Development of Phonetic Cues in Early L2 Speech Production: The Case of Korean Plosives Pronounced by Native German Speakers

Yeongeun Choi[1], Christoph Draxler[2]

[1]Department of Computational Linguistics, University of Zurich; [2]Institute of Phonetics and Speech Processing, Ludwig-Maximilians-University Munich

yeongeun.choi@uzh.ch; draxler@phonetik.uni-muenchen.de

Abstract

This study explores the developmental trends in L2 Korean stop production by German learners, focusing on VOT and F0 of the following vowel. Analyzing results from a reading task with beginners and advanced learners, it was found that beginners demonstrated an L1-like binary stop contrast, relying solely on VOT: short (fortis/lenis) versus long (aspirated). Advanced learners, however, showed notable progress by (1) lengthening VOT for lenis plosives, and (2) employing F0 as a primary cue, with lower F0 after lenis and higher F0 after fortis/aspirated plosives. Thus, learners progressively adapted both 'similar' and 'dissimilar' acoustic cues in a target-like fashion.

Index Terms: Korean stops, Voice Onset Time, fundamental frequency, phonetic cue weighting, L2 speech production

1. Introduction

The most influential second language (L2) speech models highlight the importance of (dis)similarities between learners' first language (L1) and their target language on L2 phone learning. The Speech Learning Model (SLM, [1], [2]) as well as its revised version SLM-r [3] suggest that L2 learners can easily build a new phonetic category for an unfamiliar phoneme that does not exist in their L1 sound system. In the same vein, the Perceptual Assimilation Model (PAM, [4]) predicts that the similarities between L1 and L2 sounds rather hinder the perception of unknown L2 sounds. Accordingly, the PAM-L2 model [5] claims that novice learners tend to substitute L2 phones with the closest L1 equivalent, leading to less precise L2 pronunciation at the initial stage of learning. However, as L2 learning progresses, learners filter informative acoustic features necessary for phonemic distinctions, thereby potentially improving their pronunciation accuracy, which occurs more prominently for dissimilar L2 phones than for similar ones. This study aims to explore these issues further by examining how German learners use suprasegmental cues when producing Korean homorganic plosives, tracking the phonetic shifts that may occur over the course of learning.

Both Korean ([6], [7], [8]) and German ([9], [10]) plosives are voiceless in word-initial position, mainly distinguished along the short-to-long lag *Voice Onset Time* (VOT) spectrum. Despite this similar laryngeal pattern, Korean plosives differ in several respects from the German ones, which may pose notable challenges for German speakers in L2 Korean stop production.

Most importantly, Korean has a phonemic three-way contrast between homorganic plosives as follows: fortis /p* t* k*/[1], lenis /p t k/, and aspirated /p^h t^h k^h/. These plosives are differentiated by two primary acoustic parameters, VOT and the *fundamental frequency* (F0) of the following vowel [7], [11]. Previous studies on sound change of standard (Seoul) Korean plosives (e.g., [8], [12], [13]) have demonstrated that VOT and vowel F0 alternately signal the stop contrasts as primary distinctive cues in word-initial position. The co-existence of these two primary cues has become robust over decades through a VOT merger between lenis and aspirated plosives in standard Korean (for details of VOT merger see also [6]). As a result, VOT serves as a primary cue only for discriminating short-lag VOT (fortis) from long-lag VOT (lenis/aspirated), while vowel F0 serves as a primary cue for signaling a contrast between laryngeal tension (fortis/aspirated) and laryngeal laxness (lenis).

Conversely, standard German has a two-way contrast between lenis /b, d, g/ and fortis /p^h t^h k^h/ plosives, where 'fortis' plosives phonetically resemble Korean aspirated plosives. In German, VOT is the main phonetic cue which distinguishes between homorganic plosives [10], [14], [15], whereas vowel F0 fluctuation [9], [16] occurs unintentionally and naturally as a result of obstruent-intrinsic F0 effects [17].

Drawing from these commonalities and differences between L1 German and L2 Korean in suprasegmental dimensions, as well as their typological mismatch, our research will focus on the following aspects:

(a) Categorization of L2 Korean plosive contrasts at the very early stages of learning. To our knowledge, no studies have been conducted on the production of L2 Korean stops by native German speakers. [18] is the only perception study examining German naïve listeners with Korean stop discrimination. In this test, German listeners exhibited the poorest performance with fortis-lenis pairs, while they effectively discriminated these two plosives from aspirated plosives, respectively. This indicated that German listeners distinguished the three-way contrasts of Korean plosives in a binary manner, in line with their L1 categories. This result is of particular importance, as it suggests that German beginner learners are likely to employ a binary categorization in Korean stop production as well. This expectation may find additional support in other previous research on L2 Korean stops. In both perception and production, learners whose L1 features a two-way stop contrast categorized Korean lenis plosives either with fortis plosives (e.g., English learners [19]) or with aspirated plosives (e.g., Japanese learners [20]), reflecting the stop categorization patterns of their L1. In other words, no learner group employed a two-way discrimination in perception while adopting a three-way discrimination in production. Moreover, these learners predominantly relied on the VOT cue, demonstrating nearly absent proficiency with the F0 cue. This leads us to another expectation that German speakers similarly rely more (or solely) on the VOT cue, which serves as a primary distinctive cue in their L1. We will also analyze whether the incorrectly categorized L2 phonological category – potentially lenis in our case – will be produced as simply equivalent to the single L1 phonological category, or whether it will be

accompanied by good or poorer exemplars (for further details, see PAM [4] and PAM-L2 [5]).

(b) Phonetic shifts in both VOT and F0 cues during the learning process. The developmental changes in the use of VOT and F0 cues will be examined over the course of a one-semester undergraduate course. In terms of the F0 cue, a longitudinal analysis [21] reported that native Mandarin speakers with a tonal language background failed to effectively use the F0 cue, even after approximately one year of Korean language learning, maintaining their conservative behavior in using F0. By contrast, our target group is in the opposite situation, as they do not 'actively' utilize the F0 differences in their L1. Our empirical data will therefore contribute to further insights into how speakers from non-tonal languages manage the F0 cue in L2 stop production.

(c) Change in categorization of L2 plosives at the late early stages. Lastly, the production of advanced beginners will be compared with that of novice beginners, concerning aspects (a) and (b).

To address these research aims, a laboratory speech corpus was collected from 23 German learners of L2 Korean by conducting a carefully designed reading task.

2. Method

2.1. Participants

Two groups of German learners of L2 Korean were recruited in Frankfurt, Germany: 13 beginners and 10 advanced beginners. All participants are female, aged 18-24 years (Mean = 20.3). They attended a Korean language course at Goethe University, three days a week, each session lasting 90 minutes. The 'beginner' group had completed a 10-hour Korean language course, focusing on learning the Korean alphabet and basic pronunciation. The 'advanced beginner' group, hereafter referred to as the 'advanced' group, had completed a 60-hour course, focusing on learning grammar at the sentence level.

2.2. Stimuli

In this study, we adopted a reading task previously conducted in [6] recruiting a different set of participants. The task involved 72 word pairs '**A (target)** and B' displayed within the context of interrogative carrier sentences, which always begin with the adverb *eonje* 'when'. The target word A contains a word-initial plosive, and both A and B are disyllabic words, followed by a postposition *wa/gwa* 'and' for A; an accusative case marker *eul/reul* for B. By employing a word pair framework, we were able to prevent excessive focus on the target word, such as hyperarticulation. The sentences were structured as follows: [*eonje* / A-*(g)wa* / B-*(r)eul* / verb ?]. The test sentences were presented in Korean only.

The word pair items were categorized into four types, by combining them in the following manner: real-real, real-pseudo, pseudo-real, pseudo-pseudo. Pseudo words were used to reduce lexical influence, while simultaneously facilitating the collection of a sufficient sample size. The pseudo A words are quasi-duplicates of the real A words, created by altering the coda of the real A words, for instance, *ttangkong* 'peanut' – *ttanko*; *panmae* 'sale' – *pamae*, as shown in Table 1. By contrast, none of the B words formed minimal pairs with the real and pseudo words. This minimal pair-like speaking setup was designed to analyze learners' pronunciation more precisely, and to enhance the reliability of the speech production data.

Table 1. *Examples of four types of word pair items 'A and B' where the first syllable of the target word A is bold, and the carrier sentence is omitted.*

Type	A (target)	B
real-real	**ttang**kong	hodu
real-pseudo	**pan**mae	donmae
pseudo-real	**ttan**ko	gimbap
pseudo-pseudo	**pa**mae	haneo

To control for vowel-intrinsic effects on F0 [22], the vowels after the target plosives were restricted to the vowels /a/ and /o/. Overall, a total of 1,656 tokens (3 places of articulation x 3 plosive types x 2 vowels x 4 types of word pair x 23 subjects) were collected for analysis. Following the method described by [23], 10 warm-up sentences with other initial consonants were also created, consisting solely of actual Korean words.

2.3. Procedure

The same reading task was conducted in two separate sessions, spaced approximately four months apart. Four speakers participated in both experiments; however, individual variations were not accounted for in this study.

Both experiments followed the same procedure, beginning with the warm-up sentences, followed by the sentences containing the target words. All sentences were randomly presented on a computer screen, using the *SpeechRecorder* software [24]. Generally, speakers read the given sentences aloud a single time. If they mispronounced the target word, they were asked to repeat the entire sentence; however, no hints were provided regarding what was mispronounced.

The participants were recorded in soundproof recording facilities at Goethe University Frankfurt, Germany, using a Røde Lavalier microphone attached to a Zoom H4n audio recorder. All recordings were sampled at 44.1 kHz and 16 bits.

2.4. Measurements and statistical analyses

The collected recordings were automatically segmented and labeled using a Korean forced alignment tool [25]. Subsequently, the time intervals of interest were manually corrected using Praat [26] as follows: for VOT, from the stop release burst to the voicing onset of the following vowel; for vowel F0, from the voicing onset to the offset of the following vowel. These target cues were estimated using separate statistical models as follows.

For the VOT as temporal values, the annotated VOT durations were normalized by calculating z-scores to control for speech rate across speakers. A linear mixed-effects model (LMEM) was conducted in R [27] with z-scored VOT duration as dependent variable, using the R packages *lme4* [28] and *lmerTest* [29]. The model employed as fixed factors *level* (beginner, advanced), *plosive type* (fortis, lenis, aspirated), *place of articulation* (bilabial, alveolar, velar), and *vowel* (/a/, /o/) including their interaction. The random structure included intercepts and random slopes by *speaker* and *target word*. P-values were computed using the Satterthwaite's method with Tukey adjustment for multiple comparisons. Additionally, a post-hoc Tukey's test was performed to examine the level-related VOT differences, using the R package *emmeans* [30].

Prior to statistical analysis on the F0 as non-linear values, each single vowel was divided into 30 time-points in Praat to control for variable vowel durations across speakers and target words. The F0 values measured at each time frame point were

normalized using the mean and standard deviation (SD) of the overall F0, averaged across all time frames and all speakers.

The z-scored F0 values were analyzed using generalized additive mixed models (GAMMs) with the R package *mgcv* [31]. This method is particularly effective for examining the non-linear patterns of F0 variation over time-normalized vowel intervals, allowing for the observation of how vowel F0 variations are modulated by the independent variables. The model contained three parametric terms, *level*, *plosive type*, and *vowel*, incorporating by-*target word* random intercepts, along with by-*speaker* random slopes for *Interval*. The overall F0 contours across the learner groups were plotted based on the three-way interaction of *level*, *plosive type*, and *vowel*, including a smooth term for *Interval* by this interaction. In addition, to compare the patterns of vowel F0 among *plosive type* for each group of learners and each vowel type, the parametric coefficients and approximate significance of smooth terms were separately computed considering the interaction between *level* and *vowel*.

3. Results

3.1. Voice onset time

The VOT results revealed significant main effects of *plosive type* (henceforth *type*, $F[2, 62] = 161.31$, $p < .001$) and *place of articulation* (henceforth *PoA*, $F[2, 62] = 49.18$, $p < .001$). In contrast, the effects of *level* ($F[1, 23] = 0.65$, $p = 0.43$) and *vowel* ($F[1, 62] = 2.24$, $p = 0.14$) were not significant. However, the interactions of *level*type* ($F[2, 51261] = 288.34$, $p < .001$), of *level*PoA* ($F[2, 51260] = 124.35$, $p < .001$), and of *level*type*PoA* ($F[4, 51260] = 46.67$, $p < .001$) were confirmed as significant, indicating that learners adjusted VOT durations based on articulatory manner and places of plosives, and altered their use of VOT as they progressed in learning Korean.

Specifically, as depicted in Figure 1, beginner learners distinguished the VOT durations for the three-way plosives in a binary manner: short-lag VOT for fortis/lenis plosives and long-lag VOT for aspirated plosives. Beginner learners showed no significant VOT difference between fortis and lenis plosives, as confirmed by post-hoc pairwise comparisons ($p = 0.98$). In contrast, for advanced learners, the VOT distinction between fortis and lenis plosives became evident ($p < .001$). However, it is noteworthy that the VOT values for lenis plosives are widely distributed, resulting in an overlap of VOT with both fortis and aspirated plosives. Furthermore, VOT durations were generally longer compared to beginner learners.

Considering *PoA*, its significant effect and interaction with *level* and *type* suggest that each learner group displayed different VOT patterns based on *type*. Generally, plosives with a shorter VOT, namely fortis and lenis plosives, were more influenced by *PoA* than aspirated plosives, which have the longest VOT, in terms of VOT adjustment. Commensurate with Figure 1, pairwise comparisons revealed that the VOT durations of fortis and lenis plosives increased from the anterior (bilabial/alveolar) to the more posterior places of articulation (velar). Both bilabial-velar and alveolar-velar comparisons across *level* were significant ($p < .001$). However, within the anterior places of articulation, i.e., between bilabials and alveolars, *PoA* did not significantly affect the VOT for both fortis and lenis plosives. This trend was observed in both the beginner and the advanced group (beginner – fortis: $p = 0.36$; advanced – fortis: $p = 0.65$; lenis: $p = 0.54$). The lenis plosives in beginners exhibited a statistically significant difference between bilabials and alveolars ($p < .05$); however, when compared to their fortis counterparts, the variation in VOT at these two places was not as strong.

The *PoA* effect on VOT of aspirated plosives was not significant in either group: beginner – bilabial-alveolar: $p = 0.78$; bilabial-velar: $p = 0.10$; alveolar-velar: $p = 0.34$; advanced – bilabial-alveolar: $p = 0.96$; bilabial-velar: $p = 0.26$; alveolar-velar: $p = 0.39$).

3.2. F0 contour of the following vowel

The F0 contours are illustrated in Figure 2, incorporating *level*, *type*, and *vowel*, and explaining 72.1% of the variance in the dependent variable. Given the distinctions in F0 contours observed between *level* and between *vowel*, the same model structure was fitted separately to four datasets derived from the interaction of *level*vowel*. Firstly, for the vowel /a/, the models of the beginner and advanced groups explained 68.8% and 76.6% of the variance of *type*, respectively, incorporating random intercepts for *word* and random slopes for *speaker*. The parametric coefficients indicated that in the case of vowel /a/, beginner learners displayed no significant differences in the F0 contour between the three plosives (fortis: $\beta = 0.03$, $p = 0.89$; lenis: $\beta = -0.09$, $p = 0.10$; aspirated: $\beta = 0.08$, $p = 0.14$), nor did advanced learners (fortis: $\beta = -0.34$, $p = 0.27$; lenis: $\beta = -0.06$, $p = 0.43$; aspirated: $\beta = 0.10$, $p = 0.21$). Despite the lack of statistical significance, we can still observe in the top panels of Figure 2 that aspirated plosives before the vowel /a/ consistently exhibited the highest F0 curves at both levels.

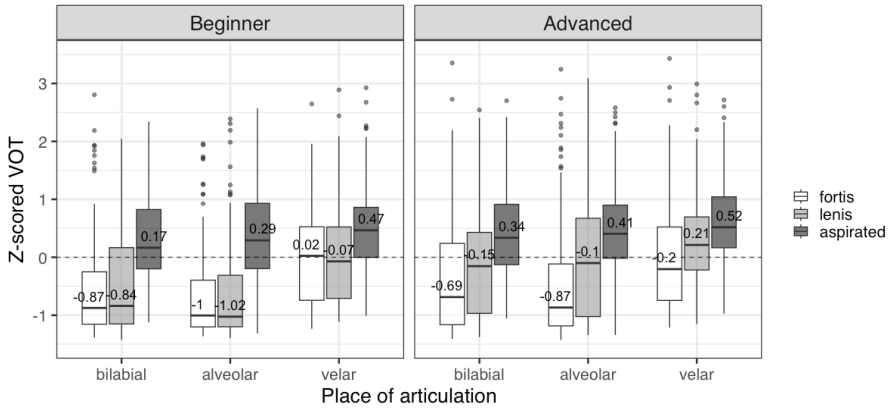

Figure 1: *Z-scored VOT as a function of (a) plosive type and (b) place of articulation across two Korean language levels*

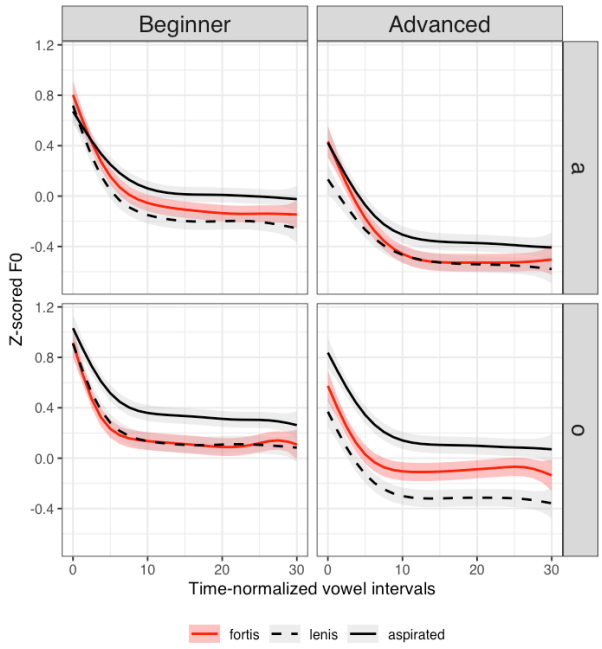

Figure 2: *Average time-normalized standardized F0 contours of the following vowels /a/ and /o/ for fortis, lenis, and aspirated plosives*

For the vowel /o/, the models of beginner and advanced groups explained 66.7% and 75.9% of the variance of *type*, respectively. As can be seen in the bottom panels of Figure 2, both beginner and advanced groups exhibited significantly higher F0 contours for aspirated plosives (beginner: $\beta = 0.22$, $p < .01$; advanced: $\beta = 0.21$, $p < .01$), which was notably distinguished from the other two plosive types. The parametric coefficients demonstrated further that advanced learners significantly discriminate between fortis (red) and lenis (dashed) plosives, using the F0 cue (fortis, intercept: $\beta = -0.04$, $p = 0.90$; lenis: $\beta = -0.20$, $p < .01$) when followed by vowel /o/. However, beginners demonstrated almost identical F0 curves for fortis and lenis (fortis, intercept: $\beta = 0.20$, $p = 0.30$; lenis: $\beta = 0.00$, $p = 0.97$).

Moreover, for advanced learners, the F0 of vowels following lenis plosives shifted downward for both /a/ and /o/, whereas it shifted upward for vowels following fortis plosives for /o/. In comparison with beginner learners, these observations suggest that advanced learners increasingly employed the F0 cue of the following vowel to distinguish between plosive types: laryngeal tension (fortis/aspirated) vs. laryngeal laxness (lenis).

4. Discussion and conclusion

The present study investigated the developmental trajectories of cue use in L2 Korean stop production by German learners during the initial two stages of learning, specifically beginner and advanced beginner levels.

At the initial stage of learning, learners tend to rely solely on the most familiar and reliable acoustic cue, VOT. The results from beginners demonstrate that the three plosives were distinctively categorized into two categories, i.e., those with short-lag VOT (fortis/lenis) and those with long-lag VOT (aspirated). The categorization of lenis plosives as short-lag VOT may be attributed to the orthographic influence being more significant in the initial stage of learning (e.g., [32]) than it is later on. Novice learners, likely influenced more by orthographic information than by the limited auditory inputs they had gathered, appeared to categorize <b d g> as "not (strongly) aspirated" in contrast to <p t k>. These incorrect binary stop categories observed in novice learners also showed a proportional relationship between VOT and vowel F0, suggesting that beginners barely adjust the F0 cue.

The novice learners' L1-like realizations of L2 sounds significantly changed as learning progressed. At later learning stage, specifically after completing a 60-hour Korean course, learners had classified fortis and lenis plosives from a single category into two distinct categories, by adjusting the cue reliance between VOT and vowel F0.

With regard to the VOT adjustment, advanced learners appeared to have acquired during their learning that Korean lenis plosives are more aspirated in word-initial position than they initially recognized, leading to a VOT split between fortis and lenis plosives. However, the VOT durations for both plosives remained widely distributed compared to those of beginners, indicating that the learners' use of VOT has not been fully stabilized yet, and their VOT distribution may continue to change at more advanced stages of learning. Similarly, the overshooting of VOT was also observed across plosive types. Especially, the excessive aspiration found in aspirated plosives may be due to learners' effort to pronounce them with distinctly longer aspiration, particularly longer than that of lenis plosive with lengthened VOT durations. These learners' efforts to precisely control VOT were also highly influenced by articulatory position. The VOT distribution for advanced learners varied according to place of articulation across plosive types: the further back the plosives were articulated, the longer the VOT produced. This contrasts with the articulatory behaviors of (young) native Korean speakers, who exhibit relatively little VOT variation depending on place of articulation [6]. Nonetheless, it is important to note that despite this unstable use of the VOT cue, the direction of changes appears to progress towards a more native-like manner.

The F0 shift of fortis and lenis plosives provides further evidence for the advancement of learners' cue use. Specifically, fortis and lenis plosives did not conform to a universal pattern of obstruent intrinsic effects on F0 (cf. [33]). That is, although the VOT of lenis plosives is as long as that of aspirated plosives, the following vowel exhibits the lowest F0 values, prominently lowering the vowel F0 for lenis plosives. For fortis plosives, by contrast, the VOT was the shortest, yet associated with a higher F0 than vowels following lenis plosives. Considering that F0 variation is an articulatory consequence in L1 German – with longer VOT accompanied by higher F0, our results indicate that German learners have begun adopting the F0 cue as a primary distinctive parameter for discriminating between Korean plosives, showing progress at the advanced beginner level.

Still, this may be a premature conclusion. The significant variation of the F0 cue is strictly interpretated on the basis of our statistical results. The F0 distinction is not as pronounced as that of native Korean speakers (cf. [13]). In addition, the magnitude of the vowel F0 distinction between lenis and fortis plosives fluctuates over vowel intervals depending on vowel types. Thus, future research on more advanced learners in similar contexts may offer further valuable evidence for understanding non-native learners' cue use in L2 speech learning.

5. Acknowledgement

This research was supported by the 2022 Korean Studies Grant Program of the Academy of Korean Studies (AKS-2022-R-061). The authors would like to thank Prof. Dr. Stephan Schmid and Prof. Dr. Eva Maria Luef for their valuable feedback on both the early and revised versions of this paper. We also extend our gratitude to two anonymous reviewers for their insightful comments and suggestions.

6. References

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

[2] Flege, J. E., "Second language speech learning: Theory, findings, and problems", in W. Strange [Ed], Speech perception and linguistic experience: Issues in cross-language research, 92, 233–277, Baltimore: York Press, 1995.

[3] Flege, J. E. and Bohn, O.-S., "The Revised Speech Learning Model (SLM-r)", in R. Wayland [Ed], Second Language Speech Learning: Theoretical and Empirical Progress, 3–83, Cambridge University Press, 2021.

[4] Best, C. T., "A direct realist view of cross-language speech perception", Speech Percept. Linguist. Exp., 171–204, 1995.

[5] Best, C. T. and Tyler, M. D., "Nonnative and second-language speech perception: Commonalities and complementarities", in O.-S. Bohn and M. J. Munro [Eds], Language Learning & Language Teaching, 17, 13–34, Amsterdam: John Benjamins Publishing Company, 2007.

[6] Choi, Y., "Intra- and intersegmental durational compensation of Korean plosives", in Proc. of the 20th Int. Congress of Phonetic Sciences (ICPhS), 2135–2139, 2023.

[7] Cho, T., Jun, S.-A. and Ladefoged, P., "Acoustic and aerodynamic correlates of Korean stops and fricatives", J. Phon., 30(2), 193–228, 2002.

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

[9] Jessen, M., Phonetics and phonology of tense and lax obstruents in German, 44, John Benjamins Publishing, 1998.

[10] Kuzla, C. and Ernestus, M., "Prosodic conditioning of phonetic detail in German plosives", J. Phon., 39(2), 143–155, 2011.

[11] Kim, M.-R. C., "Acoustic characteristics of Korean stops and perception of English stop consonants", PhD dissertation, The University of Wisconsin-Madison, 1994.

[12] Choi, J., Kim, S. and Cho, T., "An apparent-time study of an ongoing sound change in Seoul Korean: A prosodic account", PLOS ONE, 15(10), e0240682, 2020.

[13] Kang, Y., "Voice Onset Time merger and development of tonal contrast in Seoul Korean stops: A corpus study", J. Phon., 45, 76–90, 2014.

[14] Neuhauser, S., "Foreign Accent Imitation and Variation of VOT and Voicing in Plosives", in Proc. of the 17th Int. Congress of Phonetic Sciences (ICPhS), 1462–1465, 2011.

[15] Kleber, F., "VOT or quantity: What matters more for the voicing contrast in German regional varieties? Results from apparent-time analyses", J. Phon., 71, 468–486, 2018.

[16] Kirby, J., Kleber, F., Siddins, J. and Harrington, J., "Effects of prosodic prominence on obstruent-intrinsic F0 and VOT in German", in Proc. of the 10th Int. Conference on Speech Prosody, 210–214, 2020.

[17] Kingston, J., "Segmental influences on F0: Automatic or controlled?", in C. Gussenhoven and T. Riad [Eds], Tones and Tunes: Experimental Studies in Word and Sentence Prosody, 2, 171–210, Berlin, New York: Mouton de Gruyter, 2007.

[18] Seong, S.-H., "Phonological Transfer and its Hierarchy: L2 Perceptual Acquisition Process of Korean Plosives by German Native Speakers", J. Korean Lang. Educ., 16(3), 207–226, 2005.

[19] Kim, K.-H., Park, Y. and Chun, Y., "The Production and Perception of the Korean Stops by English Learners", Speech Sci., 13(4), 51–67, 2006.

[20] Holliday, J. J., "The perception and production of word-initial Korean stops by native speakers of Japanese", Lang. Speech, 62(3), 494–508, 2019.

[21] Holliday, J. J., "A longitudinal study of the second language acquisition of a three-way stop contrast", J. Phon., 50, 1–14, 2015.

[22] Whalen, D. H. and Levitt, A. G., "The universality of intrinsic F0 of vowels", J. Phon., 23(3), 349–366, 1995.

[23] Ladd, D. R. and Schmid, S., "Obstruent voicing effects on F0, but without voicing: Phonetic correlates of Swiss German lenis, fortis, and aspirated stops", J. Phon., 71, 229–248, 2018.

[24] Draxler, C. and Jänsch, K., "Speechrecorder - a universal platform independent multi-channel audio recording software", in Proc. of the 4th Int. Conference on Language Resources and Evaluation (LREC), 559–562, 2004.

[25] Yoon, T.-J., "Korean Forced Alignment System". Online: https://tutorial.tyoon.net/, accessed on 31 Mar 2024.

[26] Boersma, P. and Weenink, D., "Praat: doing phonetics by computer (version 6.3.16)". Online: http://www.praat.org/, accessed on 31 Aug 2023.

[27] R Core Team, "R: A language and environment for statistical computing (version 4.2.2)". Online: http://www.R- project.org/, accessed on 26 Dec 2022.

[28] Bates, D., Mächler, M., Bolker, B. and Walker, S., "Fitting Linear Mixed-Effects Models Using lme4", J. Stat. Softw., 67, 1–48, 2015.

[29] Kuznetsova, A., Brockhoff, P. B. and Christensen, R. H. B., "lmerTest package: tests in linear mixed effects models", J. Stat. Softw., 82(13), 2017.

[30] Lenth, R., "emmeans: Estimated marginal means, aka least-squares means. R package (version 1.8.3.)", 2022.

[31] Wood, S. N., Generalized Additive Models: An Introduction with R, Taylor & Francis, 2017.

[32] Rafat, Y., "Orthography-induced transfer in the production of English-speaking learners of Spanish", Lang. Learn. J., 44(2), 197–213, 2016.

[33] Ting, C., Clayards, M., Sonderegger, M. and McAuliffe, M., "The cross-linguistic distribution of vowel and consonant intrinsic F0 effects", PsyArXiv, 2023.

[34] Chang, C. B., "The acoustics of Korean fricatives revisited", Harv. Stud. Korean Linguist., 12, 137–150, 2008.

[1] The [+tense] feature of Korean fortis consonants is often indicated by the diacritic /*/ (e.g., in [7], [34]).