicited from the narrative task were analyzed acoustically to measure the articulation rate of each participant. The articulation rate was calculated by dividing the number of syllables in the narrative (estimated as the number of vocalic nuclei) by the duration of the narrative (without pauses) to obtain the number of syllables per second rate (see Baker-Smemoe et al. 2014; De Jong 2018; Ginther et al. 2010; Kormos and Dénes 2004; Nagy and Brook 2020 for more discussion). Each recording was analyzed by deploying a script by De De Jong and Wempe (2009) in Praat 6.1.10.

2.4. Analyses

Participants' responses in the AX discrimination task were categorized as correct or incorrect and submitted to statistical analyses performed in RStudio 1.4.1103 (RStudio Team 2020) using the 'lme4' package (Bates et al. 2015).

Three mixed-effects logistic regression models were implemented. In all three, perceptual accuracy coded as '1' (correct answer) or '0' (incorrect answer) was used as a binary categorical dependent variable. The first model was used to analyze the effect of the participant group on discrimination performance. It included group (HS—reference level, L2, L1-i), trial type (same or different—reference level), their interaction, and speaker (the four female speakers who recorded the stimuli, Speaker 1 as reference level) as fixed effects. It also included item and subject as random intercepts. Independent variables were treatment coded.

The second and the third models were conducted on the data from the HS group and the L2 group, respectively, and were implemented in order to determine whether participants' background characteristics were predictive of their perceptual accuracy in the discrimination task. These models included Korean usage (participants' self-reported percentage of choosing to speak Korean, from 0%—English only to 100%—Korean only), Korean exposure (participants' self-reported percentage of current exposure to Korean, from 0%—English only to 100%—Korean only), age of L2 acquisition or AOA (English for HSs' group and Korean for L2 learners' group), and individual articulation rate as fixed factors; subject and item were entered as random effects. Due to the low quality of recorded audio files and other technical issues, six participants were excluded from this analysis, resulting in 18 HSs in the second model and 16 L2 learners in the third model.

3. Results

3.1. Effects of Group on Perceptual Accuracy

The results of the mixed-effects logistic regression model showed that the HS group was not significantly different from the L1-i group ($p = 0.23$) while being significantly different from the L2 group ($\beta = -2.01$, $SE = 0.16$, $p < 0.001$) in terms of the accuracy of their discrimination responses. A pairwise post hoc analysis using multiplicity adjustment, averaged over type and speaker, confirmed the result while also indicating that the estimated mean of the L1-i group and that of the L2 group were significantly different ($\beta = -1.78$, $SE = 0.16$, $p < 0.001$). As shown in Figure 2, the HS group was similar to the L1-i group in terms of the accuracy in the AX task, while both were considerably more accurate than the L2 group.

3.2. Effects of Speaker on Perceptual Accuracy

The results showed that Speakers 2 and 4, but not Speaker 3, were significantly different from Speaker 1 in their effect on the perceptual judgment; listening to these speakers increased the odds of accurate perceptual response by a factor of 2.66 and 2.50, respectively, when compared to Speaker 1 (Speaker 2: $\beta = 0.98$, $SE = 0.17$, $p < 0.001$, Speaker 3: $p = 0.06$, Speaker 4: $\beta = 0.91$, $SE = 0.16$, $p < 0.001$), adjusted for group and trial type. A pairwise post hoc comparison analysis was conducted using multiplicity adjustment (Tukey) to compare other pairs of speakers. The results revealed that Speaker 3 decreased perceptual accuracy compared to Speaker 2 ($\beta = -0.68$, $SE = 0.17$, $p < 0.001$) and compared to Speaker 4 ($\beta = -0.61$, $SE = 0.17$, $p = 0.001$). This indicates that when the lenis-aspirated stop contrast was represented with audio files recorded by Speakers 1 and 3, participants found it more difficult to discriminate the contrast compared to the recordings by other speakers. Figure 3 shows the results of the pairwise analysis of the effects of speaker on perceptual judgment. The estimated mean difference indicates the difference in the coefficients of each speaker pair in the mixed-effects logistic regression model. For example, Speaker 1 has negative values in the estimated mean difference in the comparison to all the other speakers (Speakers 2, 3, and 4), indicating that participants

performed worse in discriminating the lenis–aspirated stop contrast when evaluating the stimuli recorded by Speaker 1 compared to all other speakers.

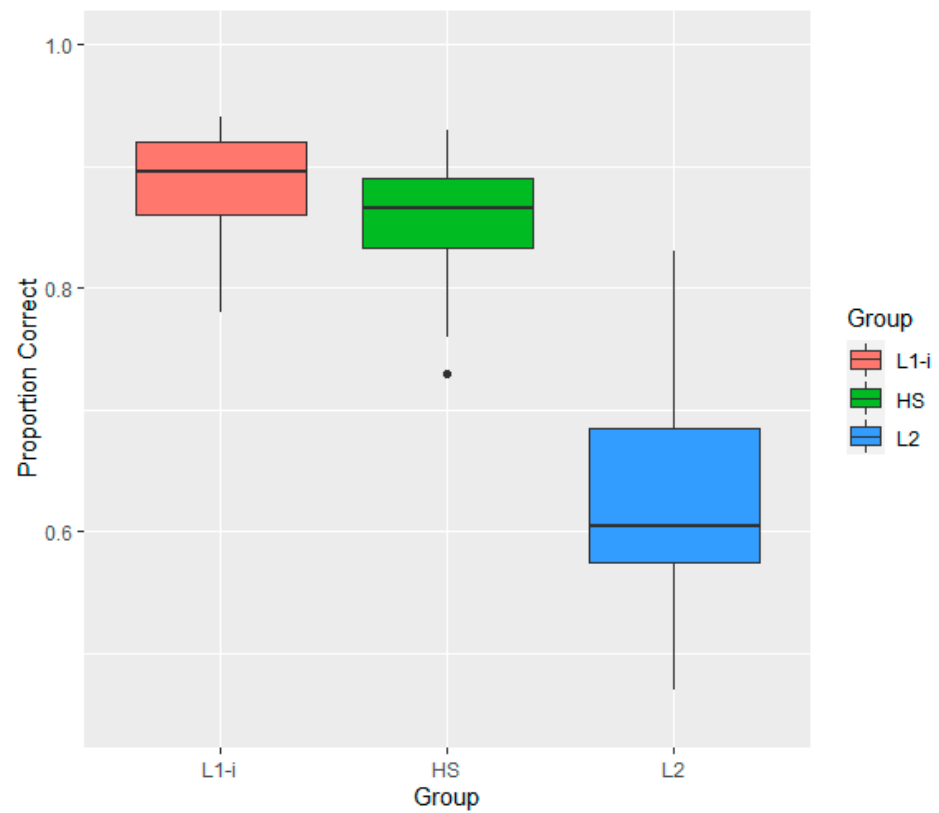


Figure 2. Boxplots showing the average perceptual accuracy by group in the AX discrimination task.

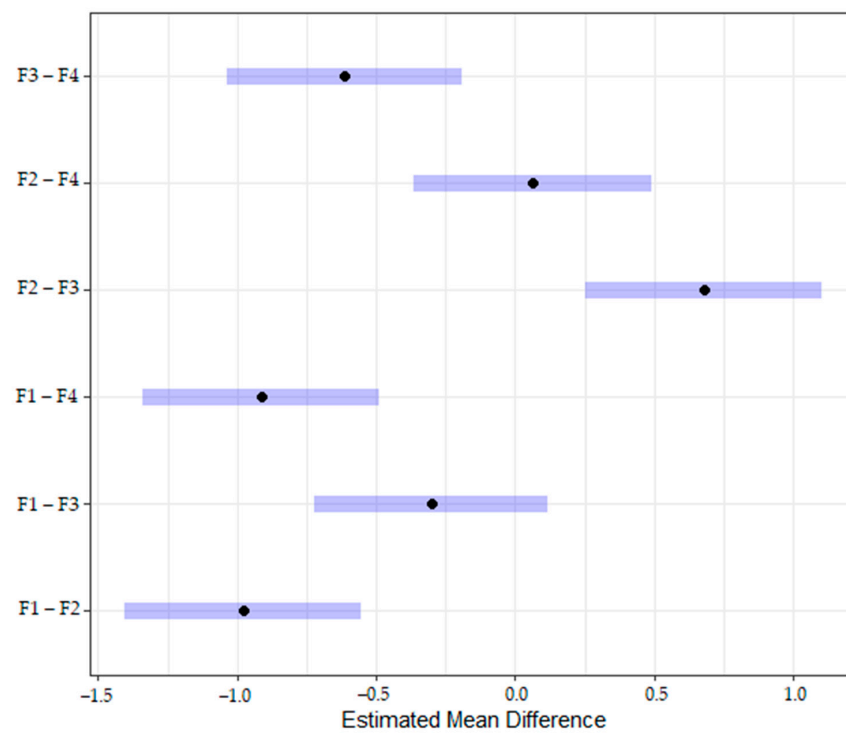


Figure 3. Estimated mean difference by each pair of speakers. F1–F4 refer to each female speaker who recorded the stimuli.

To investigate the source of this differences, we examined the VOT and F0 in all stimuli as correlates of lenis and aspirated stops for each speaker separately. A visual inspection of these values, plotted in Figure 4, reveals that Speaker 1 had the greatest amount of variability in VOT values of the lenis stops, while Speaker 3 had the least amount of difference in F0 values between the two kinds of stops. This suggests that the increased variability in VOT and the decreased distinction in F0 in the realization of these parameters as correlates of the lenis–aspirated contrast may have caused the perceptual difficulties experienced by the participants.

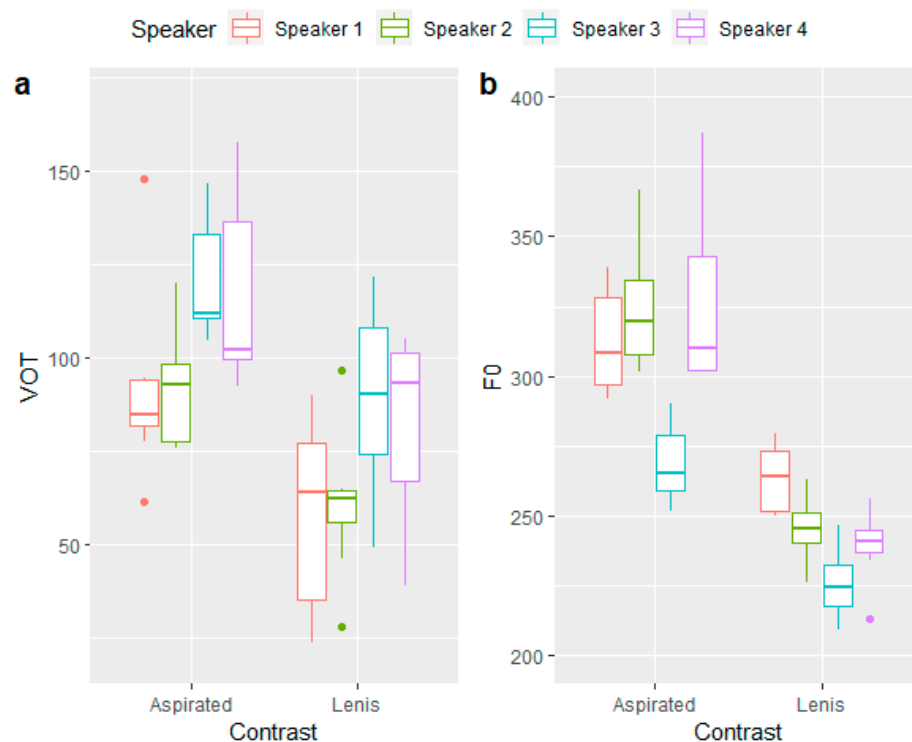


Figure 4. Box plots of the lenis–aspirated stop contrast separated by speaker. (a,b) show a comparison of VOT and F0, respectively.

3.3. Effects of Trial Type on Perceptual Accuracy

The effect of trial type (same trials/different trials) did not significantly affect perceptual judgment in the AX task ($p = 0.84$) when adjusted for group and speaker. However, the fixed effect of the interaction between group and type was found to significantly affect perceptual accuracy in the same trials for the L2 group compared to the HS group ($\beta = 1.10$, $SE = 0.10$, $p < 0.001$). On the other hand, the HS group's response did not differ by the trial type when compared with that of the L1-i group ($p = 0.05$). In order to examine the main effect of the interaction factor between group and type, a pairwise post-hoc analysis using Tukey adjustment was conducted. Figure 5 shows the linear prediction of the interaction effect between Group and Type. In the different trial type, the HS group was not significantly different from the L1-i group ($p = 0.47$). In contrast, both the HS group and the L1-i group performed significantly better than the L2 group in discriminating the stop contrasts in different trials. The odds ratios of accurate discrimination increased by a factor of 7.46 and 9.08 for these groups compared to the L2 group, respectively (HS–L2: $\beta = 2.01$, $SE = 0.16$, $p < 0.001$, L1-i–L2: $\beta = 2.21$, $SE = 0.16$, $p < 0.001$).

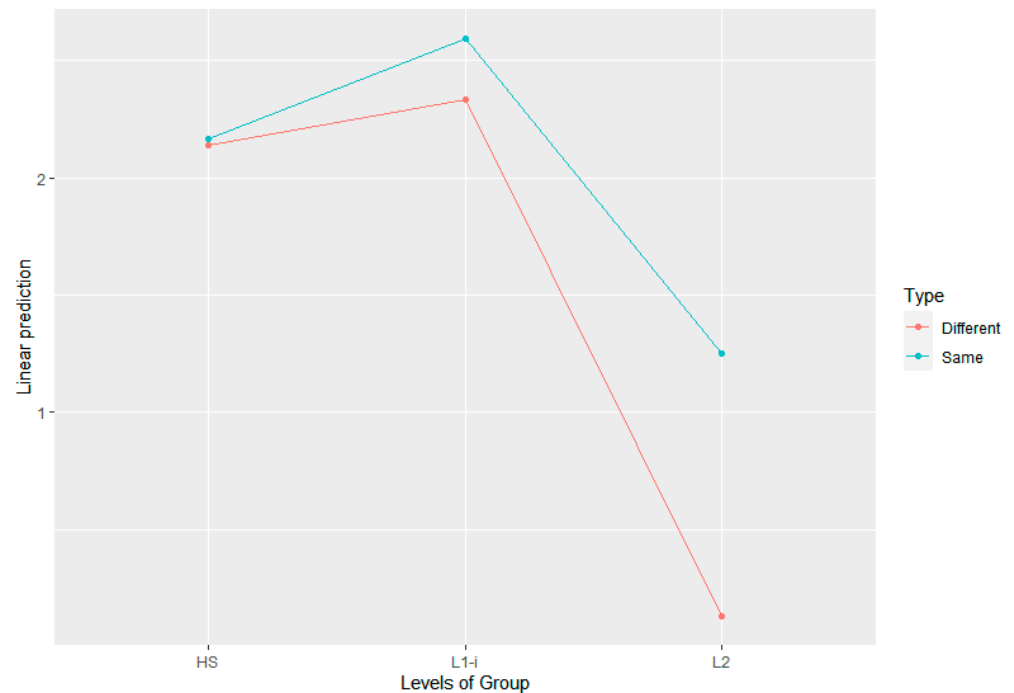


Figure 5. Linear prediction of the interaction effect between group and type. Y-axis represents the estimated means (coefficients) for each group in the pairwise post-hoc analysis of the mixed-effects logistic regression model.

In the same trial type, the HS group was significantly different from the L1-i group. The odds ratio of accurate judgment decreased by 35% for the HS compared to the L1-i group ($\beta = -0.43$, $SE = 0.17$, $p = 0.03$). Both the HS and L1-i groups significantly outperformed the L2 group on same trials. The odds ratios of accurate judgement were increased by a factor of 2.49 and 3.83 for these two groups, respectively (HS–L2: $\beta = 0.91$, $SE = 0.16$, $p < 0.001$, L1-i–L2: $\beta = 1.34$, $SE = 0.17$, $p < 0.001$).

These results indicate that while the HS group and the L1-i group performed comparably in the different trials, the L1-i group performed slightly more accurately in the same trials. On the other hand, the L2 group showed significantly lower accuracy in both same- and different-word than the HS and L1-i groups. This indicates that detecting a lenis–aspirated distinction was especially challenging for the L2 participants.

Figure 6 shows the average accuracy scores of each group by trial. In addition to the statistical analysis, there are three descriptive aspects of the result that are worth mentioning. First, the average perceptual accuracy was only slightly higher for the same trials than in the different trials in the L1-i and HS groups, while the difference was quite pronounced in the L2 group. Second, the L2 group showed the greatest difference between the same and different trials. Third, even L1-i speakers did not reach 100% accuracy in the AX task. This confirms that the lenis–aspirated stop contrast is perceptually difficult not only for L2 learners and HSs but for L1-i speakers as well. Table 2 summarizes the results of the fixed effects in the logistic mixed-effects model.

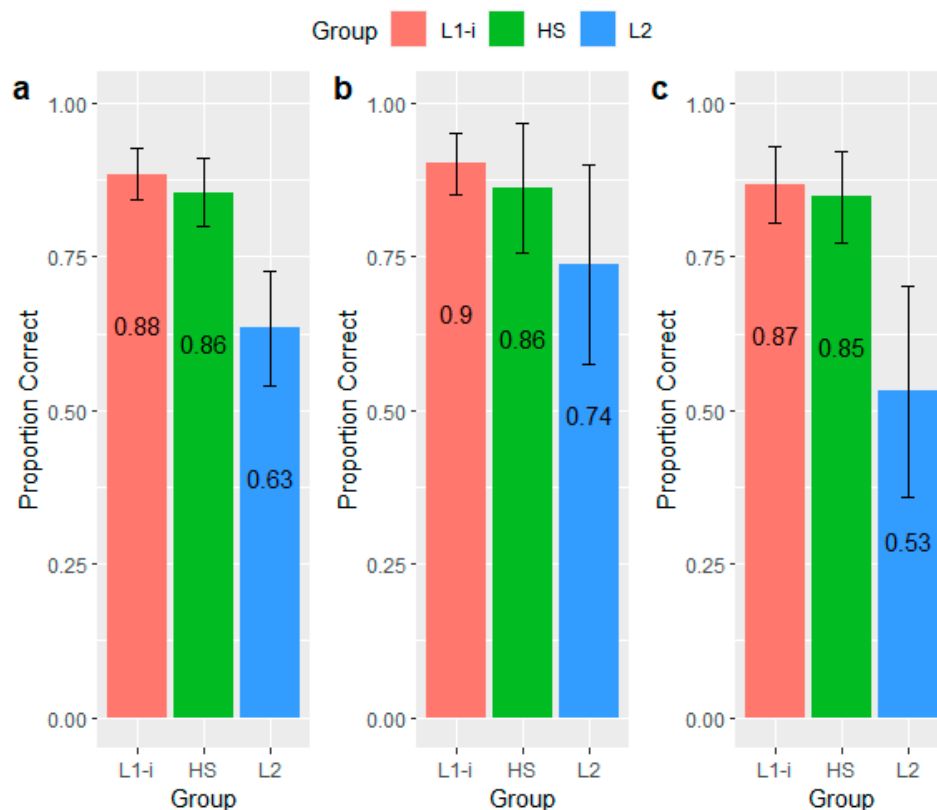


Figure 6. Bar plots of the average accuracy scores of each group by trial. (a) shows the average scores including both trial types, (b) shows the average scores in the same-word trials, and (c) shows the average scores in the different-word trials.

Table 2. Results of the fixed effects of the first mixed-effects model.

	Estimate	SE	z-Value	p-Value	Odds Ratio
Intercept	1.59	0.17	9.23	<0.001	4.90
Group L1-i	0.20	0.17	1.18	0.24	1.22
Group L2	−2.01	0.16	−12.33	<0.001	0.13
TypeSame	0.03	0.14	0.20	0.84	1.03
Speaker 2	0.98	0.17	5.92	<0.001	2.66
Speaker 3	0.30	0.16	1.85	0.06	1.35
Speaker 4	0.91	0.16	5.55	<0.001	2.50
Group L1-i:TypeSame	0.23	0.12	1.95	0.05	1.26
Group L2:TypeSame	1.10	0.10	10.56	<0.001	2.99

3.4. Effects of Articulation Rate on Perceptual Accuracy

The second mixed-effects model in the study analyzed the effects of background factors on the perceptual judgment made by HSs. The results showed that articulation rate significantly affected the perceptual accuracy of HSs by increasing the odds of accurate perceptual response by a factor of 1.83 as the articulation rate increased by 1 syllable/sec ($\beta = 0.60, SE = 0.21, p < 0.01$), adjusted for other factors. The result suggests that HSs who were more fluent in Korean (and likely, more proficient overall) were also more accurate in discriminating the lenis–aspirated stop contrast than those who were less fluent. In contrast, all other factors included in the model—AOA of English, percent Korean use, and percent Korean exposure—showed no significant relation to the perceptual accuracy in the current dataset.

The third model analyzed the L2 group and predicted the effects of the background factors on the perceptual judgment in the AX task. None of the factors in the model were significantly related to the perceptual accuracy of L2 speakers.

4. Discussion

The present study aimed to determine whether HSs of Korean in the United States find it relatively more challenging to distinguish Korean lenis stops from Korean aspirated ones in an AX discrimination task compared to the L1-i speakers in Korea. We also collected data from L2 learners of Korean with English as their native language in order to determine whether a similar disadvantage would be observed for this group but to a greater degree than for HSs. These hypotheses were motivated by several assumptions, including the fact that the lenis–aspirated contrast is proven to be a perceptually challenging one, possibly due to its extensive (and largely exclusive) reliance on onset F0 as a perceptual cue. Another underlying assumption is that English and Korean laryngeal categories can influence each other, both in production and perception of bilingual speakers. One anticipated outcome of this influence is less than optimal reliance on onset F0 in distinguishing lenis and aspirated stops by Korean HSs, due to the effect of English in which VOT serves as the dominant feature distinguishing laryngeal categories. As a result, lenis–aspirated discrimination accuracy was predicted to suffer in HSs, as well as L2 learners (for the same reason).

The results only partially supported our predictions. L2 learners of Korean did in fact demonstrate a significantly lower accuracy of lenis–aspirated discrimination compared to HSs. However, there was no difference between the groups of L1-i speakers and HSs, apart from a statistically significant but minor advantage of the L1-i group in some trials only. This result gave us no strong evidence to conclude that HSs' discrimination abilities were not comparable to those of L1-i speakers.

This result is not particularly surprising when considered in the broader context of literature on the perceptual abilities of HSs. Although some studies indicate that HSs may underperform on some aspects of native speech perception, compared to L1-i speakers (Ahn et al. 2017; Cheon and Lee 2013; Lee-Ellis 2012), there is also evidence that equivalent performance can be observed (Chang 2016; Lukyanenko and Gor 2011; Oh et al. 2003; Tees and Werker 1984; Werker 1989). For example, Oh et al. (2003) showed that HSs of Korean (childhood speakers in their terminology) were as accurate as L1-i speakers at recognizing fortis, lenis, and aspirated Korean stops in a three-choice identification task.

Furthermore, the difficulty of the task appears to play a role in determining the outcome of such comparisons, with simpler tasks, such as AX discrimination, often eliciting more comparable performance across groups (Lee-Ellis 2012). Moreover, some scholars suggest that among all types of linguistic competence, perceptual abilities benefit the most from the early and authentic exposure that characterizes heritage language acquisition, resulting in the most native-like performance in this modality, even when compared to speech production, and especially in counterposition with morphosyntactic abilities (Chang 2021; Oh et al. 2003). In addition, developmental investigations report that lenis–aspirated distinction in Korean stops is acquired before school age, when dominant language starts playing an important role in language development of HSs (Choi et al. 2019; Jun 2007; Kim and Stoel-Gammon 2009). Therefore, perceptual abilities with respect to this aspect of Korean phonology are likely to be well-established before dominant language begins exerting its influence on the heritage language. Thus, one possible interpretation of the lack of significant differences between heritage and L1-i speakers observed in the present study is that the two groups are truly equivalent in the way they perceive and process the lenis–aspirated contrast in Korean, especially under the conditions of this relatively simple task (AX discrimination).

There are, however, alternative possibilities. In particular, it is possible that HSs did rely on onset F0 less and on VOT more than L1-i speakers in their perception of the lenis–aspirated contrast as we predicted, but that it did not, contrary to prediction, affect their discrimination performance adversely. In previous work on L2 acquisition, parallels were often observed between incorrect cue weighting in perception of non-native contrasts and sub-optimal performance in the discrimination of the contrasts (Yamada and Tohkura 1990, 1992). However, there are also findings indicating that non-native cue-weighting does not inevitably lead to differences in contrast perception. For example,

Escudero (2000, 2001) showed that Spanish L2 speakers of English performed in a native-like fashion in discriminating the English lax-tense vowel contrast, even though their perceptual cue-weighting was different from that of English native speakers (they relied more on the temporal than spectral dimension). Therefore, our heritage participants could perform on par with L1-i speakers on the discrimination task, in spite of their reliance on a more English-like cue (VOT). This explanation is especially viable in light of the fact that stimuli used in the experiment did contain both VOT and onset F0 as cues to the lenis–aspirated distinction.

The acoustic analysis of the stimuli used in the AX discrimination task indicates that the contrast between lenis and aspirated stops in this dataset was implemented via both dimensions, VOT and onset F0. That is, the expected VOT merger between lenis and aspirated stops was not as pronounced in these stimuli as expected, potentially giving our heritage listeners (with their purported reliance on VOT) a leg up. It is possible that speakers who recorded the stimuli introduced some degree of hyperarticulation along the VOT dimension in their productions in order to increase the intelligibility of their speech, which is an effect sometimes observed for speech recorded in highly artificial laboratory conditions (see, e.g., Chang and Mandock 2019, for similar reasoning). In fact, there is evidence that Korean stops produced using ‘clear speech’ are differentiated via both VOT and onset F0 even for speakers who are expected to exhibit a VOT merger (Kang and Guion 2008). There is also evidence that VOT differences between lenis and aspirated stops are enhanced in child-directed speech, which is often characterized by modifications similar to those of clear speech (Ko 2018).

In addition, these speakers could have produced stronger VOT contrast between lenis and aspirated stops because they were all proficient in English and were, in fact, recorded in the United States, where they lived at the time of the recording (which is not mutually exclusive with the ‘clear speech’ hypothesis). As previous research shows, this fact alone could lead to a greater reliance on VOT in implementing Korean laryngeal contrasts by these speakers (Cheng 2017; Kang and Nagy 2016).

Finally, VOT merger in Korean lenis and aspirated stops is due to an ongoing sound change leading to tonogenesis. Some amount of variability in adopting the sound change is expected among the contemporary speakers. Therefore, some of speakers who recorded the stimuli for the experiment could be not as advanced in adopting this sound change as others (e.g., Speaker 1), relying on VOT more than they would otherwise.

Thus, the acoustic characteristics of the stimuli made alternative perceptual strategies available to the participants, specifically the use of VOT, instead of or in addition to onset F0, in determining the difference between lenis and aspirated stops, which could put the group relying mostly on F0 and the one relying mostly on VOT on equal footing.

The data we have at hand do not allow us to confirm or reject either of these interpretations at the moment. However, the two interpretations make different, clear, and testable predictions, which can be addressed in future work. For example, in a study with a comparable design which uses lenis and aspirated stimuli that differ only in F0, not in VOT (i.e., fully VOT-‘merged’ versions of lenis and aspirated stops), we would expect to find a significant difference in discrimination accuracy between HSs and L1-i speakers, if the second interpretation is correct.

Although we did not see a significant difference in the discrimination of the contrast between heritage and L1-i speakers, the heritage participants’ fluency of spoken Korean, measured as the articulation rate, was significantly predictive of their perceptual accuracy. Insofar as fluency measures can be used as indicative of the overall language proficiency (see Nagy and Brook 2020; Polinsky 2008, 2011; Polinsky and Kagan 2007), these results suggest that greater proficiency in Korean led to greater perceptual accuracy. Indirectly, this finding supports our assumption that crosslinguistic influence from English is the source of perceptual difficulties. Research indicates that bilinguals with more balanced proficiency and dominance in the two languages, in particular early simultaneous bilinguals, often demonstrate a greater ability to maintain a separation between the two sound systems,

minimizing the interference effects between the two (e.g., [Barlow et al. 2013](#); [Guion 2003](#); [MacLeod et al. 2009](#); [Sundara and Baum 2006](#)). Therefore, those of our HSs who exhibited greater proficiency in their non-dominant language could exercise better control over the interference from English, thus performing better on the perceptual task. Interestingly, for L2 learners none of the linguistic background and experience factors were found to be significantly related to perceptual accuracy.

Among other notable aspects of these results, we observed a less than perfect performance on a relatively simple AX discrimination task even among the L1-i listeners, who performed with an accuracy of 88%, on average. This outcome can partly be attributed to the overall perceptual difficulty of the lenis–aspirated distinction in the aftermath of tonogenesis, if onset F0 is a less salient perceptual cue than VOT ([Kong and Lee 2018](#); [Son 2020](#)). It nevertheless stands in contrast to the results of similar tasks, e.g., a three-choice phonemic identification of the lenis–fortis–aspirated stops in [Oh et al. 2003](#), where native monolingual speakers of Korean performed with an accuracy of 98.6%. We believe that administering the study online vs. in the laboratory may have resulted in this discrepancy. Since we did not have control over the conditions in which our participants performed the tasks, beyond ensuring that they were using the headphones, it is possible that some participated while in a noisy or distracting environment, thus failing to perform at the same level at which they would have performed in the optimal conditions of a phonetics laboratory. While this complicates comparisons with laboratory studies, these conditions are closer to those under which natural speech perception takes place, thus providing a somewhat more realistic estimate of the relevant perceptual behaviors.

It is also noteworthy that L2 learners in the present study achieved a fairly low degree of accuracy on ‘different’ trials—only 53% correct. Their moderate proficiency in Korean is undoubtedly partly responsible for this outcome (around 3 on a 7-point scale, by self-report). Nevertheless, one may question why these learners did not take advantage of the VOT as a correlate of the lenis–aspirated distinction the way HSs presumably did. This may be especially puzzling given that positive VOT is a primary cue to the voicing distinction in word–initial stops in English ([Abramson and Lisker 1985](#)). It should be considered, however, that both lenis and aspirated stops in Korean have long lag VOT and therefore fall squarely within the range of voiceless English stops. As a result, English-speaking learners are required to make a within-category discrimination decision when attempting to use VOT as a cue—in effect trying to distinguish between aspirated and slightly more aspirated stops. It should come as no surprise that it proves a difficult task for learners whose neural pathways have been trained to discriminate categorically between aspirated and unaspirated stops. In fact, [Cheon and Lee \(2013\)](#) and [Schmidt \(2007\)](#) showed that native speakers of American English, including those learning Korean as a second language, strongly associated both lenis and aspirated Korean stops with English voiceless stops. This difficulty has probably been compounded by the fact that VOT was not a very consistent correlate of this distinction in Korean. [Figure 4](#) demonstrates a fair amount of variability as well as partial overlap or close proximity between lenis and aspirated VOT ranges.

To conclude, in this study, we hypothesized, based on the assumptions of SLM, that a crosslinguistic interaction takes place between the laryngeal categories of English and Korean for those who speak both languages, specifically HSs of Korean in the United States and American learners of Korean as an L2. More precisely, we expected that, as a result of English affecting the perception of Korean, for both groups of listeners, the accuracy of their discrimination of the Korean lenis–aspirated contrast would suffer, in comparison to the group of L1-i speakers. We also expected a difference between the two groups, such that L2 learners, for whom Korean is a later acquired and non-dominant language, would discriminate Korean categories more poorly than HSs, for whom Korean is also non-dominant but an early acquired language (in addition to other differences between the groups, e.g., in terms of current exposure and use of Korean). Our prediction was partially supported, demonstrating that L2 learners of Korean were outperformed by HSs of Korean,

but the latter group was not significantly different from the L1-i speakers. While on the face of it this result adds to the body of literature arguing for equal perceptual abilities of HSs and L1-i/monolingual speakers, an alternative explanation is possible. We conjectured that the lack of differences in the performance of heritage and L1-i speakers could be explicable by the acoustic properties of the stimuli used in the present study, which could allow both groups to achieve high discrimination accuracy despite relying on distinct perceptual strategies. This is an intriguing possibility that must await future research for its definitive confirmation.

Author Contributions: Conceptualization, Y.S.; Methodology, Y.S., O.D. and A.C.; Software, Y.S.; Validation, Y.S., O.D. and A.C.; Formal Analysis, Y.S., O.D. and A.C.; Investigation, Y.S.; Resources, Y.S., O.D. and A.C.; Data Curation, Y.S.; Writing—Original Draft Preparation, Y.S.; Writing—Review & Editing, Y.S., O.D. and A.C.; Visualization, Y.S.; Supervision, O.D. and A.C.; Project Administration, Y.S.; Funding Acquisition, O.D. and A.C. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: The study was conducted according to the guidelines of the Declaration of Helsinki, and approved by the Institutional Review Board of Purdue University (IRB-2021-259 and 3 May 2021).

Informed Consent Statement: Informed consent was obtained online from all subjects involved in the study.

Data Availability Statement: The data presented in this study are available on request from the corresponding author. The data are not publicly available due to confidentiality.

Acknowledgments: We would like to thank Anna Schmidt for providing the stimuli which were used in the AX task of the present study. We are also grateful to April Ginther for sharing her expertise on fluency in second language speech.

Conflicts of Interest: The authors declare no conflict of interest.

References

- Abramson, Arthur S., and Leigh Lisker. 1985. Relative power of cues: f0 shift versus voice timing. *Phonetic Linguistics: Essays in Honor of Peter Ladefoged* 25–33.
- Ahn, Sunyoung, Charles B. Chang, Robert DeKeyser, and Sunyoung Lee-Ellis. 2017. Age effects in first language attrition: Speech perception by Korean-English bilinguals. *Language Learning* 67: 694–733. [CrossRef]
- Antoniou, Mark, Michael D. Tyler, and Catherine T. Best. 2012. Two ways to listen: Do L2-dominant bilinguals perceive stop voicing according to language mode? *Journal of Phonetics* 40: 582–94. [CrossRef]
- Anwyl-Irvine, Alexander L., Jessica Massonnié, Adam Flitton, Natasha Kirkham, and Jo K. Evershed. 2020. Gorilla in our midst: An online behavioral experiment builder. *Behavior Research Methods* 52: 388–407. [CrossRef]
- Baker, Wendy, and Pavel Trofimovich. 2005. Interaction of native-and second-language vowel system (s) in early and late bilinguals. *Language and Speech* 48: 1–27. [CrossRef] [PubMed]
- Baker-Smemoe, Wendy, Dan P. Dewey, Jennifer Bown, and Rob A. Martinsen. 2014. Does measuring L2 utterance fluency equal measuring overall L2 proficiency? Evidence from five languages. *Foreign Language Annals* 47: 707–28. [CrossRef]
- Bang, Hye-Young, Morgan Sonderegger, Yoonjung Kang, Meghan Clayards, and Tae-Jin Yoon. 2018. The emergence, progress, and impact of sound change in progress in Seoul Korean: Implications for mechanisms of tonogenesis. *Journal of Phonetics* 66: 120–44. [CrossRef]
- Barlow, Jessica A., Paige E. Branson, and Ignatius S. B. Nip. 2013. Phonetic equivalence in the acquisition of /l/ by Spanish-English bilingual children. *Bilingualism: Language and Cognition* 16: 68–85. [CrossRef]
- Bates, Douglas, Martin Mächler, Ben Bolker, and Steve Walker. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67: 1–48. [CrossRef]
- Bergmann, Christopher, Amber Nota, Simone A. Sprenger, and Monika S. Schmid. 2016. L2 Immersion Causes Non-Native-like L1 Pronunciation in German Attriters. *Journal of Phonetics* 58: 71–86. [CrossRef]
- Bernthal, John E., Nicholas W. Bankson, and Peter Flipsen Jr. 2013. *Articulation and Phonological Disorders: Speech Sound Disorders in Children*. Boston: Pearson.
- Boersma, Paul, and David Weenink. 2021. Praat: Doing Phonetics by Computer [Computer Program]. Version 6.1.40. Available online: <http://www.praat.org/> (accessed on 27 February 2021).

- Caramazza, Alfonso, Grace H. Yeni-Komshian, Edgar B. Zurif, and Ettore Carbone. 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: 421–28. [CrossRef]
- Cebrian, Juli. 2006. Experience and the use of non-native duration in L2 vowel categorization. *Journal of Phonetics* 34: 372–87. [CrossRef]
- Chang, Charles B. 2012. Rapid and multifaceted effects of second-language learning on first-language speech production. *Journal of Phonetics* 40: 249–68. [CrossRef]
- Chang, Charles B. 2016. Bilingual perceptual benefits of experience with a heritage language. *Bilingualism: Language and Cognition* 19: 791–809. [CrossRef]
- Chang, Charles B. 2021. Phonetics and phonology of heritage languages. In *The Cambridge Handbook of Heritage Languages and Linguistics*. Edited by Sivina Montrul and Maria Polinsky. Cambridge: Cambridge University Press, pp. 581–612.
- Chang, Seung-Eun, and Karina Mandock. 2019. A phonetic study of Korean heritage learners' production of Korean word-initial stops. *Heritage Language Journal* 16: 273–95. [CrossRef]
- Cheng, Andrew. 2017. VOT merger and f0 contrast in Heritage Korean in California. *UC Berkeley PhonLab Annual Report* 13: 281–311. [CrossRef]
- Cheon, Sang Yee, and Teresa Lee. 2013. The perception of Korean stops by heritage and non-heritage learners: Pedagogical implications for beginning learners. *Korean Language in America* 18: 23–39. Available online: <http://www.jstor.org/stable/42922375> (accessed on 1 August 2020).
- Cho, Taehong, Sun-Ah Jun, and Peter Ladefoged. 2002. Acoustic and aerodynamic correlates of Korean stops and fricatives. *Journal of Phonetics* 30: 193–228. [CrossRef]
- Choi, Youngon, Minji Nam, Reiko Mazuka, Hyun-Kyung Hwang, Minha Shin, Jisoo Kim, Sunho Jung, Jimin Beom, and Naoto Yamane. 2019. Development of speech perception in Korean infants: Discriminating unusual sound contrasts with diachronic change. Paper presented at the 2nd Hanyang International Symposium on Phonetics and Cognitive Sciences of Language, Seoul, Korea, May 24–25; pp. 35–36.
- Cuza, Alejandro. 2010. The L2 acquisition of aspectual properties in Spanish. *Canadian Journal of Linguistics* 55: 1001–28.
- De Jong, Nivja H. 2018. Fluency in second language testing: Insights from different disciplines. *Language Assessment Quarterly* 15: 237–54. [CrossRef]
- De Jong, Nivja H., and Ton Wempe. 2009. Praat script to detect syllable nuclei and measure speech rate automatically. *Behavior Research Methods* 41: 385–90. [CrossRef]
- Dmitrieva, Olga. 2019. Transferring perceptual cue-weighting from second language into first language: Cues to voicing in Russian speakers of English. *Journal of Phonetics* 73: 128–43. [CrossRef]
- Dmitrieva, Olga, Allard Jongman, and Joan Sereno. 2010. Phonological Neutralization by Native and Non-Native Speakers: The Case of Russian Final Devoicing. *Journal of Phonetics* 38: 483–92. [CrossRef]
- Dmitrieva, Olga, Fernando Llanos, Amanda A. Shultz, and Alexander L. Francis. 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. [CrossRef]
- Escudero, Paola. 2000. Developmental Patterns in the Adult L2 Acquisition of New Contrasts: The Acoustic Cue Weighting in the Perception of Scottish Tense/Lax Vowels by Spanish Speakers. Ph.D. dissertation, University of Edinburgh, Edinburgh, UK.
- Escudero, Paola. 2001. The role of the input in the development of L1 and L2 sound contrasts: Language-specific cue weighting for vowels. In *Proceedings of the 25th Annual Boston University Conference on Language Development*. Somerville: Cascadia Press, vol. 1.
- Flege, James Emil. 1987. The production of “new” and “similar” phones in a foreign language: Evidence for the effect of equivalence classification. *Journal of Phonetics* 15: 47–65. [CrossRef]
- Flege, James E. 1995. Second Language Speech Learning Theory, Findings, and Problems. In *Speech Perception and Linguistic Experience: Issues in Cross-Language Research*. Edited by Winifred Strange. Timonium: York Press, pp. 233–77.
- Flege, James E. 2003. Assessing Constraints on Second-Language Segmental Production and Perception James Emil Flege. In *Phonetics and Phonology in Language Comprehension and Production: Differences and Similarities*. Edited by Antje S. Meyer and Niels O. Schiller. Berlin: Mouton de Gruyter, pp. 319–58.
- Flege, James E., and Ocke-Schwen Bohn. 2020. The revised speech learning model (SLM-r). In *Second Language Speech Learning: Theoretical and Empirical Progress*. Cambridge: Cambridge University Press, pp. 3–83.
- Flege, James E., and Wieke Eefting. 1987a. Production and perception of English stops by native Spanish speakers. *Journal of Phonetics* 15: 67–83. [CrossRef]
- Flege, James E., and Wieke Eefting. 1987b. Cross-language switching in stop consonant perception and production by Dutch speakers of English. *Speech Communication* 6: 185–202. [CrossRef]
- Fowler, Carol A., Valery Sramko, David J. Ostry, Sarah A. Rowland, and Pierre Hallé. 2008. Cross Language Phonetic Influences on the Speech of French-English Bilinguals. *Journal of Phonetics* 36: 649–63. [CrossRef]
- Garcia-Sierra, Adrian, Randy L. Diehl, and Craig Champlin. 2009. Testing the double phonemic boundary in bilinguals. *Speech Communication* 51: 369–78. [CrossRef]
- Ginther, April, Slobodanka Dimova, and Rui Yang. 2010. Conceptual and empirical relationships between temporal measures of fluency and oral English proficiency with implications for automated scoring. *Language Testing* 27: 379–99. [CrossRef]
- Guion, Susan G. 2003. The vowel systems of Quichua-Spanish bilinguals. *Phonetica* 60: 98–128. [CrossRef]

- Han, Mieko S., and Raymond S. Weitzman. 1970. Acoustic features of Korean /P, T, K/, /p, t, k/ and /ph, th, kh. *Phonetica* 22: 112–28. [[CrossRef](#)]
- Hazan, Valerie L., and Georges Boulakia. 1993. Perception and production of a voicing contrast by French-English bilinguals. *Language and Speech* 36: 17–38. [[CrossRef](#)]
- Harada, Tetsuo. 2003. L2 influence on L1 speech in the production of VOT. Paper presented at the 15th International Congress of Phonetic Sciences, Barcelona, Spain, August 3–9; Edited by Maria-Josep Solé, Daniel Recasens and Joaquín Romero. Barcelona: Causal Productions, pp. 1085–88.
- House, Arthur S., and Grant Fairbanks. 1953. The influence of consonant environment upon the secondary acoustical characteristics of vowels. *The Journal of the Acoustical Society of America* 25: 105–13. [[CrossRef](#)]
- Idemaru, Kaori, Lori L. Holt, and Howard Seltman. 2012. Individual differences in cue weights are stable across time: The case of Japanese stop lengths. *The Journal of the Acoustical Society of America* 132: 3950–64. [[CrossRef](#)]
- Jun, Sun-Ah. 2007. Phonological development of Korean: A case study. *UCLA Working Papers in Phonetics* 105: 51–65.
- Kagaya, Ryohei. 1974. A fiberoptic and acoustic study of the Korean stops, affricates and fricatives. *Journal of Phonetics* 2: 161–80. [[CrossRef](#)]
- Kang, Kyoung-Ho, and Susan G. Guion. 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: 1672–83. [[CrossRef](#)] [[PubMed](#)]
- Kang, Kyoung-Ho, and Susan G. Guion. 2008. Clear speech production of Korean stops: Changing phonetic targets and enhancement strategies. *The Journal of the Acoustical Society of America* 124: 3909–17. [[CrossRef](#)] [[PubMed](#)]
- Kang, Okim, Don Rubin, and Lucy Pickering. 2010. Suprasegmental measures of accentedness and judgments of language learner proficiency in oral English. *The Modern Language Journal* 94: 554–66. [[CrossRef](#)]
- Kang, Yoonjung. 2014. Voice onset time merger and development of tonal contrast in Seoul Korean stops: A corpus study. *Journal of Phonetics* 45: 76–90. [[CrossRef](#)]
- Kang, Yoonjung, and Naomi Nagy. 2016. VOT merger in Heritage Korean in Toronto. *Language Variation and Change* 28: 249–72. [[CrossRef](#)]
- Kim, Mi-Ran Cho. 1994. Acoustic Characteristics of Korean Stops and Perception of English Stop Consonants. Ph.D. dissertation, University of Wisconsin, Madison, WI, USA.
- Kim, Mi-Ryoung. 2012. L1-L2 transfer in VOT and f0 production by Korean English learners: L1 sound change and L2 stop production. *Phonetics and Speech Sciences* 4: 31–41. [[CrossRef](#)]
- Kim, Mi-Ryoung, Patrice Speeter Beddor, and Julie Horrocks. 2002. The contribution of consonantal and vocalic information to the perception of Korean initial stops. *Journal of Phonetics* 30: 77–100. [[CrossRef](#)]
- Kim, Minjung, and Carol Stoel-Gammon. 2009. The acquisition of Korean word-initial stops. *The Journal of the Acoustical Society of America* 125: 3950–61. [[CrossRef](#)]
- Ko, Eon-suk. 2018. Mothers would rather speak clearly than spread innovations: The case of Korean VOT. Paper presented at the 1st Hanyang International Symposium on Phonetics and Cognitive Sciences of Languages, Seoul, Korea, May 18–19; pp. 31–32.
- Kong, Eun Jong. 2012. Perception of Korean stops with a three-way laryngeal contrast. *Phonetics and Speech Sciences* 4: 13–20. [[CrossRef](#)]
- Kong, Eun Jong, and Hyunjung Lee. 2018. Attentional modulation and individual differences in explaining the changing role of fundamental frequency in Korean laryngeal stop perception. *Language and Speech* 61: 384–408. [[CrossRef](#)] [[PubMed](#)]
- Kong, Eun Jong, and In Hee Yoon. 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: 81–90. [[CrossRef](#)]
- Kong, Eun Jong, Mary E. Beckman, and Jan Edwards. 2011. Why are Korean tense stops acquired so early?: The role of acoustic properties. *Journal of Phonetics* 39: 196–211. [[CrossRef](#)]
- Kormos, Judit, and Mariann Dénes. 2004. Exploring measures and perceptions of fluency in the speech of second language learners. *System* 32: 145–64. [[CrossRef](#)]
- Lang, Benjamin, and Lisa Davidson. 2019. Effects of exposure and vowel space distribution on phonetic drift: Evidence from American English learners of French. *Language and Speech* 62: 30–60. [[CrossRef](#)]
- Lee-Ellis, Sungyoung. 2012. Looking into Bilingualism through the Heritage Speaker’s Mind. Ph.D. dissertation, University of Maryland, College Park, MD, USA.
- Lee, Hyunjung, and Allard Jongman. 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: 509–30. [[CrossRef](#)] [[PubMed](#)]
- Lee, Hyunjung, Jeffrey J. Holliday, and Eun Jong Kong. 2020. Diachronic change and synchronic variation in the Korean stop laryngeal contrast. *Language and Linguistics Compass* 14: e12374. [[CrossRef](#)]
- Lee, Hyunjung, Stephen Politzer-Ahles, and Allard Jongman. 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: 117–32. [[CrossRef](#)]
- Lee, Sue Ann S., and Gregory K. Iverson. 2012. Stop consonant productions of Korean–English bilingual children. *Bilingualism: Language and Cognition* 15: 275–87. [[CrossRef](#)]
- Lehiste, Ilse, and Gordon E. Peterson. 1961. Some basic considerations in the analysis of intonation. *The Journal of the Acoustical Society of America* 33: 419–25. [[CrossRef](#)]

- Llanos, Fernando, Olga Dmitrieva, Amanda Shultz, and Alexander L. Francis. 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: 2213–24. [CrossRef]
- Lowenstein, Joanna H., and Susan Nittrouer. 2008. Patterns of acquisition of native voice onset time in English-learning children. *The Journal of the Acoustical Society of America* 124: 1180–91. [CrossRef]
- Lukyanchenko, Anna, and Kyra Gor. 2011. Perceptual correlates of phonological representations in heritage speakers and L2 learners. Paper presented at the 35th Annual Boston University Conference on Language Development, Boston, MA, USA, November 5–7; Somerville, MA: Cascadilla Press, vol. 2, pp. 414–26.
- MacLeod, Andrea A. N., Carol Stoel-Gammon, and Alicia B. Wassink. 2009. Production of high vowels in Canadian English and Canadian French: A comparison of early bilingual and monolingual speakers. *Journal of Phonetics* 37: 374–87. [CrossRef]
- Major, Roy C. 1992. Losing English as a first language. *The Modern Language Journal* 76: 190–208. [CrossRef]
- Marian, Viorica, Henrike K. Blumenfeld, and Margarita Kaushanskaya. 2007. The language experience and proficiency questionnaire (LEAP-Q): Assessing language profiles in bilinguals and multilinguals. *Journal of Speech, Language, and Hearing* 50: 940–67. [CrossRef]
- Montrul, Silvina. 2002. Incomplete acquisition and attrition of Spanish tense/aspect distinction in adult bilinguals. *Bilingualism: Language and Cognition* 5: 39–68
- Nagy, Naomi, and Marisa Brook. 2020. Constraints on speech rate: A heritage-language perspective. Special Issue: Effects of Limited Input. *International Journal of Bilingualism* 1–20. [CrossRef]
- Oh, Eunhae. 2019. Korean-English bilingual children's production of stop contrasts. *Phonetics and Speech Sciences* 11: 1–7. [CrossRef]
- Oh, Janet S., Sun-Ah Jun, Leah M Knightly, and Terry Kit-fong Au. 2003. Holding on to childhood language memory. *Cognition* 86: B53–B64. [CrossRef]
- Oh, Janet S., Terry Kit-Fong Au, and Sun-Ah Jun. 2010. Early childhood language memory in the speech perception of international adoptees. *Journal of Child Language* 37: 1123–32. [CrossRef] [PubMed]
- Oh, Mira, and Robert Daland. 2011. Stops and phrasing in Korean and English monolinguals and bilinguals. Paper presented at the 17th International Congress of Phonetic Sciences, Hongkong, China, August 17–21; pp. 1530–33.
- Ohde, Ralph N. 1984. Fundamental frequency as an acoustic correlate of stop consonant voicing. *The Journal of the Acoustical Society of America* 75: 224–30. [CrossRef] [PubMed]
- Peng, Shu-hui. 1993. Cross-language influence on the production of Mandarin /f/ and /x/ and Taiwanese /h/ by native speakers of Taiwanese Amoy. *Phonetica* 50: 245–60. [CrossRef]
- Polinsky, Maria. 2008. Gender under incomplete acquisition: Heritage speakers' knowledge of noun categorization. *Heritage Language Journal* 6: 40–71. [CrossRef]
- Polinsky, Maria. 2011. Reanalysis in adult heritage language: New evidence in support of attrition. *Studies in Second Language Acquisition* 33: 305–28. [CrossRef]
- Polinsky, Maria, and Olga Kagan. 2007. Heritage languages: In the 'wild' and in the classroom. *Language and Linguistics Compass* 1: 368–95. [CrossRef]
- Rojas, Raúl, and Aquiles Iglesias. 2013. The language growth of Spanish-speaking English language learners. *Child Development* 84: 630–46. [CrossRef]
- RStudio Team. 2020. RStudio: Integrated Development Environment for R. Boston: Rstudio Team, Available online: <http://www.rstudio.com/> (accessed on 1 May 2021).
- Sancier, Michele L., and Carol A. Fowler. 1997. Gestural drift in a bilingual speaker of Brazilian Portuguese and English. *Journal of Phonetics* 25: 421–36. [CrossRef]
- Schmidt, Anna M. 1996. Cross-language identification of consonants. Part 1. Korean perception of English. *The Journal of the Acoustical Society of America* 99: 3201–11. [CrossRef] [PubMed]
- Schmidt, Anna M. 2007. Cross-language consonant identification. In *Language Experience in Second Language Speech Learning: In Honor of James Emil Flege*. Amsterdam: John Benjamins Publishing Company, vol. 17, pp. 185–200.
- Sebastián, Eugenia, and Dan I. Slobin. 1994. Development of linguistic forms: Spanish. In *Relating Events in Narrative: A Crosslinguistic Developmental Study*. Edited by Ruth A. Berman and Dan I. Slobin. Hillsdale: Erlbaum, pp. 239–84.
- Silva, David J. 2006. Acoustic evidence for the emergence of tonal contrast in contemporary Korean. *Phonology* 23: 287–308. [CrossRef]
- Son, Gayeon. 2020. Korean-speaking children's perceptual development in multidimensional acoustic space. *Journal of Child Language* 47: 579–99. [CrossRef] [PubMed]
- Stoehr, Antje, Titia Benders, Janet G. Van Hell, and Paula Fikkert. 2017. Second language attainment and first language attrition: The case of VOT in immersed Dutch–German late bilinguals. *Second Language Research* 33: 483–518. [CrossRef]
- Sundara, Megha, and Shari Baum. 2006. Production of coronal stops by simultaneous bilingual adults. *Bilingualism: Language and Cognition* 9: 97–114. [CrossRef]
- Tees, Richard C., and Janet F. Werker. 1984. Perceptual flexibility: Maintenance or recovery of the ability to discriminate nonnative speech sounds. *Canadian Journal of Psychology* 38: 579–90. [CrossRef] [PubMed]
- U.S. Census Bureau. 2018. Detailed Languages Spoken at Home and Ability to Speak English for the Population 5 Years and over for United States: 2014–2018. Available online: <https://data.census.gov/cedsci/table?q=Languages&hidePreview=true&tid=ACST1Y2018.S1601&vintage=2018> (accessed on 1 April 2021).

- Valdés, Guadalupe. 2000. *Introduction. Spanish for Native Speakers*. AATSP Professional Development Series Handbook for Teachers K-16. New York: Harcourt College, vol. I.
- Valdés, Guadalupe. 2005. Bilingualism, heritage learners, and SLA research: Opportunities lost or seized? *Modern Language Journal* 89: 410–26. [[CrossRef](#)]
- Werker, Janet F. 1989. Becoming a native listener. *American Scientist* 77: 54–59.
- Whalen, Douglas H., Arthur S. Abramson, Leigh Lisker, and Maria Mody. 1993. F0 gives voicing information even with unambiguous voice onset times. *Journal of the Acoustical Society of America* 93: 2152–59. [[CrossRef](#)] [[PubMed](#)]
- Yamada, Reiko A., and Yoh'ichi Tohkura. 1990. Perception and production of syllable-initial English /r/ and /l/ by native speakers of Japanese. Paper presented at First International Conference on Spoken Language Processing, Tokyo, Japan, November 18–22.
- Yamada, Reiko A., and Yoh'Ichi Tohkura. 1992. The effects of experimental variables on the perception of American English /r/ and /l/ by Japanese listeners. *Perception & Psychophysics* 52: 376–92.
- Yoon, Sook-Youn. 2015. Acoustic properties of Korean stops as L1 produced by L2 learners of the English language. *Communication Science & Disorders* 20: 178–88. [[CrossRef](#)]