Crosslinguistic influence on perceptual cue-weighting in Korean-English bilinguals

Yuhyeon Seo, Olga Dmitrieva

Purdue University
seo86@purdue.edu, odmitrie@purdue.edu

ABSTRACT

The current study investigates L2 crosslinguistic influence on the perceptual weighting of the cues to L1 phonological categories. Two groups of Korean-English bilinguals participated in the study: (1) long-term immigrants in the US (LTIs, n = 25) and (2) Korean heritage speakers (HSs, n = 25). Participants completed a three-alternative forced-choice identification task using a set of resynthesized stimuli containing Korean stops varying in two acoustic parameters: onset f0 and voice onset time (VOT). The results of mixed-effects logistic regression models suggested that HSs' L1 perception was influenced by their L2, as indicated by a lesser reliance on onset f0 as a cue to some Korean laryngeal contrasts, and the overall 'regularization' of the use of both cues, compared to LTIs. HSs' perception of Korean stop contrasts was also less categorical than that of LTIs.

Keywords: bilingual speech, cue-weighting, crosslinguistic influence, phonetic drift, heritage speakers

1. INTRODUCTION

Crosslinguistic influence in bilingual speech is known to be bidirectional – the first language (L1) and the second language (L2) mutually affect each other. A substantial body of research documented crosslinguistic influence in bilingual speech production [1]–[5]. Moreover, recent work suggested that bilingual speech perception is also subject to crosslinguistic influence [6]–[8]. The present study continues this line of inquiry by investigating crosslinguistic influence in Korean–English bilinguals' perception of the two acoustic cues to Korean laryngeal contrasts in stop consonants. Two groups of bilingual speakers participated in the study: (1) a group of heritage speakers of Korean (HS) and (2) a group of long-term immigrants (LTI) to the United States.

LTIs were operationalized as bilinguals who had immigrated to an L2-speaking country at the age of 18 or later and had resided there for at least 12 years. While few investigations of L1 perception by LTIs have been conducted to date, existing research suggests that L1 perceptual strategies can be affected by an L2. For instance, [6] found that in the perception of word-final stop voicing in Russian, Russian LTIs in the US relied on a cue-weighting pattern similar to that of L1 American English speakers and distinct from that of L1-immersed Russian speakers, suggesting L2 influence on L1 perception.

HSs were operationalized as bilinguals who speak their parents' language (L1) at home, while speaking an L2 elsewhere in which they are more proficient and dominant than in the L1 [9]. One of the defining characteristics of HSs' linguistic development is the dominance reversal they typically undergo during early childhood, when an L2 starts overtaking their L1 in the amount of use and exposure. The National Heritage Language Survey [10] reports that almost 99% of Korean HSs in the US become English (L2) dominant by the age of 17.

This dominance pattern combined with early and substantive L2 exposure suggests that L2 influence on HSs' L1 perceptual strategies should be even more evident than for LTIs. Nevertheless, previous studies suggested that L1 phonetics of HSs could be immune to crosslinguistic influence from the L2, especially in the perceptual domain [1], [8], [11]. For instance, studies on the perceptual discrimination of Korean stops by HSs showed accuracy comparable to that of L1-immersed native speakers, while outperforming L2 learners [8], [11]. However, we hypothesized that the methodological choices were partially responsible for these findings. Specifically, we posited that discrimination tasks are not sensitive enough to detect L2 influence on L1 perception, as they do not illuminate the pathway to discrimination decisions.

In the present study, we adopt a cue-weighting approach to investigate the degree of an L2 crosslinguistic influence on L1 speech perception in HSs compared to LTIs. While both groups had a significant amount of exposure to the L2, they differed in the timing of L2 exposure (early childhood for HSs vs. early adulthood for LTIs), as well as the quality and quantity of L1 exposure during childhood. HSs acquired Korean from birth but a large share of their linguistic exposure and use began to be dedicated to the L2 already in early childhood, eventually leading to dominance reversal. LTIs, who came to the US in early adulthood, were exclusively Korean-dominant for a considerably longer period of their life. Therefore, the current study pursues the

hypothesis that HSs' L1 perception should be affected by the L2 more than LTI's L1 perception.

The present study focuses on the two acoustic properties, onset f0 and voice onset time (VOT), as acoustic cues to laryngeal contrasts in both Korean and English stop. As shown in Table 1, the Korean language exhibits a three-way distinction in stop consonants (lenis-aspirated-fortis, /p/-/p^{h}/-/p*/) [12], while English has a two-way voicing contrast in stops (/p/-/b/) [13]. In the word-initial position, both Korean and English stops are distinguished by the same two acoustic parameters, but the two languages differ in their relative weighting of these cues. While VOT plays a primary role in the voicing contrast in English stops, onset f0 is the primary cue to the lenis-aspirated contrast in modern Seoul Korean [12], [14]–[16].

Therefore, we hypothesized that L2 influence on the L1 will manifest itself as a greater reliance on VOT and a lesser reliance on onset f0 in the identification of Korean stops, given the difference in the functional load of these cues in English and Korean laryngeal phonology. Therefore, HSs are predicted to rely on VOT to a greater extent and on onset f0 to a lesser extent compared to LTIs.

Korean	VOT	Onset f0
Lenis	Long-lag	Low
Aspirated	Long-lag	High
Fortis	Short-lag	High
English		
Voiceless	Long-lag	High
Voiced	Short-lag	Low

Table 1: Stop contrasts in Korean and English.

2. METHODS

2.1. Participants

A total of 50 Korean-English bilinguals were recruited in the Midwestern areas of the US. The LTI group consisted of 25 Korean immigrants who moved to the US from South Korea at the age of 18 or later and were, on average, 39 years old at the time of the participation (SD = 6.2). They had resided in the US for an average of 12 years by the time of the experiment. The HS group consisted of 25 speakers who were born and raised in the US by Korean parents on both sides. HSs were, on average, 27.3 years old at the time of the experiment (SD = 6.7). Care was taken to only enroll participants whose parents were born in Korea after 1950 since the sound change that shifted the burden of differentiating lenis and aspirated stop from VOT to onset f0 is believed to have begun by 1940 [12], [16]. Participants completed a Bilingual Language Profile [17] to verify their language dominance and proficiency. The self-reports of Korean and English proficiency indicated that LTIs were more balanced bilinguals than HSs in speaking, listening, reading, and writing. Most HSs were more proficient and dominant in English than in Korean. HSs showed greater variability in terms of Korean use in daily life compared to LTIs, who used Korean more frequently than HSs. HSs also indicated a significantly less formal education of Korean than LTIs ($t(38.7) = 8.35$, $p < .001$) (HS: M = 3.3 years, SD = 3.5; LTI: M = 14.8, SD = 5.9).

2.2. Stimuli

To implement a three-alternative forced-choice (3AFC) identification task, 64 unique stimuli were created by resynthesizing a real-word minimal pair, /pul/, 'fire' and /p^{h}ul/, 'grass' produced by a male Korean speaker in his twenties, using Praat [18]. The initial stops in the stimuli differed orthogonally in terms of VOT and onset f0, based on an eight-by-eight continuum with log-scaled VOT steps (10, 14, 19, 27, 38, 53, 74, and 104 ms) and equidistant f0 steps in 10 Hz steps ranging from 120 to 190 Hz. The f0 values were resynthesized to be consistent throughout the vocalic part in each stimulus.

2.3. Procedures

All participants completed a 3AFC identification task, using PsychoPy [19] in either a speech lab or at a quiet place of their preference. Participants were wearing headphones throughout the task. In this task, a single stimulus was played on each trial, and participants identified it by clicking with a mouse on one of the three response options on a monitor screen: /pul/, 'fire,' /p^{h}ul/, 'grass,' and /p*ul/, 'horn'. Trials were self-timed. The following trial was presented only after a response was provided in the preceding trial. The response options were presented in Korean orthography, and the order of the response options was counterbalanced across trials. The full set of stimuli was randomized and repeated six times to each participant. Each participant took approximately 15 minutes to complete a total of 384 trials.

2.4. Analysis

Participants' responses were paired by stop contrast, and the resulting three pairs (lenis-aspirated, fortis-aspirated, and lenis-fortis, and) were submitted as binary dependent variables to three separate mixed-effects logistic regression analyses implemented in R (version 4.2.1) [20] using the 'lme4' package [21]. Each model was sum-coded and included Group, VOT, Onset f0, and interactions of Group with VOT and with Onset f0 as fixed effects. The models also

included by-subject intercepts and slopes as random effects to account for individual sensitivity to each acoustic cue.

3. RESULTS

3.1. Lenis-aspirated contrast

Lenis and aspirated stops in modern Seoul Korean are believed to be differentiated primarily on the basis of onset f0 values, and, to a lesser extent, VOT [12], [14]–[16]. The statistical results indicated that while both cues were significant predictors of the identification response (onset f0; β = 1.20, SE =.06, z = 19.23, p < .001; VOT; β = .99, SE =.05, z = 19.35, p < .001), onset f0 was a primary cue, as indicated by a greater coefficient than for VOT, which played a secondary role for both groups. However, the interaction between Group and Onset f0 indicated that the two groups differed in terms of their reliance on onset f0 (β = -.28, SE =.06, z = -4.48, p < .001). Specifically, HSs assigned lower weight to the vocalic cue of onset f0 contrast than LTIs. This effect is illustrated in Figure 1 (left panel), where the identification slope of HSs is visibly less steep than that of LTIs. On the other hand, the absence of an interaction between Group and VOT indicated that participant groups were not significantly different from each other in their use of VOT (p = .84).

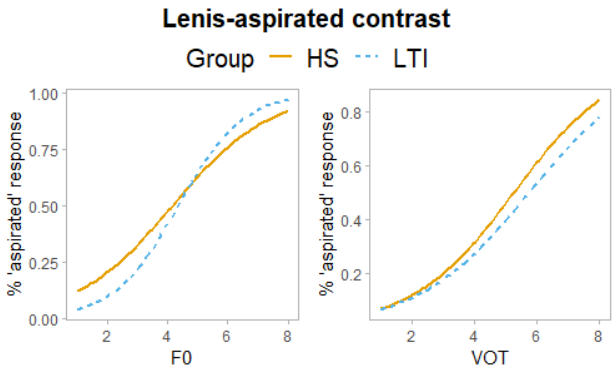

Figure 1: Identification curves for the lenis-aspirated stop contrast by group and by cue.

3.2. Lenis-fortis contrast

Lenis and fortis stops in Korean are contrasted via both acoustic correlates, onset f0 and VOT; lenis stops are characterized by longer VOT and lower onset f0 values than fortis stops [15]. The logistic regression analysis confirmed that both VOT and onset f0 were significant predictors of lenis vs. fortis responses (onset f0; β = .91, SE =.06, z = 16.07, p < .001; VOT; β = -.67, SE =.05, z = -13.5, p < .001). In addition, both interaction terms (Group by VOT and Group by Onset f0) were significant. These interactions were due to the fact that HSs assigned lower weight to onset f0 than LTIs (β = -.27, SE =.06, z = -4.88, p < .001), while assigning greater weight to VOT than LTIs (β = .17, SE =.05, z = 3.51, p < .001). This means that the two participant groups differed in whether they used VOT or onset f0 as the primary cue to the lenis-fortis contrast. These differences are illustrated in Figure 2. For example, LTIs have a visibly steeper identification curve as function of onset f0 (left panel).

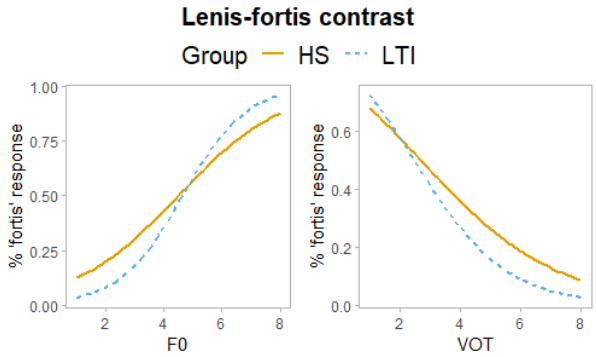

Figure 2: Identification curves for the lenis-fortis stop contrast by group and by cue.

3.3. Fortis-aspirated contrast

Fortis and aspirated stops in Korean are distinguished mainly via VOT, but not onset f0 [16]. Nevertheless, the statistical analysis demonstrated that both VOT and onset f0 were significant predictors of the response (onset f0; β = .24, SE =.04, z = 6.19, p < .001; VOT; β = 1.67, SE =.08, z = 19.99, p < .001), with VOT being the primary cue as expected. Additionally, a significant interaction of group and onset f0 (β = .12, SE =.04, z = 3.21, p = .001) showed that HSs relied on onset f0 more (by a factor of 0.25) than LTIs. In fact, as seen in Figure 3 (left panel), LTIs did not use f0 to any appreciative degree. While the HSs' identification curve went up as f0 increased, LTIs function remained almost level. In contrast, the significant interaction between Group and VOT was due to a greater reliance of LTIs than HSs on VOT.

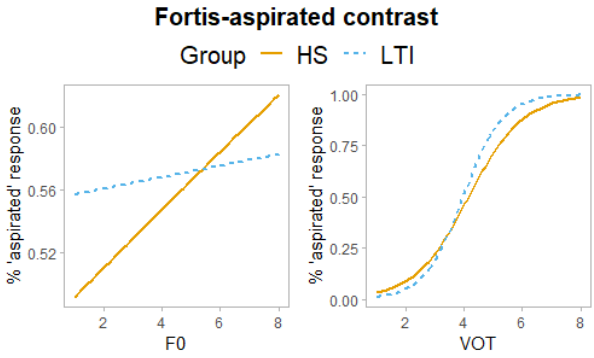

Figure 3: Identification curves for the fortis-aspirated stop contrast by group and by cue.

4. DISCUSSION

We hypothesized that L2 influence would be imposed on HSs' L1 perception to a greater degree than on LTIs' L1 perception, due to the differences between the two groups in terms of the timing, quantity, and quality of exposure to both L1 and L2. Specifically, we predicted that HSs would be less reliant on onset f0 and more reliant on VOT in identifying Korean stop categories, compared to LTIs. This hypothesis was based on previous studies demonstrating that onset f0 plays a secondary role in cueing laryngeal stop categories in English unlike in Korean [7].

The results partially supported these hypotheses. First and foremost, HSs presented a markedly different cue-weighting pattern in identifying the three Korean stop categories, as compared to LTIs. Second, as predicted, in two out of three binary contrasts, HSs relied on onset f0 significantly less than LTIs in making their perceptual decisions.

Specifically, the lenis-aspirated contrast was perceived primarily on the basis of onset f0 by both groups, indicating a presence of tonal contrast for both bilingual groups. Nevertheless, HSs relied on onset f0 significantly less than LTIs.

The lenis-fortis contrast is cued by both parameters in Korean stops [15]. The results confirmed this cue-assignment in both participant groups to a significant degree in identifying stops. Nevertheless, HSs used onset f0 significantly less than LTIs.

These findings are compatible with the prediction that lower functional load of onset f0 in English laryngeal phonology would lead to a lesser reliance of HSs on f0 when identifying Korean stops. Nevertheless, an alternative explanation should be acknowledged. HSs' native language input differed from that of the LTIs. Critical to our investigation is the possibility that participants' parents did not participate in the sound change that increased reliance on onset f0 as a cue to laryngeal categories, or that they spoke a dialect of Korean that did not undergo this sound change [16]. Moreover, the parents' Korean speech could already be influenced by English at the time of its transmission to our participants. While we cannot rule out the latter, some of the results discussed below speak against the sound change non-participation hypothesis.

Regarding the use of VOT, contrary to the expectation, there was little evidence that HSs were more reliant on VOT as a cue to laryngeal categories than LTIs, despite the central role that VOT plays in cuing English voicing. Thus, the overall pattern suggests HSs rely less on onset f0 without increasing their reliance on VOT, as opposed to LTIs.

A puzzling case is presented by the fortis-aspirated contrast, normally cued exclusively by VOT in Korean [16]. LTIs behaved accordingly, relying exclusively on VOT in making the fortis-lenis decisions. Surprisingly, HSs made use of onset f0, in contrast to LTIs. While this finding continues the pattern of a distinct perceptual behavior by HSs compared to LTIs, it does not align with the hypothesis that influence from English will lead to pervasive underutilization of onset f0 in L1 Korean. It also does not support the hypothesis that parents' non-participation in the relevant sound change leads to underuse of f0. Instead, HSs 'regularized' their cue-weighting approach to all Korean laryngeal contrasts, always relying to some degree on both cues, VOT and onset f0. To a certain extent, this strategy could also be viewed as English-based. Perceptual studies suggest that, in English, both VOT and onset f0 are often attended to by native listeners when making voicing decisions, albeit to a different degree, [7], [13], unlike in Korean, where, depending on the contrast, listeners are supposed to attend to one cue while largely ignoring the other one [14].

Finally, it should also be noted that, in most cases and independently of the cue, HSs' identification curves were shallower than those of LTIs, suggesting a less categorical perception overall. This observation is also in accordance with the hypothesis that HSs' early exposure and greater dominance in their L2, English, has consequences for their L1 perception. In this case, the consequence is the decreased overall sensitivity to Korean laryngeal contrasts.

Therefore, the results presented here provide evidence for the hypothesis that bilinguals' L2 can affect their L1 perception and that earlier exposure and greater dominance in L2 lead to stronger crosslinguistic influence on L1 perception.

This study is among the first few to uncover the effects of bilingualism on the perception of native speech [1], [4], [6]. The results confirm that HSs' L1 perception is not immune to the influence of L2, similarly to other aspects of their native language competence [22]. Finally, the present findings fit well within theories of L2 acquisition, such as SLM-r [23], which postulates that L1 and L2 sound categories co-exist in the same phonetic space and, therefore, are predicted to affect each other in a bidirectional manner. While in the past such predictions were made and tested almost exclusively on the basis of speech production data, the recent revision of SLM (SLM-r) proposed that production and perception co-evolve during L2 acquisition. The present findings lend support to this theoretical assumption.

5. ACKNOWLEDGEMENT

We are very grateful to Dr. Jennifer Cabrelli (University of Illinois at Chicago) for invaluable assistance with data collection in Chicago, Illinois. We also thank the anonymous reviewers for constructive suggestions.

6. REFERENCES

[1] Chang, C. B. 2016. Bilingual perceptual benefits of experience with a heritage language, *Bili.: Lang. and Cogn* 19, 791–809.

[2] Chang, S.-E., Weiss-Cowie, S. 2021. Hyper-articulation effects in Korean glides by heritage language learners, *International. J. of Bilingualism.* 25, 3–20.

[3] Flege, E. 1987. The production of 'new' and 'similar' phones in a foreign language: Evidence for the effect of equivalence classification, *J. of Phon.* 15, 47–65.

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

[5] Cheng, A. 2017. VOT merger and f0 contrast in Heritage Korean in California, *UC Berkeley PhonLab Annu. Rep.* 13, 1.

[6] Dmitrieva, O. 2019. Transferring perceptual cue-weighting from second language into first language: Cues to voicing in Russian speakers of English, *J. of Phon.* 73, 128–143.

[7] 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, *J. of Phon.* 49, 77–95.

[8] 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, 6.

[9] Valdés, G. 2005. Bilingualism, heritage language learners, and SLA research: Opportunities lost or seized? *Modern Language* 89, 410–426.

[10] 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, 40–64.

[11] Oh, S., Jun, S.-A., Knightly, L. M., Au, T. K. 2003 Holding on to childhood language memory, *Cognition* 86, B53–B64.

[12] Silva, D. J. 2006. Acoustic evidence for the emergence of tonal contrast in contemporary Korean. *Phonology* 23, 287–308.

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

[14] Kong, E., Holliday, J., H. Lee, H. 2022 Post-adolescent changes in the perception of regional sub-phonemic variation, *J. of Phon.* 90.

[15] Kang, K.-H., Guion, S. G. 2006. Phonological systems in bilinguals: Age of learning effects on the stop consonant systems of Korean-English bilinguals, *J. Acoust. Soc. Am.* 119, 1672–1683.

[16] Kang, Y., Nagy, N. 2016. VOT merger in Heritage Korean in Toronto. *Language Variation and Change* 28, 249–272.

[17] Birdsong, D., Gertken, L. M., Amengual, M. 2012. Bilingual language profile: An easy-to-use instrument to assess bilingualism, *COERLL Univ. Tex. Austin*.

[18] Boersma, P., Weenink, D. 2012. *Praat Doing Phon. by Computer*.

[19] Peirce, J. W. 2007. PsychoPy—psychophysics software in Python. *J. of Neuroscience Methods* 162.

[20] R Core Team. (2022). *R: A language and environment for statistical computing*. URL https://www.R-project.org/.

[21] Bates, D., Mächler, M., Bolker, B., Walker, S. 2014. Fitting linear mixed-effects models using lme4, *ArXiv Prepr. ArXiv14065823*.

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

[23] Flege, J. E., & Bohn, O.-S. (2021). The revised speech learning model (SLM-r). Second Language Speech Learning: Theoretical and Empirical Progress, 3–83.