



## A study of L1 phonetic drift in the voice onset times of Korean learners of English with long L2 exposure\*

Mi-Ryoung Kim\*\*

*Department of Practical Foreign Languages, Korea Soongsil Cyber University, Seoul, Korea*

### Abstract

This study examines the voice onset times (VOTs) of Korean stops produced by Korean learners of English with high language proficiency and long L2 exposure (i.e., Korean-English bilinguals) to assess whether the VOTs of their lax and aspirated stops are merging and, if so, which types of stops are being changed. Thirteen Korean speakers (six female and seven male) who had studied in the USA for more than three to ten years participated. The results show that the speakers in this study with long L2 exposure are participating in the VOT merger, in which VOTs for aspirated stops are reduced while those for lax stops are increased. In other words, change in VOT affects not only aspirated stops but also lax stops. The results indicate that L1 phonetic drift may not be primarily affected by the amount of L2 exposure, and language contact may not be the primary factor triggering a sound change in the Korean stop system. Further study is necessary focusing on the phonetic shift of the “lax” category because it may play a pivotal role in a tonogenetic-like sound change in present-day Korean.

**Keywords:** phonetic drift, sound change, tonogenesis, Korean stops, VOT merger, Korean-English bilinguals, VOT

### 1. Introduction

Many studies have discussed that there are contact-induced diachronic sound changes (Flege, 1995, 2002; Sapir 1921; Weinreich, 1953). Sapir (1921:165) states, “...individual variations arising at linguistic borderlands—whether by the unconscious suggestive influence of foreign speech habits or by the actual transfer of foreign sounds into the speech of bilingual individuals—have gradually been incorporated into the phonetic drift of a language.” The term ‘phonetic drift’ refers to the L2-influenced phonetic change in an

individual’s L1 system. Several studies have shown that the L1 can, in fact, be affected by L2 learning. The effect of L2 on L1 speech is first discussed extensively in the work of Flege (1995, 2002). Flege’s findings suggest that the L1 system does not become static and invariable, but instead remains dynamic and ever-changing. For the effects of late second-language learning on first-language speech production, Chang (2012) shows that the effects can be found to be influenced by even brief experience with L2 Korean. His results suggest that not only advanced but also novice learners may have pervasive influence of second-language experience on first-language

\* This work was presented at Seoul International Conference on Speech Sciences (SICSS 2019), Seoul National University, Seoul, Korea.

\*\* kmrg@mail.kcu.ac, Corresponding author

Received 13 October 2019; Revised 20 November 2019; Accepted 29 November 2019

© Copyright 2019 Korean Society of Speech Sciences. This is an Open-Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/4.0>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

representations. If this is plausible, we may expect that learners with long L2 exposure may have more influences in an L1 system than those with no or short L2 exposure (see Kim, 2014, 2017).

Recent studies show that, for Seoul Korean speakers, Korean stops undergo a tonogenetic-like sound change in terms of voice onset time (VOT) and fundamental frequency (f<sub>0</sub>): the differences of VOT between aspirated and lax stops are being reduced while f<sub>0</sub> differences between the two stops are being maximized. As a result, the well-known minimal pair can be distinguished not by the traditional consonantal cues (VOT and others) but by fundamental frequency (f<sub>0</sub> or tone) (e.g., [t<sup>h</sup>ál] ‘mask’ vs. [t<sup>h</sup>ál] ‘moon’). The phonetic shift (or cue weighting) from VOT to f<sub>0</sub> has been discussed in many studies about tonogenesis in Korean (Bang et al., 2018; Kang, 2014; Kang & Han, 2013; Kim, 2000; Kim et al., 2002; Silva, 2006) and that in other languages (Beddor, 2009; Coetzee et al., 2014; Kirby, 2014; Matisoff, 1973). Concerning the change of VOT, there are two controversial issues: one is which stops truly undergo a change and the other is whether language contact plays a role triggering a sound change.

For the phonetic drift (or change) of Korean stops, there is a consensus that the VOT differences between lax and aspirated stops are reduced. However, there seems to be a controversy that both stops undergo a change. Most of studies have argued that, among the three stops in Korean, change in VOT affects only aspirated stops (i.e., VOT reduction from aspirated stops), rather than lax stops or both stop classes (Bang et al., 2018; Kang, 2014; Silva, 2006). A couple of studies have shown that not only aspirated stops but also lax stops undergo a sound change in which VOTs in aspirated stops are being decreased whereas those in lax stops are being increased (Kim, 2014, 2017). Because of different findings, it is not clear which stops truly undergo a sound change. Much attention was given to the change of aspirated stops, instead of that of lax counterparts. Furthermore, the role of the “lax” category has been neglected in a tonogenetic-like sound change in Korean. In this study, the issues of which stops undergo a change and how the change of the lax category is related with tonogenesis are discussed.

With respect to the role of language contact in sound change, a number of studies have shown that the sound change mainly occurs in Seoul Korean, the standard dialect of Korean, but not in non-Seoul Korean dialects (Choi, 2002; Kim, 2014; Lee & Jongman, 2012). Some studies have shown that the change can also occur to Korean speakers residing abroad, or heritage Korean speakers in Shenyang or Toronto (Jin, 2008; Kang & Guion, 2008; Kang & Nagy, 2016; Oh & Yang, 2013). As for the influence of L2 English on L1 Korean, Kim (2011) reports that the merger phenomenon in the L1 system is more robust for L2 advanced speakers than for L2 beginners because of the amount of L2 experience. If this is true, some language contact or L2 experience can be the primary external factor that triggers the sound change in the L1 system. Thus, it might be worth examining whether VOT drift also occurs for Korean learners of English with long L2 exposure.

### 1.1. Background of Korean Stops

Languages in the world are common to have a voicing contrast in stops. The phonological voicing contrast is phonetically realized as prevoiced, voiceless unaspirated, and voiceless aspirated stops. In order to separate the stops categories of a number of languages, Lisker & Abramson (1964) introduced three stages in VOT, voicing lead (negative VOT) for prevoiced stops, short-lag voicing (positive

VOT) for voiceless unaspirated stops, and long-lag voicing (positive VOT) for voiceless aspirated stops. His classification accounted well for two- to four-category languages. When stops have a three-way contrast, they are categorized mostly in terms of voicing and aspiration: “voiced”, “voiceless unaspirated” or “aspirated” in the languages of the world (e.g., Keating, 1984; Ladefoged & Maddieson, 1996; Lisker & Abramson, 1964). Korean stops, however, are known to be cross-linguistically atypical in that they have a three-way contrast where they are all voiceless in phrase-initial position. VOT has been well-documented as a primary cue to contrast stops across the world’s languages. However, it is not good enough to separate the three Korean categories (Cho et al., 2002; Han & Weitzman, 1970; Lisker & Abramson, 1964).

Different from the world’s languages, Korean is the only language that does not have a voicing contrast and that reportedly has three kinds of voiceless stops. They are usually described as aspirated /p<sup>h</sup> t<sup>h</sup> k<sup>h</sup>/, tense or fortis /p\* t\* k\*/, and lax or lenis /p t k/. All Korean stops are phonetically voiceless, except that lax stops voice intervocalically under certain prosodic conditions (Jun, 1993). In addition, according to the traditional approach, there is a consonant-tone correlation in that fortis consonants are correlated with High tone whereas lenis consonants are correlated with Low tone. This kind of correlation is not common but Korean-specific. In addition, the phonetic behaviors of the voiceless lax stop are very hard to be accounted for why they become heavily aspirated and how they obtain a Low tone. In addition, the early acquisition of the tense stop is also a mystery. Along with the tradition approaches, so many Korean-specific phonetics and phonology are required but may issues still remain unsolved. Arguing against that Korean does have an unusual phonology, Kim & Duanmu (2004) propose a voicing-aspiration system where underlyingly “tense” stops are regular voiceless unaspirated, and “lax” stops are regular voiced stops (see also Kim, 2000). With this proposal, a number of issues are resolved: the consonant-tone correlation is another case of voiceless-high and voiced-low and the phonetic behaviors of lax and tense stops are also accounted for (see Bang et al., 2018; Kang, 2014; Silva, 2006 for different proposals). Under the theory of standard tonogenesis, /d/ becomes [t<sup>h</sup>] leaving a low tone (e.g., [t<sup>h</sup>ál] ‘mask’ vs. [t<sup>h</sup>ál] ‘moon’). This explanation can account for why lax stops are correlated with low tones and why they become heavily aspirated (see Kim, 2014 for more discussion). Therefore, to account for a tonogenetic-like sound change in Korean, the role of the “lax” category is crucial. The issue of how Korean stops and tones should be analyzed still remains controversial. Whatever phonological assumptions are, it is very important to pay more attention to the role of the lax stop in sound change.

### 1.2. VOT Drift: A Diachronic Perspective

Numerous investigations have explored the precise phonetic characteristics of the three Korean stops. Among them, the summary of mean VOT values for each stop type is presented in Table 1 where, on average, VOTs are shortest (about 10-15 ms) for Korean tense stops, longer (30-71 ms) for lax, and longest (68-115 ms) for aspirated. At the same time, despite systematic mean differences, VOT values for individual tokens of lax and tense stops, and to a lesser extent lax and aspirated stops, overlap (Cho et al., 2002; Han & Weitzman, 1970; Kim, 1965; Kim, 1995; Lisker & Abramson, 1964). This trend still holds for present-day Korean stops except that, instead of the overlap between the lax and tense stop, there is a

complete or partial VOT merger between the lax and aspirated (Kang & Guion, 2008; Kim, 2014, 2017; Oh, 2011; Silva, 2006). The diachronic change on VOT shift is presented well in Table 1 in which the mean VOT difference between the lax and aspirated stop ( $VOT_{asp-lax}$ ) has decreased from 77 ms to 5 ms.

**Table 1.** Mean VOTs (ms) and their ranges in parentheses of the three stops in Korean from 1964's to 2017's findings

	Tense	Lax	Aspirated	$VOT_{asp-lax}$
Lisker & Abramson (1964)	12 (0–35)	30 (10–65)	103 (65–200)	73
Kim (1965)	12 (9–15)	35 (23–45)	93 (90–98)	58
Han & Weitzman (1970)	7 (0–52)	30 (10–80)	107 (40–148)	77
Kim (1995)	11 (7–14)	64 (31–108)	115 (88–158)	51
Cho et al. (2002)	21	58	113	55
Kang & Guion (2008)	11	63	68	5
Oh (2011)	15 (9–24)	52 (32–75)	70 (39–97)	18
Kim (2014)	15 (5–48)	71 (14–146)	83 (28–145)	12
Kim (2017)	13 (10–15)	69 (14–153)	79 (29–143)	10

Table 1 presents that, overall, VOTs for tense stops are stable from 11 ms to 21 ms whereas those for lax and aspirated stops are not. VOTs for the aspirated category are getting shorter from 115 ms to 70 ms. In contrast, VOTs for the lax category in recent studies has become twice as long as before (30 ms vs. 71 ms).

### 1.3. VOT Variations: A Synchronic Perspective

The VOT shift for Korean stops has been synchronically documented in terms of various socio-phonetic factors such as age, gender, dialect, speech rate, L2 proficiency, and word-frequency (Bang et al., 2018 for word-frequency; Choi, 2002; Cho et al., 2002; Kang & Guion, 2008; Kim, 2011 for L2 proficiency; Lee & Jongman, 2012 for dialect; Oh, 2009 for speech rate; Oh, 2011 for gender; Silva, 2006 for age). For the factor “age”, this change is more advanced in younger speakers’ speech than in older speakers’ speech. Since Silva (2006), numerous studies have shown that older speakers tend to maintain a three-way VOT categorization (tense < lax < aspirated) whereas younger speakers tend to show a two-way VOT categorization (tense < lax = aspirated) (see also Kang & Guion, 2008 and among others). With respect to the effect of gender on VOT, the VOT merger is more robust in female speakers’ speech than in male speakers’ speech (Kang & Nagy, 2016; Kim, 2014, 2017; Oh, 2011; and many others). For the effect of dialect on VOT, the merger is more advanced in Seoul dialect than in non-Seoul dialects (Choi, 2002; Kang & Han, 2013; Lee & Jongman, 2012; Oh & Yang, 2013). Taken all aforementioned factors into consideration together, young female Seoul speakers are found to be the most active adopters or propagators in sound change. Besides the effect of age, gender, and dialect, it is also reported that the change is more advanced in proficient L2 English speakers than non-proficient speakers (Kim, 2011), more advanced in more fast speech than in slow speech (Oh, 2009), and more advanced in frequent words than in non-frequent words (Bang et al., 2018). Compared with the factors of age, gender, and dialect, however, these factors are not stable because of speaker variations. Kim (2017) reports that there

are huge inter- and intra-speaker variations on VOT that the VOT differences were remarkably huge even among young male and female Seoul speakers in the same generation. This suggests that VOTs for lax and aspirated stops are not stable and are synchronically varied.

With respect to the VOT merger in sound change in Korean, this is a follow-up experiment of Kim’s (2014, 2017) with different speakers in which native Korean speakers with no L2 exposure participated in Kim (2014) and those with short L2 exposure participated in Kim (2017). This study exclusively examines VOTs of Korean stops produced by Korean learners of English with high language proficiency and long L2 exposure to assess whether their VOTs between lax and aspirated stops are being merged and, if any, what types of stops are being changed. The results of the present study are compared with those of Kim’s (2014, 2017).

## 2. Method

For a direct comparison regarding a sound change in Korean, the same methodology as Kim (2014, 2017) was used in this study.

### 2.1. Participants

Thirteen Korean speakers (seven male and six female), studying at a university in the United states, participated. At the time of testing, they all had resided in the mid-western area of the U.S.A approximately for more than three years to ten years as presented in Table 2.

**Table 2.** Speakers’ language background

Spks.	Gender	Year	AOA	Profi.	Dialect
THK	F1	3	21	High	Gyeonggi
GEK	F2	8.3	18	High	Gyeonggi
SJK	F3	3	21	High	Seoul
ETK	F4	9	21	High	Seoul
WKC	F5	7	20	High	Seoul
JNC	F6	6	13	High	Seoul
DRK	M1	10	18	High	Seoul
SMK	M2	5	21	High	Seoul
JHP	M3	9	17	High	Seoul
JHB	M4	6	16	High	Seoul
SHO	M5	7.5	15	High	Seoul
JSY	M6	8	16	High	Seoul
DKC	M7	3	22	High	Seoul
Mean		6.5	18.4		

Year, residence in America; AOA, age of arrival; profi., proficiency).

They were all university students at a college. The mean age was 23 years, 7 months and the individuals ranged from 20 years, 5 months to 33 years, 6 months. Since most of them were born and had grown up in Seoul (the capital city of Korea) or the Gyeonggi area (near Seoul), they were purely native Korean speakers and used the standard dialect of Seoul Korean. They came to the United States after their puberty and one female speaker who came to the US at the age of 13. According to their self-reported questionnaires and interviews, they all were very fluent in English and evaluated as advanced Korean learners of L2 English (i.e., High proficiency). They could communicate with native speakers of English with confidence and comfort. None of speakers had any history of speech pathology or phonetic training.

## 2.2. Speech Materials

For a comparative purpose, the same speech materials as in Kim (2014) were used in this study. Eighteen monosyllabic Korean words were constructed with a balanced list of the three stop types (lax, aspirated, and tense) and the three places of articulations (labial, alveolar, and velar) followed by a vowel /a/ context. The syllable type was either CV or CVC where the final coda was /t/ or /k/ (a neutralized unreleased stop). All words were real words, as presented in Table 3.

Table 3. Target words

Type	Stops	Bilabial	Alveolar	Velar
CV(C)	Asp.	[p <sup>h</sup> at] ‘red bean’ [p <sup>h</sup> a] ‘to dig’	[t <sup>h</sup> at] ‘blame’ [t <sup>h</sup> a] ‘to ride’	[k <sup>h</sup> at] ‘stop’ [k <sup>h</sup> a] ‘car’
	Lax	[pat] ‘field’ [pa] ‘to see’	[tat] ‘anchor’ [ta] ‘all’	[kat] ‘hat’ [ka] ‘to go’
	Tense	[p*ak] ‘head’ [p*a] ‘to grind’	[t*ak] ‘precisely’ [t*a] ‘to pick’	[k*ak] ‘croak’ [k*a] ‘to peel’

Tense stops=/p\* t\* k\*/, lax or lenis=/p t k/, and aspirated=/p<sup>h</sup> t<sup>h</sup> k<sup>h</sup>/.

In order to record the target words in a constant prosodic environment, they were recorded in the carrier sentence [igo \_\_\_\_\_ hasejo] “Say this \_\_\_\_\_” where the words were located in phrase-initial position. In the position, the target words were expected to be fully stressed and emphasized. Speech materials were presented in *Hangul*, the writing system of Korean.

## 2.3. Procedure

Each of the speakers was recorded using a Shure (model SM 10A) head-mounted microphone and a Marantz digital recorder (PMD 670). The recordings were made in a sound-attenuated booth in the Phonetics Lab at a sampling rate of 44,100 Hz. The target words in carrier sentence were presented on a monitor in quasi-random order and were automatically popped up at a 3-second interval to control participants’ speech rate. The monitor connected to the computer was inside the lab but the computer itself was outside the lab to minimize background noise. Each speaker familiarized sentences and was asked to read the target words in a natural intonation. A total of 936 tokens (18 words×4 repetitions×13 speakers) were obtained for analysis. All recorded utterances were analyzed using Praat 6.0.37, a speech analysis program (Boersma & Weenink, 2018).

VOT was measured from the release burst of a stop to the onset of periodicity in the waveform (Lisker & Abramson, 1964). The onset of the vowel in the waveform was determined by the onset of the first full glottal pulse of the vowel as well as the pitch of the spectrogram. The onset of the voicing energy in the second formant shown in a time-locked spectrogram was used to help determine voicing onset in conjunction with the waveform. The onset of voicing (=vowel onset) was defined as the first and periodic pulse of a vocalic waveform that show features typical of a vowel.

The acoustic data obtained were statistically tested using repeated measures analysis of variance (RM ANOVA) in the context of a univariate context of a general linear model (SPSS/PASW, 2018). The main goal of the present study was to examine whether there is any main effect of stop type (i.e., aspirated, lax, and tense), place (labial, alveolar, and velar), gender (male vs. female), LOR, and individual speakers on VOT and whether there was any interaction effect between factors. Repeated measures ANOVA includes both

“between” subjects effects (gender and speaker) and “within” subject effects (i.e., stop type and place). Their main and interaction effects were statistically analyzed at a 0.05 significant level. Post hoc tests were also run to answer the following questions: (i) whether any differences in pairs among three stop types, between gender, and places were significant, and (ii) whether any differences in pairs among individual speakers were significant. For statistical analysis, the VOT differences and speaker variability on sound change according to the three stop types were mainly focused.

## 3. Results

### 3.1. Effects of Stop, Place, and Gender on VOT

For the pool data averaged across thirteen participants, results of univariate repeated measures (RM) ANOVA showed a main effect of stop on VOT [ $F(2, 933)=1,228.761, p<.0001$ ]. Post hoc Tukey HSD multiple comparisons revealed that differences in each pair among the three stop types were statistically significant ( $p<.0001$ ): VOT is longest (80 ms) for the aspirated stop, longer (64 ms) for the lax stop, and shortest (15 ms) for the tense stop (tense<lax<aspirated) (Figure 1).

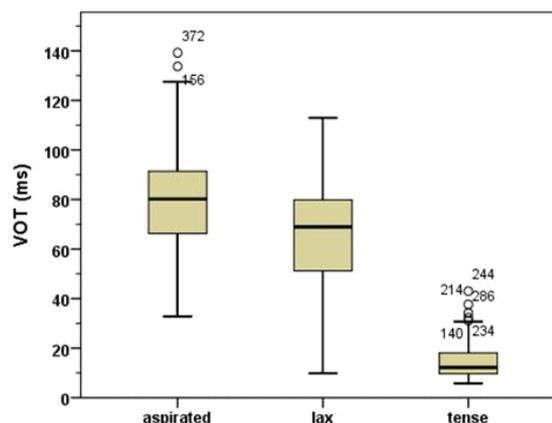


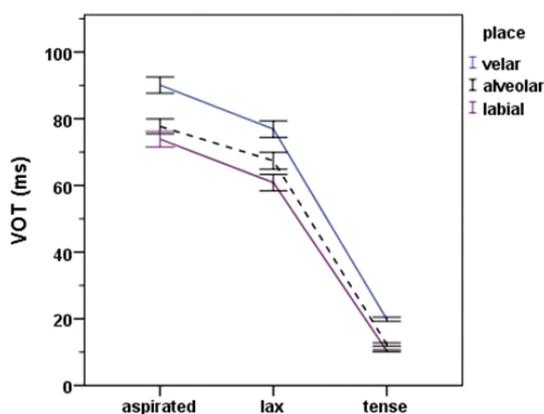
Figure 1. Mean VOT values of the aspirated, lax, and tense stop for the productions of 13 speakers (7 males and 6 females) with long L2 exposure for more than three years. Error bars indicate 95% confidence intervals.

Overall, the pooled results in Figure 1 correspond to Kim’s (2014, 2017) and previous findings in that, although the three stops statistically differed in terms of VOT, their VOT values between the aspirated and lax stop were largely overlapped in their ranges. The mean difference between the aspirated and the lax stop is getting shorter, as reported in Kang (2014), Kim (2014), Silva (2006), and among many others, supporting that stops in Korean undergo a sound change as an aspirated-lax merger on VOT. However, the present results are comparable to earlier findings on whether change in VOT affects only aspirated stops (i.e., VOT reduction from aspirated stops), rather than lax stops or both stop classes (Bang et al., 2018; Kang, 2014; Silva, 2006).

Silva (2006: 294) reports, “For younger speakers, the ‘innovators’, the mean VOT associated with tense and lax stops is nearly equal to that for the traditionalists, but the mean VOT for aspirated stops is markedly lower (69.7 ms vs. 94.0 ms).” His findings suggest that aspirated stops only undergo a VOT change. Bang et al., (2018: 137) also discuss, “During this sound change, VOT ‘reduction’ (over

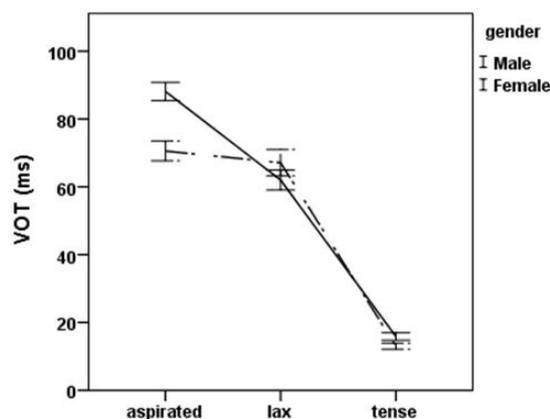
time) comes largely from aspirated stops, while lax stops show little change.” Different from their findings, however, Kim (2008, 2014, 2017) shows that not only aspirated stops but also lax stops undergo a sound change in terms of VOT. The present findings correspond to Kim’s in that both stops under a VOT change or shift: aspirated stops undergo a shortening process (115 ms to 80 ms) whereas lax stops undergo a lengthening process (30 ms to 64 ms) (see Table 1). That is to say, the change occurs for both aspirated and lax stops.

The effect of stop on VOT held well for that of place, as see in Figure 2. There was a main effect of place of articulation on VOT [ $F(2, 918)=95.829, p<.0001$ ]. Post hoc Tukey HSD multiple comparisons showed that mean VOT values averaged across the three stops are significantly longest (62 ms) for the velar stop, intermediate (51 ms) for the alveolar stop, and shortest (46 ms) for the labial stop (labial<alveolar<velar) (Cho et al., 2002). There was also an interaction between stop type and place of articulation [ $F(4, 918)=3.842, p<.0001$ ], indicating that the effect held well for each stop type. For each place of articulation, mean VOT values were longest for the aspirated, longer for the lax stop, and shortest for the tense stop, as clearly illustrated in the figure below.



**Figure 2.** Mean VOT values of the aspirated, lax, and tense stop for the productions of 7 male and 6 female speakers. Error bars indicate 95% confidence intervals.

The VOT shift does not occur for all speakers’ speech materials as clearly shown below. There was also a main effect of gender on VOT [ $F(1, 930)=22.650, p<.0001$ ] and an interaction between stop type and gender [ $F(2,930)=38.132, p<.0001$ ] on VOT. Mean VOTs for aspirated stops are significantly longer for male speakers than female speakers. In addition, the VOT merger between lax and aspirated stops are robust for female speakers but not for male speakers, as can be seen in Figure 3 (see also Figure 4).



**Figure 3.** Mean VOT values of the aspirated, lax, and tense stop for the productions of male and female speakers. Error bars indicate 95% confidence intervals.

The effects of gender on VOT are analogous to Oh (2011) in that the VOT merger between lax and aspirated stops are much greater for female speakers than for male speakers. Out of six female, five speakers showed a statistical merger between lax and aspirated stops whereas, out of seven male, none of speakers showed a merger. Note that most of speakers are under or in their twenties. This suggests that the socio-phonetic factor “gender” can be more important than the factor “age.” This gender effect corresponds to that of Kim’s (2014, 2017). The results of stop, place, and gender on VOT are summarized in Table 4.

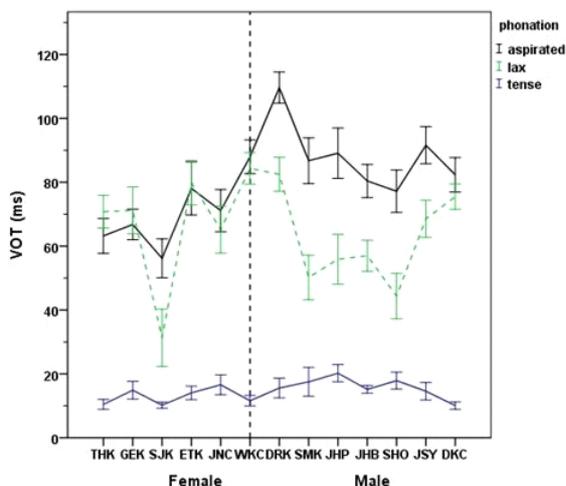
**Table 4.** Mean VOTs with standard deviations (in parentheses) by stop, place, and gender averaged over the six female and seven male speakers of Korean

Gender	Place	Aspirated	Lax	Tense
Female	Labial	62 (15)	56 (21)	9 (2)
	Alveolar	70 (17)	69 (23)	11 (3)
	Velar	79 (17)	77 (21)	18 (5)
	<b>Average</b>	71 (18)	67 (24)	13 (5)
Male	Labial	81 (14)	54 (18)	12 (4)
	Alveolar	83 (16)	58 (21)	13 (4)
	Velar	100 (16)	73 (13)	22 (7)
	<b>Average</b>	88 (18)	62 (19)	15 (7)
Mean		80(20)	64 (21)	15 (7)

Overall, the group-normative data in Table 4 show that, despite systematic mean differences among the three stops in statistics, VOT differences between lax and aspirated stops are largely decreased showing a merger between the two stops. With respect to VOT, tense stops are well separated from either aspirated or lax stops whereas aspirated stops are not distinguished well from lax stops. Tense stops are always realized with short-lag voicing ( $M_{tense}=15$  ms (5.8 min.–43 max.)) while aspirated and lax stops both are realized with long-lag voicing ( $M_{asp}=80$  ms (32.8 min.–139.2 max.);  $M_{lax}=64$  ms (9.7 min.–113 max.)). The results on the effect of gender on VOT suggest that an apparent merger between aspirated and lax stops is still in progress among female speakers’ productions. Concerning the sound change of the Korean system, the results on VOT aggregated by stop, place, and gender partly correspond to previous findings such as Silva (2006), Oh (2011), Kim (2014, 2017). However, there were inter- and intra-speaker variations on VOT, as discussed below.

### 3.2. Effects of Individual Speakers on VOT

Considering the remarkable gender differences, it is especially worth considering individual speakers' data. Figure 4 shows mean VOT values of the three stops for each speaker separately. The huge individual variations correspond to Kim's (2017). A rather obvious point is that the absolute values of VOT vary with speakers, showing individual speaker variations. Statistical analysis confirmed significant differences among individual speakers [ $F(12, 897)=32.788, p<.0001$ ] and an interaction between stop and individual speakers on VOT [ $F(24, 897)=17.126, p<.0001$ ].



**Figure 4.** Mean VOT values of each stop type for the productions of 6 female (left) and 7 male (right) speakers. Error bars indicate 95% confidence intervals.

In Figure 4, for all speakers, among the three stops, the productions of tense stops are very stable whereas the productions of both aspirated and lax stops are not. There are remarkable individual differences on VOT, heavily dependent on the factor “gender.” There are also intra-speaker variations on VOT even within the same gender. Note that, out of six, one female speaker SJK only shows a clear-cut three way stop system. In contrast, out of seven, only one male speaker, DKC, shows a merger-like pattern of aspirated and lax stop. Synchronic variations on sound change are robust between males and females: male speakers maintain a traditional three-way contrast on VOT while female speakers do not. Compared with male speakers, female speakers tend to produce shorter VOT for aspirated stops but longer VOT for lax stops as evidence of a sound change. This shortening and lengthening process made the two stops be merged. The merger is statically supported, as presented in Table 5.

**Table 5.** Mean VOTs (ranges in parentheses) and statistical groupings for individual speakers

Spkr.	VOT <sub>aspirated</sub>	VOT <sub>lax</sub>	VOT <sub>asp-lax</sub>	Merger
THK (F1)	63 (41–85)	71 (48–98)	–8	YES
GEK (F2)	67 (39–88)	71 (39–99)	–4	YES
SJK (F3)	56 (33–82)	31 (9.7–82)	25	NO
ETK (F4)	78 (41–126)	80 (32–113)	2	YES
WKC (F5)	88 (66–112)	84 (59–110)	4	YES
JNC (F6)	71 (47–112)	65 (31–97)	–6	YES
DRK (M1)	110 (92–134)	82 (55–111)	28	NO
SMK (M2)	87 (58–121)	50 (25–79)	37	NO
JHP (M3)	89 (55–139)	56 (17–80)	33	NO
JHB (M4)	81 (54–104)	57 (38–77)	24	NO
SHO (M5)	77 (53–113)	44 (20–78)	33	NO
JSY (M6)	92 (73–115)	69 (40–89)	23	NO
DKC (M7)	82 (62–128)	75 (55–94)	7	NO
<b>Mean</b>	<b>80 (33–139)</b>	<b>64 (9.7–113)</b>		

Taken the individual data in Figure 4 and Table 5 into consideration together, VOTs in the production of the three Korean stops are different from one speaker to the other in the following three aspects. Firstly, does a speaker show an aspirated-lax merger or not? In Table 5, out of thirteen, five speakers statistically have the merger and they all are females (F1, F2, F4, F5, F6). For speaker F1, F2, and F6, they produced lax stops even longer than aspirated counterparts. We can say that these participants, who show the complete merger, are considered as “radical innovators” in sound change. Compared to the female individuals, male speakers tend to maintain a three way contrast in terms of VOT. Thus, they are “conservative traditionalists” in sound change (see also Oh, 2011). From Kim's (2014) findings, we can expect that these male speakers do not develop a tonal contrast in the vowel. Overall, female speakers truly undergo a sound change but male speakers do not.

Secondly, individuals can differ which stops undergo a phonetic change. Consider the reference point from previous findings on VOT. Lisker & Abramson (1964) report that mean VOT for the aspirated stop is 103 ms ( $M_{\text{aspirated}}=103$  ms) and mean VOT for the lax stop is 30 ms ( $M_{\text{lax}}=30$  ms). When comparing the current results with Lisker & Abramson's, it is clear that overall mean VOT for the aspirated stop is much shorter from 103 ms to 80 ms whereas mean VOT for the lax stop is much longer from 30 ms to 64 ms.

For aspirated stops, there is a shortening process. As presented in Table 6, speaker F1, F2, F3, F4, F6, and M5 produced aspirated stops with relatively shorter VOTs than the rest speakers. Note that speaker M1, M2, M3, and M6 maintain relatively long VOTs. In contrast, there is a lengthening process for the lax stop. Speaker F1, F2, F4, M1, and M7 produced lax stops with relatively longer VOTs than the rest speakers. For these speakers, the merger did not occur because, unlike lax stops, aspirated stops did not undergo a sound change. To sum up, taking group mean VOTs (80 ms) as a reference point in consideration, six out of thirteen speakers (e.g., F1, F2, F3, F4, F6, and M5) show a shortening process on VOT for the aspirated stop. In contrast, eight out of thirteen speakers (e.g., F1, F2, F4, F6, M1, M6, and M7) produced lax stops with relatively longer VOTs (above group mean 64 ms). Four speakers undergo both lengthening and shortening processes.

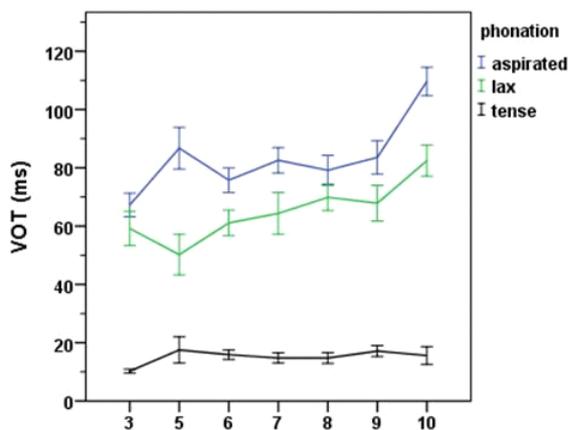
Once stops undergo a sound change, the process of aspirated-shortening and/or lax-lengthening seems to occur simultaneously. When there is only one process, the merger does not occur. For example, female speaker F3 had an aspirated-shortening process (56 ms) but she did not have a lax-lengthening process (31 ms). Unless a

lax-lengthening process occurs, she will tend to preserve a contrast instead of a merger. Similarly, speaker M1 took part in the lax-lengthening process (82 ms) but he did not participate in the aspirated-shortening process (110 ms). As a result, his stop system did not undergo a merger. It is very important to capture the essence of the sound change in the Korean stop system. Because of this shortening or lengthening process among the three stops, the overall VOT duration of the three stops is different from one speaker to the other. Post hoc Tukey multiple comparisons among the three stops showed three statistical groupings in duration, as presented in Table 6.

**Table 6.** Tukey HSD statistical subsets of participants for the overall VOT duration averaged across 936 tokens

Spkr.	N	1	2	3
SJK (F3)	72	32.565		
SHO (M5)	72	46.486	46.486	
THK (F1)	72	48.146	48.146	
JHB (M4)	72		50.853	
GEK (F2)	72		50.960	
JNC (F6)	72		50.969	
SMK (M2)	72		51.491	
JHP (M3)	72		55.059	55.059
DKC (M7)	72		55.950	55.950
ETK (F4)	72		57.317	57.317
JSY (M6)	72		58.238	58.238
WKC (F5)	72		61.296	61.296
DRK (M1)	72			69.233
Sig.		.151	.213	.276

Thirdly, individuals' productions on VOT are slightly different depending on their residential years. Figure 5 shows mean VOT values of the three stops in terms of speakers' L2 exposure period.



**Figure 5.** Mean VOT values of each stop type according to L2 exposure period from 3 to 10 years. Error bars indicate 95% confidence intervals.

There was a main effect of L2 exposure on VOT [ $F(6, 915)=23.116, p<0.0001$ ] and an interaction between stop and L2 exposure on VOT [ $F(12, 916)=7.276, p<0.0001$ ]. Post hoc Tukey multiple comparisons among the three stops revealed two statistical subsets, as presented in Table 7. The overall VOT is significantly longer for a 10-year participant than the rest ones.

**Table 7.** Tukey HSD statistical subsets of L2 exposure for the production of stops

Year	N (s)	1	2
3 year	216 <sub>(3)</sub>	45.554	
6 year	144 <sub>(2)</sub>	50.991	
5 year	72 <sub>(1)</sub>	51.491	
7 year	144 <sub>(2)</sub>	53.891	
8 year	144 <sub>(2)</sub>	54.599	
9 year	144 <sub>(2)</sub>	56.188	
10 year	72 <sub>(1)</sub>		69.233
Sig.		.158	1.000

N, a total number of speech materials; s, number of speakers.

Taken Figure 5 and Table 7 together, individuals may differ from one to the other depending on their L2 exposure period. Among speakers, one five- and ten-year speaker conservatively maintain the three-way pattern. All the other speakers tend to undergo a phonetic merger. Recall that one of the issues in the current study is whether there are any phonetic changes in one's native language due to the influence of another language (L2). Based on the results in Figure 5, it seems to hard to tell that the phonetic change (in particular VOT) in the L1 Korean system is purely due to the influence of L2 English. The current results on the effect of L2 exposure correspond to Kim's (2017) findings. If there is some influence of language contact in the L1 system, we may expect that the longer their L2 exposure, the more the merger. Note that a ten-year speaker tend to maintain the most clear-cut three way stop contrast in VOT. The effect of L2 exposure on VOT seems to be marginal and linguistically unimportant. Since the number of participants were neither balanced nor enough to be more than three for each year, it is hard to obtain a correlation effect between stop and L2 exposure. Further study may be necessary to make the claim.

Finally, individuals differently produced the same tokens. For example, speaker F3 produced stops inconsistently. For the token [pat] 'field' of a lax stop, she repeated it in the range of 9 ms to 80 ms. The inconsistent VOT values are due to her recording style of target words. When VOT became much shorter than normal, the target words were produced without some pause between the carrier sentence. She often located them in a phrase-medial position rather than a phrase-initial position. Remarkable inter-speaker differences show some gender dependency. Similar to Oh's (2011) and Kim's (2014, 2017) findings, female speakers tend to show more merger than male speakers, as illustrated in Figure 3 and 4. Let us compare the results speaker THK (F1) with those of speaker SMK (M2). Except for gender, their language background were almost the same: They both were graduate students studying in America. They were in their twenties (23 vs. 20) and they grew up in Seoul (i.e., same dialect). In addition, they had lived in the USA for about three years (same length of residence). In spite of their similar language background, their productions of Korean stops were remarkably different: speaker F1 showed a complete merger while speaker M2 didn't. The female speaker undergoes a sound change whereas the male speaker is not. Similar phenomena can be observed from other speakers' productions.

Taken together, inter-speaker and intra-speaker differences are shown to be spectacular. The within-speaker variability in females appears to be high. Note that speakers employed in the present study were from the same Seoul (including Gyeonggi) dialects and the same generations in their twenties and thirties. In addition, their L2 exposure period was controlled between three to ten years, and

speech rate was also controlled. Despite the fact that such socio-phonetic factors as age, dialect, L2 exposure, and even speech rate were controlled, there were huge inter-speaker and intra-speaker variations on the merger and duration, indicating a synchronous speech sound change. It seems clear that VOTs for the production of stops are internally changing for each speaker. Hence, we can say that individual speakers play an important role in taking a part in sound change of their native system.

#### 4. Discussion and Conclusions

This study was designed to examine VOT of three Korean stops produced by advanced Korean learners of English with long L2 exposure. With respect to the VOT shift for speakers employed in this study, the present results replicate Kim's (2014, 2017) findings in the following aspects: firstly, the VOT differences between aspirated and lax stops are being reduced toward a consonantal merger (Silva, 2006 and many others). Secondly, the consonantal merger is mostly led by female speakers (Oh, 2011 and many others). Thirdly, change in VOT affects not only aspirated stops but also lax stops (Kim, 2014). Fourthly, there are huge inter-speaker and intra-speaker variations on VOT, especially between male and female speakers (Kim, 2017).

With respect to the research question on which stops truly undergo a sound change, the present findings are inconsistent with Bang et al.'s (2018), Kang (2014), and Silva (2006). They reported that, during the sound change, VOT 'reduction' (over time) comes largely aspirated stops, while lax stops show little change. Instead, the present findings replicate Kim's (2014, 2017) where both aspirated and lax stops truly undergo a sound change. The change is not uni-directional but bi-directional, called the aspirated-shortening process and the lax-lengthening process. The diachronic change can be synchronically observed in terms of various socio-phonetic factors such as age, gender, dialect and so on.

The present results also replicate Oh's (2011) findings that the sound change is mostly led by female speakers (i.e., the effect of gender on VOT shift). When 'age' is controlled, we can say that 'gender' is the most prominent factor that varies the sound change. Labov (1994) hypothesizes, "if socio-phonetic factors are working well, young female speakers are leaders in sound change." As seen in Figure 3 and 4, the fact that the sound change showed a strong gender dependency is robust. Almost all female speakers (five out of six) show a complete merger whereas none of male speakers do not. The present findings support Labov's (1994) assertion that women rather than men play a leading and active role in sound change.

However, the current results do not replicate Kim's (2011) where speakers with L2 proficiency in English lead over those without. She reports that the VOT shortening for aspirated stops in the L1 system is closely related with the L2 English influence: more proficient speakers tend to produce shorter for L1 Korean aspirated stops than less proficient speakers. If there are large influences of L2 English on L1 change, we may obtain different results. The current findings are also different from Chang's (2012) rapid and multifaceted effects of second-language learning on first-language speech production. In contrast to his results that novice learners have pervasive influence of second-language experience on first-language representations, advanced learners with long L2 exposure do not show much influence of L2 on the L1 system. With respect to the VOT change, the present findings suggest that there is little

relationship between L1 change and L2 influence. In other words, the L1 system was not largely affected by L2 exposure or L2 proficiency. This study suggests that the effect of L2 English on L1 Korean can be marginal.

In addition, the new findings of the study are that, regardless of the residential L2 period or L2 proficiency, some advanced Korean learners of English are actively participating in changing their pronunciation for aspirated and lax stops. Five out of thirteen show the statistical merger between the two stops in which almost all female speakers show the merger but all male speakers do not. With respect to the aspirated-lax merging phenomenon as evidence of a sound change, there is little relationship between L1 change and L2 influence that some male speakers with long L2 exposure tend to maintain clear VOT differences between lax and aspirated stops whereas some female speakers with long L2 exposure tend to show the merger. This indicates that the merger may not be closely related with the amount of L2 exposure or the level of L2 proficiency but with gender, as clearly illustrated in Figure 4. However, since a number of speakers were not balanced for each year and were not enough to be compared in terms of L2 exposure, further study is necessary to generalize.

In the current study, certain socio-phonetic factors such as age, dialect, gender, and speech rate that might influence VOT were controlled to minimize any effects as much as possible. Even with controlled speakers' language background, some speakers made a clear lax and aspirated distinction whereas others showed the VOT merger of the two stops. Some speakers had relatively longer VOTs for both stops than others. Remarkable inter-speaker and intra-speaker variability suggest that speakers themselves are internally changing their pronunciations. Thus, some speakers may be earlier adopters or more active propagators of the sound change than others. In addition, inter- and intra-speaker variations indicate that a synchronous speech sound change is in progress in the stop system of Korean. Although there are various (methodological) difficulties in comparing the present data to those published in the literature decades ago, there can be no question that a synchronous change in the articulation of the lax and aspirated stop is in progress. Taken the previous findings of socio-factors together, we might say that young female Seoul speakers are the most active initiators and propagators in sound change in the Korean stop system.

The present study has an implication that Korean stops are still in progress in sound change, instead of being completed and, among the three stops, the "lax" category seems to play a pivotal role in a sound change (or tonogenesis) in present-day Korean. Further inquiry is necessary to figure out how the phonetic drift of "lax" stops in Korean can be phonologically accounted for.

#### Acknowledgements

I thank Allard Jongman, Sookhang Lee, Jongho Jun, Byunggon Yang, and three anonymous reviewers for their helpful comments on an earlier draft. I'm grateful to all participants. All the remaining errors are mine.

## References

- Bang, H., Sonderegger, M., Kang, Y., Clayards, M., & Yoon, T. J. (2018). The emergence, progress, and impact of sound change in progress in Seoul Korean: Implications for mechanisms of tonogenesis. *Journal of Phonetics*, 66, 120-144.
- Beddor, P. S. (2009). A coarticulatory path to sound change. *Language*, 85(4), 785-821.
- Boersma, P., & Weenink, D. (2018). Praat: Doing phonetics by computer (version 6.0.37) [Computer program]. Retrieved from <http://www.praat.org/>
- Chang, C. B. (2012). Rapid and multifaceted effects of second-language learning on first-language speech production. *Journal of Phonetics*, 40(2), 249-268.
- Cho, T., Jun, S. A., & Ladefoged, P. (2002). Acoustic and aerodynamic correlates of Korean stops and fricatives. *Journal of Phonetics*, 30(2), 193-228.
- Choi, H. (2002). Acoustic cues for the Korean stop contrast: dialectal variation. *ZAS papers in Linguistics*, 28, 1-12.
- Coetzee, A. W., Beddor, P. S., & Wissing, D. P. (2014). Emergent tonogenesis in Afrikaans. *Journal of Acoustical Society of America*, 135, 2421.
- Flege, J. E. (1995). Second language speech learning: Theory, findings, and problems. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 233-277). Baltimore, MD: York Press.
- Flege, J. E. (2002). Interactions between the native and second-language phonetic systems. In P. Burmeister, T. Piske, & A. Rohde (Eds.), *An integrated view of language development: Papers in honor of Henning Wode* (pp. 217-244). Trier, Germany: Wissenschaftlicher Verlag.
- Han, M. S., & Weitzman, R. S. (1970). Acoustic features of Korean /P, T, K/, /p, t, k/ and /p<sup>h</sup>, t<sup>h</sup>, k<sup>h</sup>/. *Phonetica*, 22, 112-128.
- Jin, W. (2008). *Sounds of Chinese Korean: A variationist approach* (Doctoral dissertation). The University of Texas, Arlington, TX.
- Jun, S. A. (1993). *The phonetics and phonology of Korean prosody* (Doctoral dissertation). Ohio State University, Columbus, OH.
- Kang, K. H., & Guion, S. G. (2008). Clear speech production of Korean stops: Changing phonetic targets and enhancement strategies. *The Journal of the Acoustical Society of America*, 124(6), 3909-3917.
- Kang, Y. (2014). Voice onset time merger and development of tonal contrast in Seoul Korean stops: A corpus study. *Journal of Phonetics*, 45, 76-90.
- Kang, Y., & Han, S. (2013). Tonogenesis in early contemporary Seoul Korean: A longitudinal case study. *Lingua*, 134, 62-74.
- Kang, Y., & Nagy, N. (2016). VOT merger in heritage Korean in Toronto. *Language Variation and Change*, 28(2), 249-272.
- Keating, P. A. (1984). Phonetic and phonological representation of stop consonant voicing. *Language*, 60(2), 286-319.
- Kim, C. W. (1965). On the autonomy of the tensity feature in stop classification (with special reference to Korean stops). *Word*, 21, 339-359.
- Kim, M. R. (2000). *Segmental and tonal interactions in English and Korean: A phonetic and phonological study* (Doctoral dissertation). The University of Michigan, Ann Arbor, MI.
- Kim, M. R. (2008). Lax stops in Korean revisited: VOT neutralization. *Studies in Phonetics, Phonology, and Morphology*, 14(2), 3-20.
- Kim, M. R. (2011). The relationship between cross-language phonetic influences and L2 proficiency in terms of VOT. *Phonetics and Speech Sciences*, 3(3), 3-11.
- Kim, M. R. (2014). Ongoing sound change in the stop system of Korean: A three- to two-way categorization. *Studies in Phonetics, Phonology, and Morphology*, 20(1), 51-82.
- Kim, M. R. (2017). Inter-speaker and intra-speaker variability on sound change in contemporary Korean. *Phonetics and Speech Sciences*, 9(3), 25-32.
- Kim, M. R., Beddor, P. S., & Horrocks, J. (2002). The contribution of consonantal and vocalic information to the perception of Korean initial stops. *Journal of Phonetics*, 30(1), 77-100.
- Kim, M. R. and Duanmu, S. (2004). "Tense" and "lax" stops in Korean. *Journal of East Asian Linguistics*, 13, 59-104.
- Kim, Y. (1995). *Acoustic characteristics of Korean coronal stops, affricates, and fricatives* (Master's thesis). The University of Texas, Arlington, TX.
- Kirby, J. P. (2014). Incipient tonogenesis in Phnom Penh Khmer: Acoustic and perceptual studies. *Journal of Phonetics*, 43, 69-85.
- Labov, W. (1994). *Principles of linguistic change, Internal factors*. Oxford: Basil Blackwell.
- Ladefoged, P., & Maddieson, I. (1996). *The sounds of the world's languages*. Oxford, UK: Blackwell.
- Lee, H., & Jongman, A. (2012). Effects of tone on the three-way laryngeal distinction in Korean: An acoustic and aerodynamic comparison of the Seoul and South Kyungsang dialects. *Journal of the International Phonetic Association*, 42(2), 145-169.
- Lisker, L., & Abramson, A. (1964). A cross-language study of voicing in initial stops: Acoustical measurements. *Word*, 20, 384-422.
- Matisoff, J. A. (1973). Tonogenesis in Southeast Asia. *Consonant types and tone: Southern California Papers in Linguistics*, 1, 72-95.
- Oh, E. (2009). Voice onset time of Korean stops as a function of speaking rate [written in Korean]. *Phonetics and Speech Sciences*, 1(3), 39-48.
- Oh, E. (2011). Effects of speaker gender on voice onset time in Korean stops. *Journal of Phonetics*, 39(1), 59-67.
- Oh, M., & Yang, H. (2013). The production of stops by Seoul and Yanbian Korean speakers. *Phonetics and Speech Sciences*, 5(4), 185-193.
- Sapir, E. (1921). *Language: An Introduction to the study of speech*. Mineola, NY: Dover Publications.
- Silva, D. J. (2006). Acoustic evidence for the emergence of tonal contrast in contemporary Korean. *Phonology*, 23(2), 287-308.
- SPSS/PASW Statistics. (2018). *IBM SPSS statistics for Windows (Version 25.0)* [Computer program]. Armonk, NY: IBM.
- Weinreich, U. (1953). *Languages in contact*. The Hague : Mouton.

• **Mi-Ryoung Kim** Corresponding author  
 Department of Practical Foreign Languages  
 Korea Soongsil Cyber University  
 23 Samil-daero 30-gil, Jongno-gu, Seoul, Korea  
 Tel: 02-708-7845  
 Email: kmrg@mail.kcu.ac  
 Fields of interest: Phonetics, Phonology