Title:

Tonogenesis in early Contemporary Seoul Korean: a longitudinal case study

Authors:

Yoonjung Kang

University of Toronto Scarborough/University of Toronto, Canada

Sungwoo Han

Inha University, Korea

Mailing Address and Phone number of the lead author:

1265 Military Trail

University of Toronto Scarborough, Department of Humanities

Toronto, ON, M1C 1A4, Canada

416-287-7172

Email:

yoonjung.kang@utoronto.ca

Short title:

Tonogenesis in early Contemporary Seoul Korean

Title:

Tonogenesis in early Contemporary Seoul Korean: a longitudinal case study

Abstract:

Recent apparent time studies observed that Seoul Korean is undergoing a tonogenetic sound change whereby the VOT contrast between aspirated and lenis stops is being merged and the contrast is more reliably signaled by difference in F0 of the following vowel. This paper presents an instrumental phonetic study of stops in early 20th century Seoul Korean based on audio recordings of textbooks from 1935 and a longitudinal case study of a child speaker from 1935 rerecorded in 2005. The results show that speakers of 1935 exhibit a more conservative pattern of stop realization than Present Day Seoul Korean speakers of comparable age and gender confirming that a tonogenetic sound change has been in progress over the last century. Also, the speaker examined in the longitudinal case study underwent change in the direction of sound change over his lifespan indicating that the apparent time studies likely underestimate the rate of change. The study is significant in being the first longitudinal instrumental phonetic study of tonogenetic sound change and also the first ever instrumental phonetic study of the Korean language from prior to 1945.

Acknowledgements:

SSHRC Standard Research Grant

Ontario Work Study Program

Mary Bowden

Hyesun Cho

Bethany MacLeod

Ron Smyth

Audience at Harvard Workshop on Korean Linguistics

Audience at Linguistics Colloquium at Seoul National University

INTRODUCTION

Tonogenesis is a commonly attested sound change whereby consonantal contrasts of voicing, aspiration and glottalization, or phonation contrast of breathiness and creakiness give rise to and eventually become replaced by tonal contrasts (Haudricourt 1971; Hombert, 1978; Hombert, Ohala, & Ewan, 1979; Kingston, 2011; Maran, 1973; Matisoff, 1973; Thurgood 2002). In many cases, tones start out as redundant phonetic attributes of consonantal contrast (e.g., ba vs. pa), then develop into a robust distinction coexisting with the original consonantal contrast (bà vs. pá), and finally become the primary contrastive feature (pà vs. pá), as the original consonantal features are lost. Many studies reconstruct tonogenetic sound change by comparing linguistic data from different time periods of a language or synchronic data from related dialects and languages, where some of the dialects and languages retain the earlier state of consonantal contrast while others have replaced the consonantal contrasts with tonal contrasts (Haudricourt, 1971; Karlgren, 1966; Kingston, 2005; Matisoff 1973). There are also a number of instrumental phonetic studies of languages in a state of transition from consonantal to tonal constrast (Chen, 2011; DiCanio, 2012; Mazaudon & Michaud, 2008) and also a phonetic study that compares different endpoints of sound change by examining related dialects of a language, one that retains the original consonantal contrast and another that replaced the consonantal contrast with a tonal contrast (Svantesson & House, 2006). Also, there are studies that examine contact-induced synchronic variation in the realization of tonal contrast along a tonogenetic trajectory (Brunelle, 2009; Pearce, 2009). Relatively rare are instrumental studies that examine the diachronic sound change within a single speech community. Hyslop (2009) and Abramson, L-Thongkum, & Nye (2004) are notable exceptions in that they examined variation within a speech community with an eye to tracking diachronic sound change but these studies are limited in their scope. Also, as far as we know, there is no instrumental study that examines a tonogenetic sound change in a single dialect of a language across real time or a longitudinal study of the same speakers over time.[i]

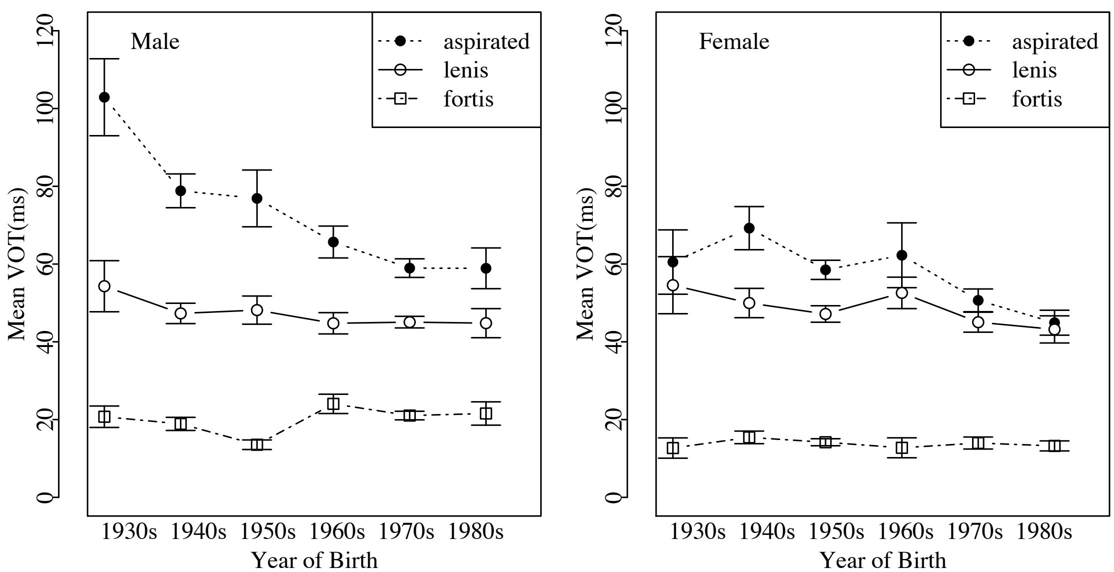
Contemporary Seoul Korean (1890-present) presents an ideal opportunity to examine this sound change in progress. Korean has a three-way laryngeal contrast of voiceless stops, among aspirated stops /p^h,t^h,k^h/, lenis stops /p, t, k/, and fortis stops /p’,t’,k’/. Previous studies show that the contrast is signalled by a combination of acoustic cues including voice onset time (VOT) of the stop and fundamental frequency (F0) and amplitude difference between the first and the second harmonics (H1-H2) of the following vowel; aspirated stops have the longest VOT values, fortis stops have the shortest values, and lenis stops have intermediate values; F0 on the following vowel is higher for aspirated and fortis stops than for lenis stops; vowels following aspirated and lenis stops have breathier voice quality than vowels following fortis stops as indicated by higher H1-H2 values (Cho, Jun, & Ladefoged, 2002; Han & Weitzman, 1970; Hardcastle, 1973; Kang & Guion, 2008; Kim, 1965; Kim, 1994; Lisker & Abramson, 1964).

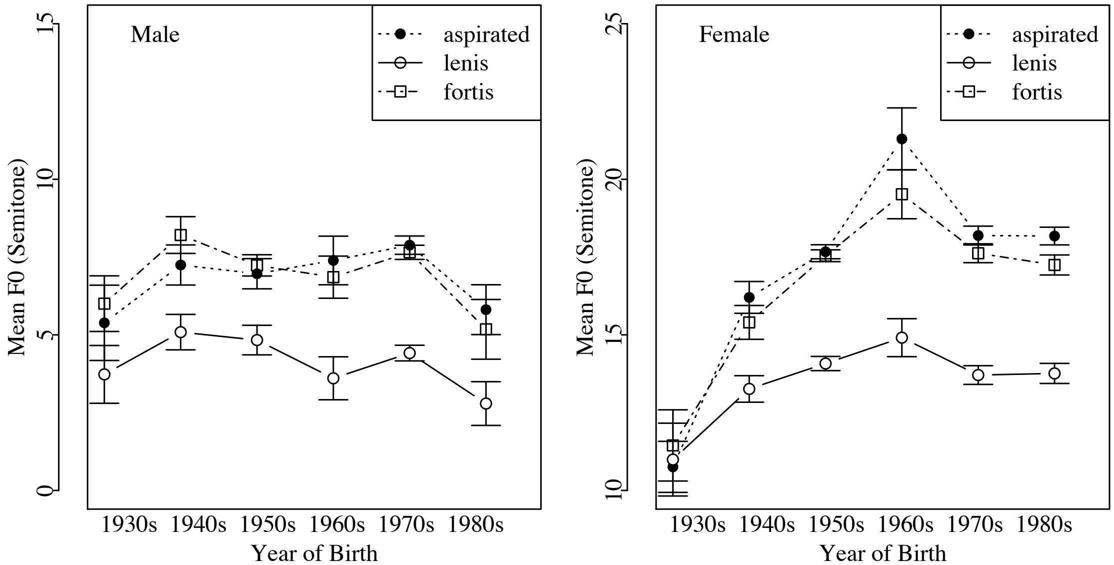
There are a number of recent apparent-time studies (Bailey, Wikle, Tillery & San, 1991; Labov, 1963) on the emergence of tonal contrast in Seoul Korean (Kang & Guion, 2008; Silva, 2002, 2006; Wright, 2007; ____, submitted). These studies show that Seoul Korean is in fact in the process of losing the VOT (Voice Onset Time) contrast between aspirated and lenis stops (/p^h t^h k^h/ vs. /p t k/) in phrase-initial position and the formerly redundant pitch difference—a high pitch on vowels following aspirated stops and a low pitch on vowels following lenis stops—is replacing the VOT difference as the primary cue of the contrast. A similar change is reported for a dialect of Chinese Korean (Jin, 2008).

An Accentual Phrase (AP) is a crucial unit in the realization of tones in Present Day Korean. An AP is a unit of prosodic structure that is above Prosodic Word and below Intonational Phrase in the prosodic hierarchy of Seoul Korean (Jun, 1993). Accentual Phrases (AP) are marked by boundary tones and the initial boundary tones vary between LH and HH depending on the laryngeal feature of the AP-initial consonant; APs that begin with a lenis stop or affricate or a sonorant are marked by LH boundary tones with L falling on the initial syllable and H on the second, while phrases that begin with a fortis or aspirated stop or affricate, /h/, or a coronal fricative, are marked by HH boundary tones with H falling on each of the first two syllables (Jun, 1993). The consonant-induced F0 difference extends far beyond the initial portion of the immediately following vowel indicating that the F0 perturbation is not an automatic consequence of physiological restrictions in laryngeal articulation (Jun, 1996; Kim, 2000; Silva, 2006). Recent perception studies also find that F0 is a crucial perceptual cue for stop distinction, for lenis-aspirated stops in particular (Kim, 2004; Kim, Beddor, & Horrocks, 2002; Kong, Beckman, & Edwards, 2011; Lee & Jongman, 2011). So, this recent change in Seoul Korean is consistent with a process of tonogenesis, whereby consonant-induced F0 perturbation is exaggerated and reinterpreted by learners as a primary contrast, eventually replacing the original voicing or phonation contrast of consonants (Hombert, Ohala, & Ewan, 1979; Kingston, 2011).

The merger of VOT is almost complete in the speech of younger female speakers, a population known to lead many sound changes, and the least advanced in the speech of older male speakers. The F0 contrast is further enhanced in younger and female speakers' speech. Figure 1, reproduced from _____ (submitted), shows the variation in VOT values of sentence-initial stops and F0 at the midpoint of the following vowel by speakers' age and gender based on the NIKL (The National Institute of the Korean Language) acoustic corpus of Standard Seoul Korean (2005).

Figure 1. (a) Mean VOT (ms) and (b) mean F0 (semitone, reference=100Hz) at following vowel midpoint of three stop categories by speakers' year of birth, aggregated by 10 year bands, for speakers of Seoul Korean recorded in 2003. The error bars represent 95% confidence interval. (From _______, submitted)





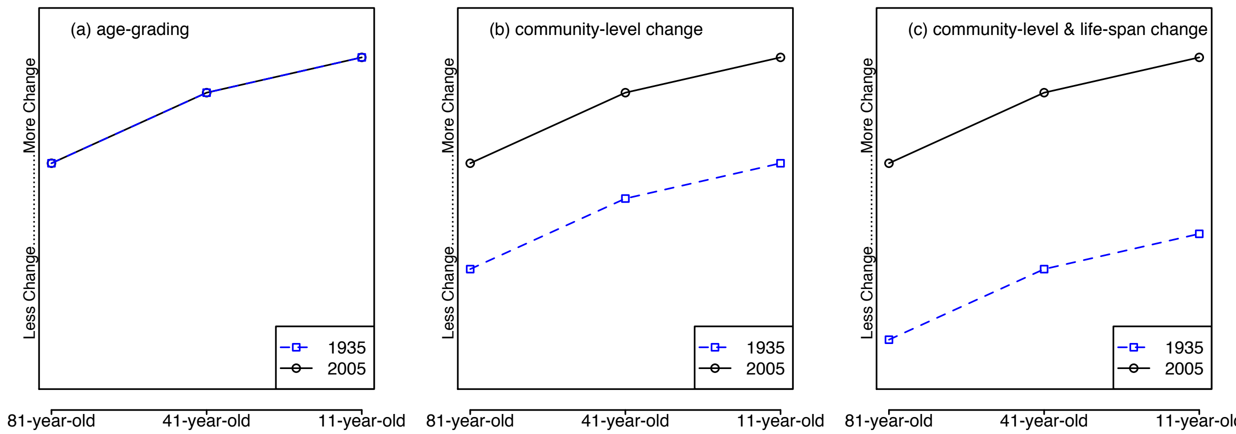
It is notable that even the oldest male speakers shown in Figure 1, i.e., those born in the 1930s, exhibit a substantial consonant-induced pitch difference of around 2 semitones at the midpoint of the following vowel, indicating that the tonogenetic process has already advanced fairly far in Present Day Korean. At the same time, these speakers show a robust contrast of VOT between the two stops, showing a difference of over 40ms in mean VOT values, indicating that the sound change has not progressed far enough for F0 to overtake VOT as the primary cue. Given what we know about the general trajectory of tonogenetic sound change (cf. Maran 1973), we expect that the preceding stage of Seoul Korean had a similarly more robust VOT contrast between aspirated and lenis categories but a less distinctive F0 difference between the two stop categories.

In apparent-time studies, it is assumed that the individual speakers' speech remains more or less stable across the lifespan reflecting the state of language at the time they learned the language and that differences among speakers of different generations at a single point in time mirror diachronic change in real time. While it has been shown that the apparent time study of linguistic change is a reliable reflection of linguistic change in real time for the most part (Bailey et al., 1991), there are factors that affect the age-dependent variation other than community-level language change, such as age-grading and change across the life span (Labov, 2001; Chambers, 2003; Boberg, 2004; Sankoff, 2005; Sankoff & Blondeau, 2007). Age-grading refers to the type of variation where as speakers grow older, they take on a speech style that is appropriate for their age and gender. In this case, we can get age-dependent variation synchronically while the language at the community level does not undergo any change. Change across the life span refers to a situation where adult speakers undergo language change over the course of their lifetime in a direction consistent with the general direction of community-level sound change. In the latter case, an apparent time study of language change underrepresents the rate of change showing less differentiation between older and younger speakers than the corresponding difference over real time. Therefore, to gain a fuller picture of sound change, a comparison of data from different time periods (a trend study) and a longitudinal study of individual speakers across their lifespan (a panel study) are helpful.

In this paper, we examine the stop consonant production in an earlier stage of Seoul Korean than has been examined by previous studies. We present an acoustic study of the speech of one 41-year-old male and one 11-year-old male speaker of Seoul Korean from 1935 and the speech of the latter speaker re-recorded in 2005 at the age of 81. We examine whether and how the three sets of speech data differ from one another and from the speech of Present Day Seoul Korean speakers of comparable age and gender. As mentioned above, in Present Day Korean, older male speakers retain a relatively large VOT difference between aspirated and lenis stops and produce substantial F0 contrast on the vowels following the stops, although the degree of F0 contrast is less than in younger speakers. We expect that the speech of the 81-year old male (year of birth: 1924), recorded in 2005 will show a similar pattern and the contrast of aspirated and lenis stops is signaled by a reliable difference both in VOT and F0.

Several predictions can be made about the speech of the speakers in 1935. If the age-conditioned variation in 2005 is not a reflection of community-level sound change but due to age-grading, assuming that a similar age-grading pattern was active in 1935 we expect that speakers of 1935 will mirror the speech patterns of comparable age groups in 2005. This prediction is schematically represented in Figure 2, adapted from Boberg (2004). In the graphs, the x-axis represents the age of the speaker, the y-axis represents the degree of progress along the relevant change, i.e., a combination of VOT and F0 realization patterns, the blue dotted line represents the 1935 pattern and the black solid line represents the 2005 pattern. If age-grading is responsible for the age-conditioned variation, the two lines will coincide with each other as shown in Figure 2(a). So, we expect more a conservative pattern in the speech of the 81-year-old speaker in 2005 than in the speech of the two 1935 speakers.[ii]

Figure 2. Schematic representations of expected age-conditioned variation in 1935 and 2005 under (a) age-grading, (b) community-level change, and (c) community-level change along with lifespan change.



On the other hand, if the synchronic variation in Present Day Korean is due to sound change in progress, we expect that the 41-year-old speaker in 1935, who is a whole generation ahead of the 81-year-old speaker of 2005 in real time, to produce a more conservative speech pattern than the 81-year-old speaker as schematically shown in Figure 2 (b) and (c). Less clear is what to expect from the 11-year-old speaker in 1935. It has been suggested that during the age of language acquisition, assumed to be up to and around the age of 4, children acquire the speech of their caregiver, female in most cases, but beyond this initial acquisition period, they start to deviate from their caregiver's speech and take on the speech style of their peer groups in their adolescent years, injecting further changes to the language (Labov, 2001; Tagliamonte and D'Arcy, 2009). It has been noted that in female-led sound changes, this change during adolescence is not limited to but most notable with girls, creating the commonly attested gender asymmetry whereby female speakers push the change further than male speakers. Even after the teenage years, individual speakers may continue to change their speech. Sankoff and

Blondeau (2007), for example, showed that some, but not all, speakers of Montreal French changed their pattern of [r] production along the overall direction of sound change, during their adult years. Since our speaker is male, there is no clear prediction as to whether the speakers' speech may have changed over the 70 years between the two recordings. But, in either case, if there is indeed a sound change in progress at the community level, we expect that the 1935 child speech is at least as conservative (without any change across life span) or even more conservative than his speech 70 years later (with change across life span). These two predictions are schematically represented in Figure 2. (b-c).

2. Material and Methods

The data from 1935 are drawn from the recording of elementary school textbooks, *Pothonghakkyo Cosunetokpon* 'Korean reader for normal school', which were first published in 1911 and underwent several revisions over the years. The recording consists of 27 tracks and the speakers were a group of 5th graders attending *Kyengseng Sapem Pusok Pothong Hakkyo*, an elite elementary school in Seoul and a 41 year-old male, Mr. Shim Uy-Lin, who was a linguist and also a teacher at the school. They were all speakers of Seoul Korean born and raised in Seoul. The recording was published as an SP by a Japanese recording company, OKEH in 1935, and reproduced in CD format by Koyang Cultural Foundation in 2004 (Koyang Cultural Foundation, 2004). A detailed phonological study of the recording is found in Han (2005). The data for the current paper is from track #2, which is a reading of the Korean syllabary. In this track, Mr. Shim, the teacher, recites the Korean syllabary and Mr. Chung Kye-Whan, then an 11-

year-old boy, repeats after the teacher after each two-syllable sequence. Once they have gone through the entire syllabary, the younger speaker recites the entire syllabary by himself. In 2005, the second author re-recorded the reading of the textbook by Mr. Chung Kye-Whan, by then 81 years old. The recording was initially made in an analogue tape and later digitized for acoustic analysis. For the rest of the paper, we will refer to Dr. Shim Uy-Lin as T (for Teacher), Mr. Chung Kye-Whan of 1935 as S (for Student), with S1 referring to the first round of production and S2, the second, and Mr. Chung Kye-Whan of 2005 as A (for Adult).

Fortunately for our purpose, the Korean syllabary provides a controlled dataset for the study of Korean consonants, and their effect on the pitch of the following vowels. The syllabary consists of 140 syllables in total, which are combinations of 14 onsets (/p p^h t t^h k k^h c c^h s h m n l/ and null) and 10 vowels or glide-vowel sequences (/a ja ə jə o jo u ju ɨ i/). For example, the first line of the syllabary is /ka-kja kə-kjə ko-kjo ku-kju kɨ-ki/, with each two-syllable sequence read as a unit, roughly corresponding to an Accentual Phrase (AP). As the AP-medial lenis stops are allophonically voiced and only AP-initial consonants induce substantial tonal contrast on the following vowel in Present Day Korean, we will focus on the realization of AP-initial consonants. Due to extraneous noise, a few tokens are excluded from the analysis. As a result, our data consists of a total of 70 two-syllable phrases for speaker T(eacher), 69 for speaker A(dult), and 138 for speaker S(tudent) (68 for S1 and 70 for S2).

Note that not all consonants and vowels of Korean are represented in the syllabary. Table provides the full inventory of Korean consonants and monophthongal vowels. Missing from the syllabary are /ŋ/, which does not occur in onset position, all fortis consonants /p' t' k' c' s'/ and non-high front vowels /e ɛ/ as well as various glide+vowel sequences. The absence of fortis consonants is particularly unfortunate but as the main focus of our study is the change in the relationship between the lenis and the aspirated obstruents, the data is adequate in that regard.

Table 1 *The Inventory of Korean Consonants and Monophthongal vowels*

p p' p^h	t t' t^h	k k' k^h	
	c c' c^h		
	s s'		h
m	n	ŋ	
	l		

i	ɨ	u
e	ə	o
ɛ	a	

The acoustic analysis was conducted using *Praat* (Boersma & Weenink, 2011). For all phrase-initial CV sequences in the data, the onset of stop release, the onset of voicing on the vowel and the offset of the vowel were manually identified. The quality of the recording was less than ideal and while the vowel onset and offset can be identified reliably, the onset of stop release was often not clearly visible on the waveform or spectrogram. We tried the noise removal function of *Audacity* on the data but it did not make a noticeable improvement. To ensure the reliability of analysis, the manual segmentation was conducted twice, once by a research assistant, who was naïve as to the purpose of the study and did not know Korean, and once by the first author. Pearson's product-moment correlation between the VOT (Voice Onset Time) measurements based on the two segmentations was highly significant ($t = 21.1395$, $df = 114$, $p < 0.0001$, r=0.8926). Here, we will report the result based on the segmentation by the first author with the assurance that the results based on the alternative segmentation support the same overall conclusion.

The acoustic measurements included the Voice Onset Time (VOT) of the AP-initial stops, and the duration and fundamental frequency (F0) of the vowel following the initial stop. VOT was defined as the duration from the onset of the stop release to the zero crossing at the upward swing of the first periodic cycle of the following vowel. The F0 was measured at several time points during the vowel, 15ms and 30ms after the vowel onset, as well as at 25%, 50%, and 75% point of vowel duration, to examine how far into the vowel the consonant-induced pitch difference extends. Here we report the measurements at 15 ms after the vowel onset and at the temporal midpoint of the vowel using *Praat*'s pitch tracking function with pitch range set at 75-500 Hz and followed by an octave jump removal process and a manual check of pitch tracks for any irregularity or error. All measurements were extracted using the scripting function of *Praat.* As we are comparing the pitch pattern of speakers at three different ages, F0 measurements in Hz are converted to semitones (St), a logarithmic scale of pitch, which allows for comparison of pitch range across gender and age (Oh, 2011; Whalen & Levitt, 1995). St was calculated with 100Hz as a reference pitch using the formula, $\log_2(Hz/100)*12$. Of primary interest to us is how our speakers use the combination of the acoustic cues, VOT of the initial stops and the F0 of the following vowel to signal the contrast between the aspirated and lenis stops. Statistical tests were conducted in *R* (R Development Core Team 2011).

3. RESULTS

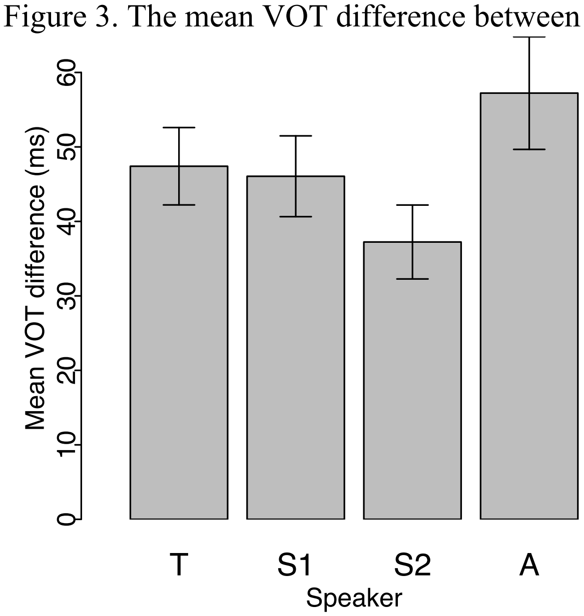
We first examine the VOT and F0 values of the stops separately and then examine how the weighting of the two cues change from speaker to speaker. The data from the two sets of productions by speaker S are reported separately as S1 and S2 because the two repetitions are produced under different conditions, S1 as repetitions of teacher's speech, and S2 as independent production but as a second reading of the material. Table summarizes the mean VOT values of each stop category for each speaker. In all four datasets, the VOT values are substantially longer for aspirated stops than for lenis stops. A paired-samples t-test was conducted for each dataset to compare the VOT values of aspirated and lenis stops and there was a significant difference in all four datasets (T: $t(12)= 9.1287$, $p<0.0001$; S1: $t(14)=8.4903$, $p<0.0001$; S2: $t(13)=7.501$, $p<0.0001$; A: $t(13)=7.5701$, $p<0.0001$). In line with cross-linguistic tendencies, the VOT values are longer for stops with more posterior places of articulation (Dorsal>Coronal>Labial).

Table 2 *The mean and standard deviation (in parentheses) of VOT in ms*

Speaker		Labial	Coronal	Dorsal	Total
T	Lenis	20.4 (5.8)	40.4 (9.4)	51.8 (25.1)	37.1 (21.1)
	Aspirated	73.3 (13.2)	82.6 (5.1)	97.8 (15.4)	84.3 (15.0)
S1	Lenis	23.2 (10.3)	37.3 (6.7)	43.2 (11.7)	34.6 (12.6)
	Aspirated	78.4 (10.7)	78.9 (10.4)	84.6 (18.0)	80.7 (12.8)
S2	Lenis	23.5 (11.5)	32.6 (72.)	33.9 (11.4)	30.5 (10.4)
	Aspirated	67.7 (14.6)	56.3 (8.8)	78.8 (16.5)	67.6 (15.8)
A	Lenis	34.3 (14.7)	44.8 (14.5)	51.1 (28.6)	44.0 (20.4)
	Aspirated	79.3 (14.9)	84.7 (13.7)	137.0 (18.2)	100.3 (30.6)

Figure . shows the mean VOT difference between aspirated and lenis stops in corresponding contexts for each speaker. The mean difference is 47.4 ms for T, 46.1ms for S1, 37.2 ms for S2, and 57.2 ms for A. It is notable that the second production of Speaker S, S2, shows the least difference and this is likely due to the fact that this was his second repetition and his speech was less careful than the first repetition, as is common with repeated speech material (Fowler, 1988; Fowler & Housum, 1987; Liberman, 1963). To examine whether the VOT difference by stop category is significantly different from speaker to speaker, a one-way ANOVA was conducted with the VOT difference between aspirated and lenis stop in corresponding contexts as a dependent variable and speaker as an independent variable and there was no significant effect of speaker on VOT difference [$F(3,52)=1.9305$, $p=0.1361$]. In other words, overall, there was no clear trend of VOT contrast reduction from older to younger speaker in 1935 or from the speech of a single speaker from 1935 to 2005.

Figure 3. The mean VOT difference between lenis and aspirated stops by each speaker.



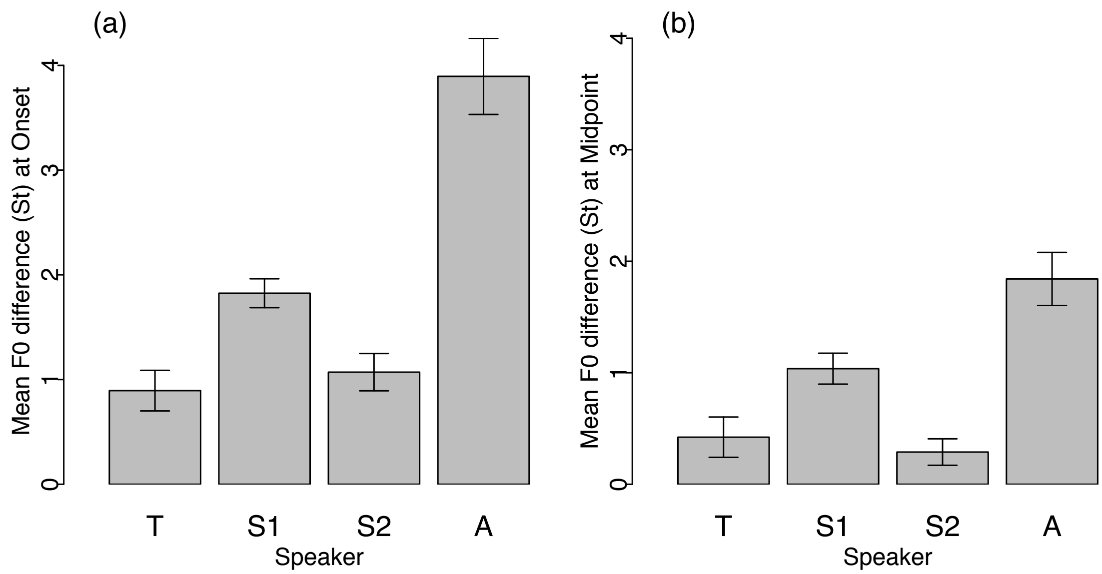
Now turning to F0 values, Table summarizes the mean F0 values both in Hz and in Semitone units at vowel onset (measured at 15ms after vowel of voicing) and at temporal midpoint of vowel for each stop category for each speaker. In all four datasets, the F0 values are substantially higher for aspirated stops than for lenis stops, both at the onset and at the midpoint of vowel. The statistical analyses reported here are conducted on the F0 values in Semitone units but the results are comparable in analyses based on F0 values in Hz units. A paired-samples t-test was conducted for each dataset to compare the F0 values of aspirated and lenis stops and there was a significant difference in all datasets both for the measurements at vowel onset (T : $t(14)= 4.6158$, $p=0.0004$; S1: $t(14)=13.2$, $p<0.0001$; S2: $t(14)=6.012$, $p<0.0001$; A: $t(10)=10.7$, $p<0.0001$) and at vowel midpoint (T : $t(14)= 2.3366$, $p=0.0348$; S1: $t(14)=7.4771$, $p<0.0001$; S2: $t(14)=2.4321$, $p=0.0290$; A: $t(13)=7.7392$, $p<0.0001$).

Table 3 *The mean and standard deviation (in parentheses) of F0 in Hertz and Semitone scales*

		Hertz		Semitone	
Speaker		Onset	Midpoint	Onset	Midpoint
T	Lenis	128.0 (5.4)	123.4 (4.1)	4.3 (0.7)	3.6 (0.6)
	Aspirated	134.8 (5.7)	126.5 (4.9)	5.2 (0.7)	4.1 (0.7)
S1	Lenis	262.0 (8.5)	262.2 (8.1)	16.7 (0.6)	16.7 (0.5)
	Aspirated	291.1 (9.8)	278.4 (9.3)	18.5 (0.6)	17.7 (0.6)
S2	Lenis	292.1 (5.7)	291.2 (3.9)	18.6 (0.3)	18.5 (0.2)
	Aspirated	310.8 (9.1)	296.2 (6.8)	19.6 (0.5)	19.9 (0.4)
A	Lenis	165.9 (13.4)	175.5 (8.2)	8.7 (1.4)	9.7 (0.8)
	Aspirated	206.9 (14.7)	195.4 (7.0)	12.5 (1.2)	11.6 (0.6)

While the F0 difference is significant in all contexts for all speakers, the F0 difference was noticeably larger for speaker A, 41.0 Hz or 3.8 St at onset and 19.9 Hz or 1.9 St at midpoint, than for the others. Also, the F0 difference was larger at vowel onset than at vowel midpoint. Figure 4. summarizes the mean difference of F0 values between aspirated and lenis stops in corresponding contexts, at the onset and the midpoint of the following vowel.

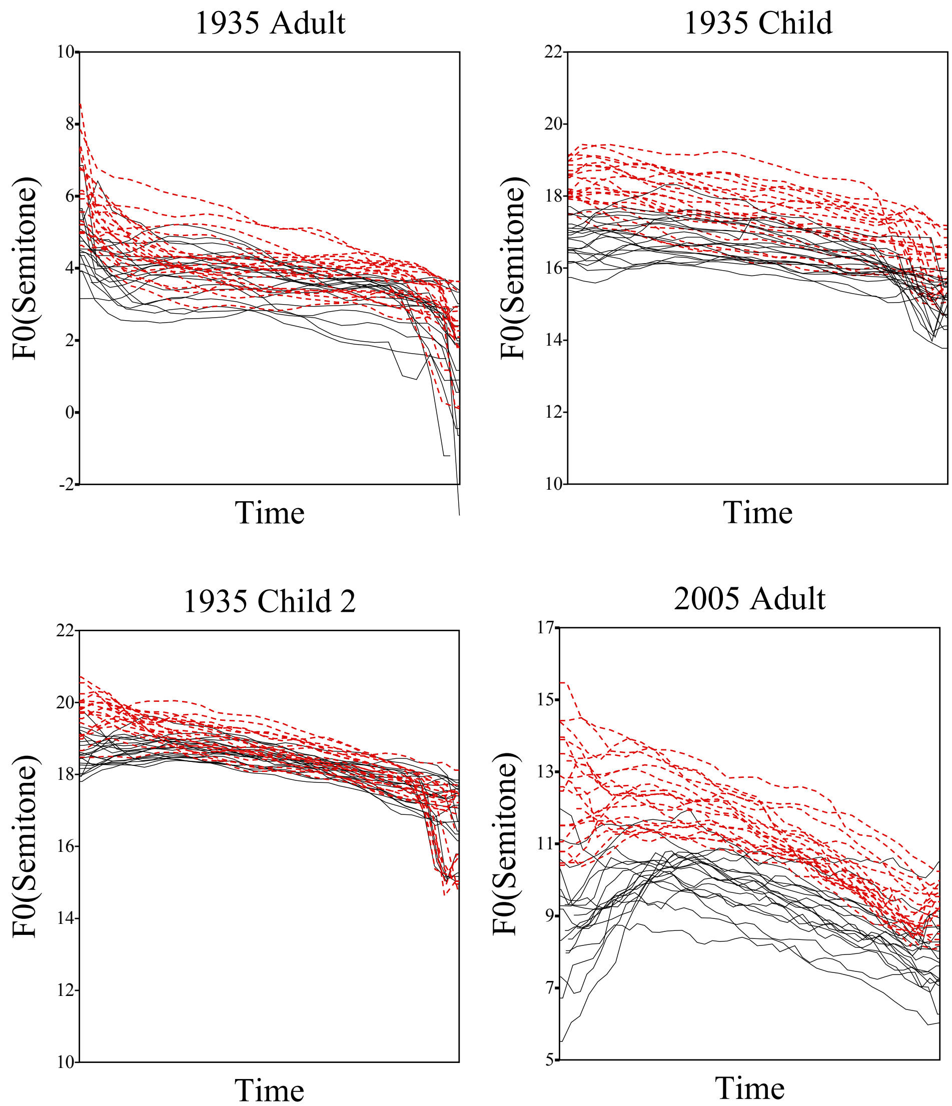
Figure 4. The mean F0 (semitone) difference between lenis and aspirated stops by each speaker: (a) at 15ms after the vowel onset and (b) at vowel midpoint.



To examine whether the F0 difference by stop category is significantly different from speaker to speaker and from the onset to the midpoint, a two-way ANOVA was conducted with the F0 differences between aspirated and lenis stops in corresponding contexts as a dependent variable and speaker (T, S1, S2, and A) and the position in vowel (onset and midpoint) as independent variables. There was a significant effect of speaker on F0 difference [$F(3,110)=41.996$, $p<0.0001$] and a significant effect of position on F0 difference [$F(1,110)=44.880$, $p<0.0001$] with the difference being larger at vowel onset than at vowel midpoint. A Tukey's HSD post-hoc test showed that the F0 difference was largest for Speaker A, smallest for Speaker T and S2, and intermediate for S1 (A> S1> S2, T).

The different realization of F0 by speaker is also clearly visible in Figure 5. The graphs show the pitch tracks of all tokens of vowels that follow AP-initial stops in normalized time with the red dashed lines representing tokens of aspirated stops and the black solid lines representing tokens of lenis stops. Throughout the vowels, the pitch tracks of the two stop categories are very well separated for the 2005 adult speaker (A) while for the 1935 adult (T) and the second repetition of 1935 child (S2), the two largely overlap. For the first production of 1935 child (S1), the pattern is intermediate, showing more overlap than Speaker A, but more differentiation than Speaker T. As was the case with VOT contrast, the child speaker is making a less clear distinction between stop categories in the second repetition (S2) than in the first (S1).

Figure 5. The F0 (semitone) contours of aspirated (red dashed lines) and lenis (black solid lines) contexts in normalized time encompassing the entire vowel duration



The significant speaker effect on F0 difference is in contrast to the lack of significant speaker effect on VOT difference. In other words, the VOT difference between aspirated and lenis stops remained generally stable while the F0 difference increased over time, both real and apparent, with the adult male of 1935 showing the most conservative pattern, the 11 year boy of 1935 showing an ambiguous pattern and the older male of

2005 showing an increased contrast in F0. This indicates that in this early stage of tonogenesis, the original consonantal contrast of VOT was still retained and remained stable while the F0 contrast emerged and further amplified. The 11-year-old boy's speech underwent change to expand the F0 contrast over the 70 years.

So far, we examined the two acoustic characteristics, VOT and F0, separately but as we are interested in change in the function of these cues as primary contrast, it will be instructive to examine both VOT and F0 cues together and see how each speaker combines the two acoustic cues together to signal the contrast between the stop categories and which of the two cues are relied on more for signaling the contrast. For this purpose, a linear discriminant analysis (LDA) was conducted with the laryngeal category as the response variable and the VOT and F0 at vowel midpoint as the independent variables to determine to what extent each of the two acoustic cues contributes to the distinction of the stop categories. The same analysis with F0 at vowel onset, instead of F0 at vowel midpoint, produces comparable results and is not separately reported. The VOT and F0 values were normalized using z-score transformation so that the effect size of the two can be directly compared from the coefficients of linear discriminants. The coefficients values are summarized in Table . For Speaker T, the 41-year-old speaker from 1935, the coefficient for VOT is larger than F0 by a ratio of over 10 to 1, indicating that the speaker mostly relies on VOT for the stop distinction. For Speaker A, the 81-year-old speaker from 2005, on the other hand, the coefficient for F0 is slightly larger than that for VOT indicating that he relies on both VOT and F0 with slightly more weight on F0. Speaker S, the 11-year-old speaker of 1935, shows the intermediate pattern with VOT outweighing F0 but the difference is not as drastic as is the case in Speaker T's speech.

Table 4 Discriminant Coefficients for VOT and F0 by speaker

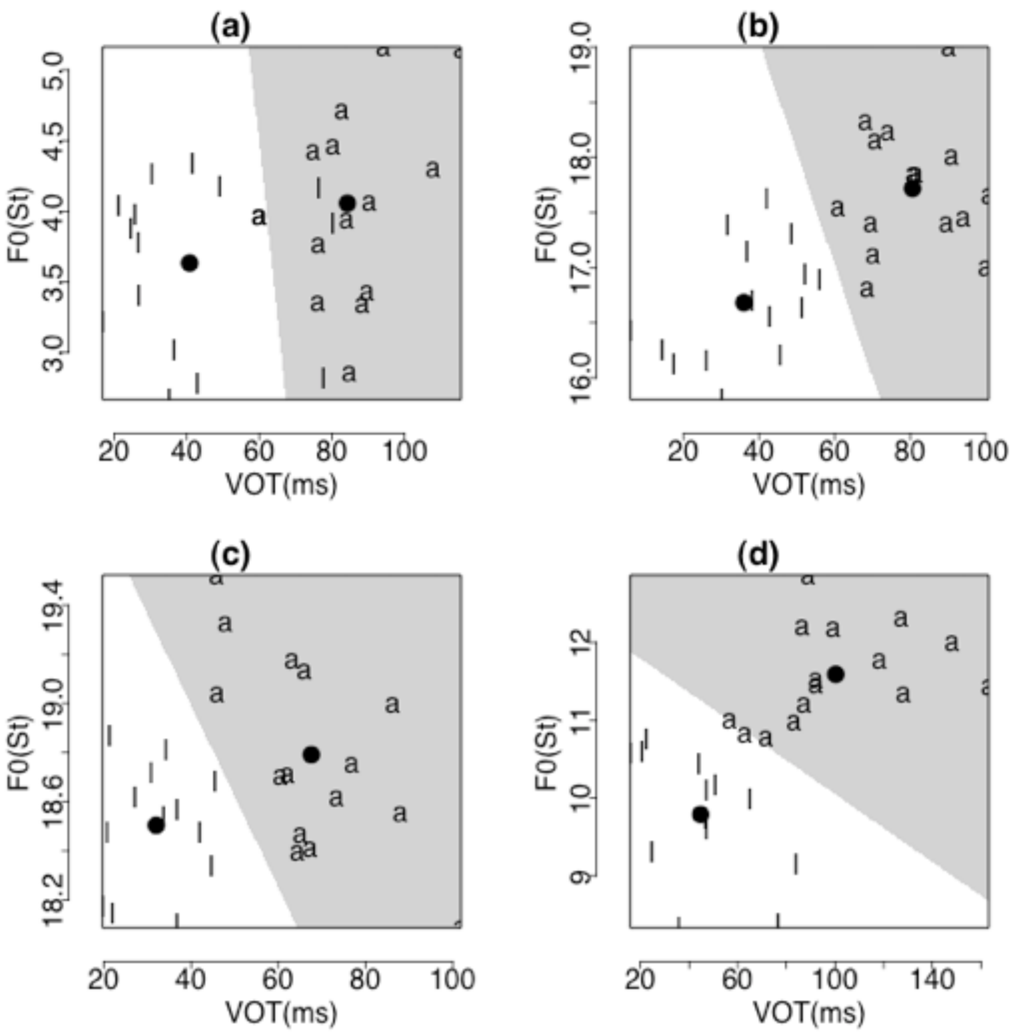
Speaker	VOT	F0	Ratio: VOT/F0
T	1.5106	0.1413	10.7
S1	1.6435	0.4614	3.6
S2	1.8614	0.7936	2.3
A	0.9255	1.3060	0.7

It is also notable that the speaker S's two rounds of production patterns yield a comparable discriminant coefficient values while his VOT and F0 values alone give the appearance of a drastic difference. This suggests that while the raw values of VOT and F0 can vary significantly due to repetition effect, the relationship between the cues as revealed by linear discriminant analysis remains relatively stable across speech types, indicating that the latter measures may provide a more reliable measure of tonogenetic sound change or other sound change involving cue trading. Also, it is notable that the Speaker S' coefficient ratio is more similar to that of Speaker T in S1, when the speaker was repeating after speaker T than in S2 indicating that when speaker S was repeating and possibly imitating speaker T, he made use more use of VOT and less use of F0 than when he was not repeating after speaker T. This imitation effect was not noticeable when the VOT and F0 were examined as S1 was more similar to T than S2 was in VOT differences but S2 was more similar to T than S1 in F0 differences.

The results of the discriminant analyses are visualized in Figure 6., which shows how the two stop categories are distinguished in a two-dimensional acoustic space of VOT by F0 for each speaker. For Speaker T, the boundary between the two categories is almost vertical, i.e. the distinction between the two categories is made mainly according to their VOT values, while for Speaker A, the boundary is diagonal, showing that both VOT and

F0 play a role in defining the contrast. Speaker S shows an intermediate pattern, with the boundary being more vertical than that for Speaker A but not as vertical as for Speaker T.

Figure 6. The classification of lenis ("l") and aspirated ("a") stops in two-dimensional acoustic space of VOT (ms) by F0 (st) based on linear discriminant analyses for (a) Speaker T, (b) Speaker S1, (c) Speaker S2, and (d) Speaker A



4. DISCUSSION

We found that the speakers in 1935 show a more conservative pattern of stop contrast than the 81-year-old speaker of 2005 as well as younger speakers of Present Day Korean.

Therefore, we can confirm that the age-conditioned variation in synchronic Seoul Korean cannot be attributed to age-grading (cf. Figure 2(a)) but a reflection of genuine sound change in progress whereby the VOT contrast of aspirated and lenis stops is replaced by tonal contrast on the following vowel. We also found that Mr. Chung's speech underwent substantial change during the 70 years between the two recordings along the direction of general community-level sound change, consistent with the view that an individual speaker's speech does not necessarily remain stable as they exit the critical period of language acquisition but undergo further changes through their lifespan (Figure 2(c)). This result is also in agreement with the impressionistic observation by the author that the speech of 1935 recording sounds quite different from Present Day Seoul Korean, especially in the intonational patterns as well as the speech of Mr. Chung's in 2005.

The results suggest that the synchronic age-dependent variation of stop production likely underestimate the rate of sound change in real time and the emergence of tonal contrast in Seoul Korean is in fact fairly recent, contrary to the suggestion by Silva (2006). In other words, even though many speakers of Present Day Seoul Korean in their 70s and 80s show substantial F0 contrast between stop categories, this is not because this stage of sound change was already complete during the period of acquisition for these speakers (1920s and 1930s) but the change was just emerging during the 1920s and 1930s. Partial support for this view comes from the fact that even in the synchronic data presented in Figure 1, some of the older female speakers in fact show minimal F0 difference between lenis and aspirated stops, similar to the pattern shown by the speakers in 1935. Sankoff and Blondeau (2007) showed that in their study of Montreal /r/, while some speakers underwent further change during their adult years in the direction of general sound change, there are speakers who retained their older pattern. It seems that we are also seeing this mixed pattern from our oldest speakers in Present Day Korean, some undergoing further changes while others not.[iii] While the present study is limited to the speech of two male speakers, the study is valuable in that it is the only instrumental acoustic study of Korean stops from data before 1945 and also the only longitudinal instrumental phonetic study of tonogenetic sound change and confirmed that Seoul Korean has been indeed undergoing tonogenetic sound change.

REFERENCES

Abramson, Arthur S., L-Thongkum, Theraphan, & Nye, Patrick W. (2004). Voice register in Suai (Kuai): an analysis of perceptual and acoustic data. Phonetica 61:147-171.

Bailey, Guy, Wikle, Tom, Tillery, Jan, & San, Lori. (1991). The apparent time construct. Language Variation & Change 3:241-264.

Boberg, Charles. (2004). Real and apparent time in language change: late adoption of changes in Montreal English. American Speech 79(3):250-269.

Boersma, Paul, & Weenink, David. (2011). Praat: doing phonetics by computer [Computer program]. Version 5.2.25. <http://www.praat.org/>

Brunelle, Marc. (2009). Contact-induced change? Register in three Cham dialects, Journal of the Southeast Asian Linguistics Society 2:1-21.

Chambers, J. K. (2003). Patterns of variation including change. In J. K. Chambers, Peter Trudgill, & Natalie Schilling-Estes (Eds.), The Handbook of Language Variation and Change. Oxford: Blackwell Publishers. 349-372.

Chen, Yiya. (2011). How does phonology guide phonetics in segment–f0 interaction? Journal of Phonetics 39:612-625.

Cho, Taehong, Jun, Sun-Ah, & Ladefoged, Peter. (2002). Acoustic and aerodynamic correlates of Korean stops and fricatives. Journal of Phonetics 30:193–228.

Choi, Hansook. (2002). Acoustic cues for the Korean stop contrast: dialectal variation. ZAS papers in linguistics 28:1–12.

DiCanio, Christian T. (2012). Coarticulation between tone and glottal consonants in Itunyoso Trique. Journal of Phonetics 40:162-176.

Fowler, Carol A. (1988). Differential shortening of repeated content words produced in various communicative contexts. Language and Speech 31.4:307-319.

Fowler, Carol A., & Housum, Jonathan. (1987). Talkers' signaling of "new" and "old" words in speech and listeners' perception and use of the distinction. Journal of Memory and Language 26:489-504.

Han, Sung-Woo. (2005). Phonological study of the acoustic data from Pothonghakkyo cosenetokpon. Emunyenku 33:29-58. [In Korean]

Han, Mieko S., & Weitzman, Raymond S. (1970). Acoustic features of Korean /P, T, K/, /p, t, k/ and /ph, th, kh/. Phonetica 22:112-128.

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

Haudricourt, André G. (1971). Two-way and three-way splitting of tonal systems in some far-eastern languages. In J. G. Harris & R. B. Noss (Eds.), Tai phonetics and phonology. Bangkok: Central Institute of English Language, Mahidol University. 58–86.

Hombert, Jean-Marie. (1978). Consonant types, vowel quality, and tone. In V. A. Fromkin (Ed.), Tone: A linguistic survey. New York: Academic Press. 77-111.

Hombert, Jean-Marie, Ohala, John J., & Ewan, William G. (1979). Phonetic explanations for the development of tones. Language 55:37-58.

Hyslop, Gwendolyn. (2009). Kurtöp tone: a tonogenetic case study. Lingua 119:827–845.

Jin, Wenhua. (2008). Sounds of Chinese Korean: a variationist approach. Doctoral dissertation, University of Texas at Arlington.

Jun, Sun-Ah. (1993). The phonetics and phonology of Korean prosody. Doctoral dissertation, The Ohio State University.

Jun, Sun-Ah. (1996). Influence of microprosody on macroprosody: a case of phrase initial strengthening. UCLA Working Papers in Phonetics 92:97-116.

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

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

________ (Submitted). Tonogenesis in Seoul Korean: a corpus study.

Karlgren, Bernhard. (1966). Analytic dictionary of Chinese and Sino-Japanese. Paris: P. Geuthner.

Kim, Midam. (2004). Correlation between VOT and F0 in the perception of Korean stops and affricates. In proceedings of the 8th international conference on spoken language processing (INTERSPEECH-2004), Jeju Island, Korea: International Speech Communication Association. 49–52.

Kim, Mi-Ryoung. (2000). Segmental and tonal interactions in English and Korean: a phonetic and phonological study. Doctoral dissertation, University of Michigan, Ann Arbor.

Kim, Mi-Ryoung. (2008). “Lax” stops in Korean revisited: VOT neutralization. Studies in Phonetics, Phonology and Morphology 14:185-203.

Kim, Mi-Ran Cho. (1994). Acoustic characteristics of Korean stops and perception of English stop consonants. Doctoral dissertation, University of Wisconsin-Madison.

Kim, Mi-Ryoung, Beddor, Patrice Speetor, & Horrocks, Julie. (2002). The contribution of consonantal and vocalic information to the perception of Korean initial stops. Journal of Phonetics 30:77-100.

Kim, Minjung, & Stoel-Gammon, Carol. (2009). The acquisition of Korean word-initial stops. Journal of the Acoustical Society of America 125: 3950–3961.

Kingston, John. (2005). Athabaskan tonogenesis. In Sharon Hargus and Keren Rice (Eds.), Athabaskan Prosody, John Benjamins.

Kingston, John. (2011). Tonogenesis. In M. van Oostendorp, C. J. Ewen, E. Hume, & K. Rice (Eds.), The Blackwell Companion to Phonology. Blackwell Publishing: Blackwell Reference Online. http://www.companiontophonology.com/subscriber/tocnode?id=g9781405184236_chunk_g978140518423699

Kong, Eun Jong, Beckman, Mary E., & Edwards, Jan. (2011). Why are Korean tense stops acquired so early?: The role of acoustic properties. Journal of Phonetics 39:196–211.

Koyang Cultural Foundation. (2004). Our language, Seoul language, Koyang language from the 1930s [CD]. Koyang Cultural Foundation.

Labov, William. (2001). Principles of linguistic change, vol.2: Social factors. Oxford: Blackwell.

Lee, Sue Ann S., & Iverson, Gregory K. (2012). Stop consonant productions of Korean–English bilingual children. Bilingualism: Language and Cognition 15:257-287.

Lee, Hyunjung, & Jongman, Allard. (2011). Perception of initial stops in tonal and non-tonal Korean. Poster presented at the 162nd Meeting of the Acoustical Society of America.

Lisker, Leigh, & Abramson, Arthur S. (1964). A cross-language study of voicing in initial stop: acoustic measurements. Word 20:384-422.

Liberman, Philip. (1963). Some effects of semantic and grammatical context on the production and perception of speech. Language and Speech 6:I72-I87

Maran, Laraw. (1973). On becoming a tone language: A Tibeto-Burman model of tonogenesis. In Hyman, Larry (ed.), Consonant Types and Tone, Southern California Occasional Papers in Linguistics 1:98-114.

Matisoff, James A. (1973). Tonogenesis in Southeast Asia. In Hyman, Larry (ed.), Consonant Types and Tone, Southern California Occasional Papers in Linguistics 1:73-95.

Mazaudon, Martine, & Michaud, Alexis. (2008). Tonal contrasts and initial consonants: a case study of Tamang, a ‘missing link’ in tonogenesis. Phonetica 65:231-256.

Oh, Eunjin. (2011). Effects of speaker gender on voice onset time in Korean stops. Journal of Phonetics 39:59-67.

Pearce, Mary. (2009). Kera tone and voicing interaction. Lingua 119: 846-864.

R Development Core Team. (2011). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. < http://www.r-project.org/>

Sankoff, Gillian. (2005). Cross-sectional and longitudinal studies in sociolinguistics. An international handbook of the science of language and society, vol.2, Ulrich Ammon, Norbert Dittmar, Klaus J. Mattheier, and Peter Trudgill (Eds.),. Berlin: Mouton de Gruyter. 1003-1013.

Sankoff, Gillian, & Blondeau, Hélène. (2007). Language change across the lifespan: /r/ in Montreal French. Language 83:560-588.

Silva, David J. (2002). Consonant aspiration in Korean: a retrospective. In Sang-Oak Lee & Gregory K. Iverson (Eds.) Pathways into Korean language and culture: essays in honor of Young-Key Kim-Renaud. Revised edn. Seoul: Pagijong Press. 447-469.

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

Svantesson, Jan-Olof, & House, David. (2006). Tone production, tone perception and Kammu tonogenesis. Phonology 23:309-333.

Tagliamonte, Sali A., & D'Arcy, Alexandra. (2009). Peaks beyond phonology: adolescence, incrementation, and language change. Language 85(1):58-108.

The National Institute of the Korean Language. (2005). A Speech Corpus of Reading-Style Standard Korean [DVDs]. Seoul, Korea: The National Institute of the Korean Language.

Thurgood, Graham. (2002). Vietnamese and tonogenesis. Diachronica 19:333-363.

Whalen, D. H., & Levitt, Andrea G. (1995). The universality of intrinsic F0 of vowels. Journal of Phonetics 23:349-366.

Wright, Jonathan. (2007). Laryngeal contrast in Seoul Korean. Doctoral dissertation, University of Pennsylvania.

Zsiga, Elizabeth C. (2008). Modeling diachronic change in the Thai tonal space. In J. Tauberer, A. Eilam, and L. MacKenzie (Eds.) University of Pennsylvania Working Papers in Linguistics 14.1:395-408.

---

[i] Zsiga (2008) examined the change in the realization of contour tones in Vietnamese but not the emergence of tones themselves.

[ii] Studies on the stop production by children, ranging in age from 2 to 10 years, in Present Day Korean show that children produce a substantial overlap of VOT values between aspirated and lenis stops and more reliable difference in F0 (Kim and Stoel-Gammon,

[ii] Studies on the stop production by children, ranging in age from 2 to 10 years, in Present Day Korean show that children produce a substantial overlap of VOT values between aspirated and lenis stops and more reliable difference in F0 (Kim and Stoel-Gammon, 2009, Kong et al. 2011, Lee and Iverson, 2011) in line with younger adult speakers' speech pattern.

[iii] In this respect, it is interesting to note that in an interview with Mr. Kye-Whan Chung produced by Koyang Cultural Foundation included in the CD of textbook recordings, we

can see that he is a highly-educated and upwardly mobile person; he is an avid photographer and a life-long lover of classical music, well-known in classical music scene as a fixture at the *Seoul Arts Centre,* a premier cultural performance venue in Seoul.