

Tonogenesis in Seoul Korean and L3 Production of Korean stops by Cantonese-English Bilinguals

Xinran Ren, Peggy Mok

The Chinese University of Hong Kong

renxinr@link.cuhk.edu.hk, peggymok@cuhk.edu.hk

Abstract

This paper aims to investigate how tonogenesis in Seoul Korean affects non-native language acquisition in a multilingual context with L1 Cantonese, L2 English and L3 Korean. We are also interested to examine the source and direction of cross-linguistic influences in the acquisition of VOT and F0 patterns which are used for the stop contrasts. Sixteen L3 Korean learners differing in proficiencies (beginning vs. intermediate) were recorded reading a list of words embedded in carrier sentences in the three languages. Results from acoustic analyses found that the learners did not merge VOTs for the Korean lax and aspirated stop contrast, but they showed the tendency for tonogenesis, such as their heavy use of f0 in producing different Korean stops. Furthermore, comparison between the beginners and the intermediate learners revealed various cross-linguistic influences: regressive transfer of f0 on L1 and L2, bidirectional interaction of f0 values between L2 and L3, and L1 influence on the production of VOT in L2 and L3. These findings suggest that L3 transfer happens in a cue-by-cue manner, i.e. VOT and f0 need to be accounted for separately.

Index Terms: tonogenesis, Seoul Korean, third language (L3) acquisition, VOT, cross-linguistic influence.

1. Introduction

Korean is well known to have a three-way contrast among voiceless stops: lax, tense and aspirated. Previous studies have shown a systematic VOT merger between the lax and the aspirated stops in Seoul Korean while pitch differences emerge in the following vowels. The trade-off relationship between VOT and pitch is an example of tonogenesis. As in Figure 1, speakers born before the 1950s contrasted lax, tense and aspirated stops only on the VOT dimension. For younger female speakers born after the 1950s, pitch difference emerged and VOT values between the lax and the aspirated stops began to overlap and totally overlapped in the female speakers born after the 1970s. However, younger male speakers still maintained some VOT differences, i.e. the VOT merger was not complete in them [1].

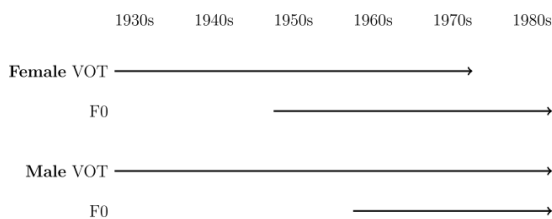


Figure 1: Timeline of the VOT merger and emergence of pitch difference. (Borrowed from [2])

Such tonogenesis should be taken into consideration in the acquisition study, otherwise the VOT merger between Korean lax and aspirated stops may be treated as a non-native-like pattern resulting from L1 interference, as is the case in some studies.

Previous studies on perception [3, 4] showed that it was very difficult for a speaker whose L1 only contains a two-way stop contrast to acquire the three-way contrast in Korean. The difficulty lies in the poorly discriminated Korean lax and aspirated contrast since both of them will be assimilated to a single L1 category. However, some production studies reported a wide range of patterns in stop categorization using VOT and f0. A study on the production of Korean stops by native Mandarin speakers [5] suggested that the experience with a tone language may not help the learning of f0 as a cue for stop contrast, probably because f0 activities in a tone language are largely constrained by the demands of the lexical tone system [6].

The current study investigated the case of L1 speakers of Cantonese, a tone language with a two-way stop contrast between short and long lag VOT, learning English as their second language and Korean as a third language. Based on the relevant studies, several predictions can be made about their production patterns of Korean stops. Along the VOT dimension, English and Korean contrasts will be mapped onto Cantonese unaspirated and aspirated stops accordingly. And Korean lax and aspirated stops will be similarly mapped on the Cantonese aspirated category. The tonogenesis in Korean seems 'beneficial' for the learners in that they can substitute similar categories in their native language for Korean stops in terms of VOT, and manipulation of f0 for the stop contrast seems not very difficult for native speakers of a tone language. However, this needs to be tested.

We then asked the following research questions: (1) Can the learners in this study produce different VOT and f0 values for different stop categories in a native-like way? (2) Do the learners' Korean stops present tonogenesis patterns? (3) How do the L3 beginning and intermediate learners differ in their production patterns using VOT and f0?

These questions are also interesting in the aspect of L3 phonological acquisition because previous studies mainly focus on the VOT dimension for stop contrast, with no study examining the cross-linguistic influences for other cues, e.g. f0, which are also used in stop categorization.

2. Method

2.1. Subjects

Sixteen native Cantonese university female students (with a mean age of 21.5 years) who have been learning English as their L2 and Korean as their L3 (eight at beginning level, eight at intermediate level) were recruited in the experiment. The beginning group finished 120 hours of formal classroom learning of Korean while the intermediate group finished 240. Moreover, 4 students in the intermediate group have participated in a 3-week intensive Korean language program at various universities in South Korea, and 2 students in the beginning group have reported a few years of study abroad experience in an English-speaking country.

The two groups started to learn English at around 2-3 years old at kindergarten and they did not differ from each other in IELTS scores: 6.78 for beginners, 6.89 for intermediate learners with $t=-0.36$, $p=0.72$. Although they learned English at a very early age, their English productions have a noticeable Hong Kong accent, at least some.

Besides English and Korean, they also speak Mandarin due to the Bi-literacy and Tri-lingualism language policy in Hong Kong. However, their Mandarin is not considered here, because Mandarin stop contrast is quite identical to Cantonese in that Mandarin also have unaspirated and aspirated stops and mainly uses the VOT cue for the contrast.

2.2. Materials

The subjects were asked to read a list of target words embedded in carrier phrases in all the three languages: L1 Cantonese, L2 English and L3 Korean. The simplest possible real words to the learners were chosen. For the Cantonese target words, words with Cantonese high level tone 1 (55) were used. The target words mostly consist of a CVCV sequence with the stops of all phonation types appearing in both word-initial and medial positions (English /stV/ included in the word-medial positions). Each target word begins with an alveolar stop followed by vowel /a/ and /i/. In total, each speaker produced 44 target words, which is twice (2 repetitions) the sum in Table 1. All the words were randomized and the order of languages tested were counterbalanced for each subject.

Table 1: *Number of target words for each stop × position × vowel combination.*

	Cantonese		English		Korean		
	/t/	/t ^h /	/d/	/t/	/t*/	/t/	/t ^h /
word-initial	2	2	2	2	2	2	2
word-medial	-	-	2	4	-	2	-
Carrier phrase	請讀 _ 比我聽。		Please say _ to me.		자, _ 해 봐.		

2.3. Acoustic measurements

VOT of the target alveolar stop and f0 of the following vowel were measured. VOT varies from positive (voicing lag) to negative (voicing lead) ranges. For intervocalic stops, voicing lead includes not only full voicing during the whole closure duration, but also partial voicing with both voiced and voiceless closure present. The partially voiced stops can be further divided into several voicing shapes, such as the classic negative VOT, as well as a ‘bleed’ pattern, which is defined as ‘voicing continues from the preceding sonorant and dissipates before the following stop release’ [7]. It is inappropriate to measure only the voicing closure as a negative VOT value in this case,

because by doing so, word-initial VOT is compared with word-medial voicing gap, rather than VOT. Therefore, the crucial thing here is to make a decision whether the bleed VOT is counted as positive or negative. We followed the criteria by [8] that VOT is considered negative if at least half of the closure duration is voiced, otherwise positive.

Fundamental frequency was measured as the mean over 15ms after the vowel onset following [5], using the ‘To Pitch...’ function in Praat. To make the automatic pitch tracking more accurate, the pitch range was set between 100-300Hz for the female speakers, on the premise that the speech samples all fall within this pitch range, because it is found that such pitch range setting can produce the closest results between careful hand analysis and automated batch processing [9]. If the speech samples fall outside this pitch range, the setting was adjusted according to each speaker’s individual pitch range.

Some data, especially the English productions, were omitted mostly due to creaky voice or production errors and variations: if the phonetic segment is no longer a stop, but is lenited to a flap, replaced by an affricate, or deleted in a consonant cluster, measurements should not be given.

2.4. Statistical analyses

Two types of analyses were performed on the production data. Firstly, in order to examine whether the learners can successfully use the VOT and the F0 values to categorize different stops separately in each language, we adopted within-subject ANOVAs or paired-samples t-tests (no p-value corrections for multiple comparisons) with each data point representing for each stop category by each subject. To this aim, the VOT and the f0 values were normalized using z-scores within each subject for each language separately.

The second analysis is for the comparison of VOT and F0 values used for stop contrasts in different languages. Therefore, the VOT and the f0 values were converted to z-scores within each subject’s productions across all the languages.

These analyses show how the multilingual speakers categorized the stop contrasts using VOT and F0. In addition, by comparing the VOT and the F0 values used by L3 beginners and intermediate learners, we can better understand the sources and the directions of cross-linguistic influences in third language phonological acquisition.

3. Results

3.1. Stop categorization

Before examining whether the L3 Korean learners present native-like tonogenesis patterns and how VOT and f0 values for stops transfer in a multilingual context, we should first find out how well the L3 learners, with different proficiencies, use these two cues to categorize stop contrasts in certain languages.

Regarding word-initial stops, as is shown in Figure 2, both beginning (left) and intermediate learners (right) produced different VOTs for different stop types in all the three languages. Compared with Cantonese and English, their Korean stop categories were more obviously differentiated by f0. It is also noteworthy that the intermediate group became more dependent on f0 for English stop contrasts than the beginning group, though there was still some overlap on f0 in the intermediate learners. This is confirmed by two t-tests examining whether f0 values are significantly different between the English /d/-/t/ contrast: for the intermediate, $t(14)=2.79$,

$p < .05$; for beginners, $t(14)=2.14$, $p=.07$. Regarding the Cantonese stops, the data show that both groups produced Cantonese aspirated stops with a significantly higher f_0 than the unaspirated ones, with $t(14)=3.01$, $p < .05$ for beginners and $t(14)=2.50$, $p < .05$ for the intermediate. This is inconsistent with the findings in [6] and [10] that the native Cantonese speakers produced the aspirated stops with an onset f_0 no higher than the unaspirated ones.

For the Korean aspirated-lax pair of interest, t-tests for VOT and f_0 in each group were performed: (VOT) beginners with $t(14)=4.40$, $p < .005$; the intermediate with $t(14)=3.30$, $p < .05$; (f_0) beginners with $t(14)=-5.86$, $p < .001$; the intermediate with $t(14)=4.13$, $p < .005$, which shows that these L3 learners, whether beginning or intermediate, were still using different values on both VOT and f_0 dimensions to produce the aspirated and the lax stops. This pattern is consistent with the traditional description of Korean three-way stops, but deviates from some recent studies reporting VOT merger for this pair.

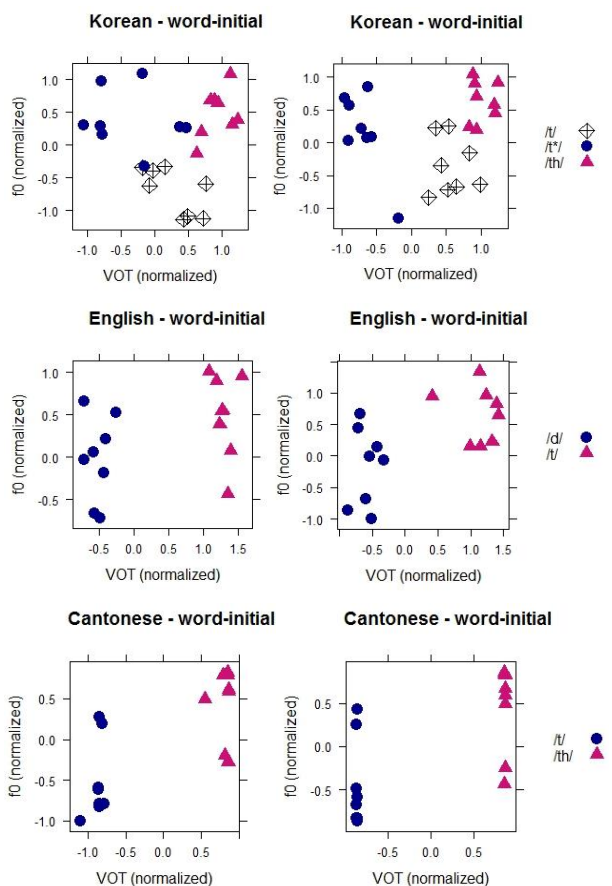


Figure 2: Korean, English, Cantonese – f_0 by VOT produced by beginners (left) and intermediate learners of Korean (right) in word-initial positions.

Since Cantonese stops and Korean tense and aspirated stops do not vary with different positions, we only examined English stops and Korean lax stops in word-medial positions. As can be seen from Figure 3, for the Korean medial /t/, the L3 Korean learners at beginning level produced a wider range of f_0 values while the intermediate had f_0 values clustering towards the lower end, which approximates the native pattern.

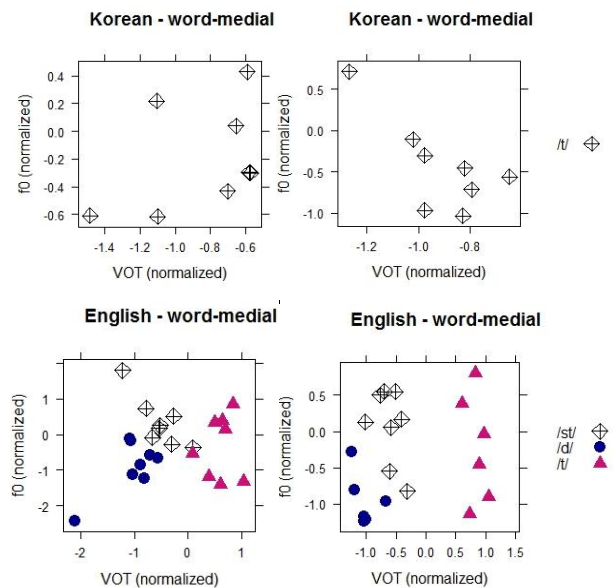


Figure 3: Korean, English – f_0 by VOT produced by beginners (left) and intermediate learners of Korean (right) in word-medial positions.

As for English medial stops, it shows that the beginners formed distinct categories on both VOT and f_0 dimensions for the intervocalic /t/ and /d/, as well as the /t/ after an /s/ sound in a consonant cluster, which is confirmed by corresponding statistical tests: $F(2,14)=27.84$, $p < .0001$ and $F(2,14)=6.09$, $p < .05$. As for the intermediate learners, they did not use f_0 as a cue to differentiate among these categories, with $F(2,14)=126.02$, $p < .0001$ for VOT but $F(2,10)=3.30$, $p=.08$ for f_0 . Importantly, it seems that the categorization of stops in English word-medial positions rely more heavily on VOT than f_0 , which is indicated by the effect size: in the beginners, $\eta_p^2 = 0.80$ for VOT larger than $\eta_p^2 = 0.47$ for f_0 ; in the intermediate learners, $\eta_p^2 = 0.95$ for VOT larger than $\eta_p^2 = 0.40$ for f_0 .

3.2. Comparison of L1, L2 and L3 categories

It was found out that generally, the speakers did quite good in categorizing stop contrasts within the three languages, using both VOT and f_0 . Thus, we can now examine the sound mappings in different languages by comparing VOT and f_0 values for different categories across languages.

3.2.1 F0

Since Cantonese is a tone language and f_0 patterns are heavily influenced by lexical tones, we did not include the Cantonese data here to compare with English and Korean.

In the following line chart, three patterns can be observed: firstly, comparing the two proficiency groups, the beginners tend to produce more distinct f_0 values between English and Korean stops while the intermediate learners made them closer by assimilating the Korean initial tense /t*/ with the English initial voiceless /t/ as well as the Korean medial /t/ with the Korean initial /t/, which is indicated by respective insignificant t-test results: $t(14)=0.12$, $p=.91$ and $t(14)=-0.53$, $p=.61$. Secondly, for the newly formed f_0 categories in the intermediate learners, i.e. word-initial Korean /t*/ and English /t/, they were formed the way that the originally distinct two values merged into their intermediate. Finally, f_0 difference for

the Korean initial stops was greater than that for the English initial stops, which may be related to the fact that f0 is a primary cue in Korean but a secondary cue in English.

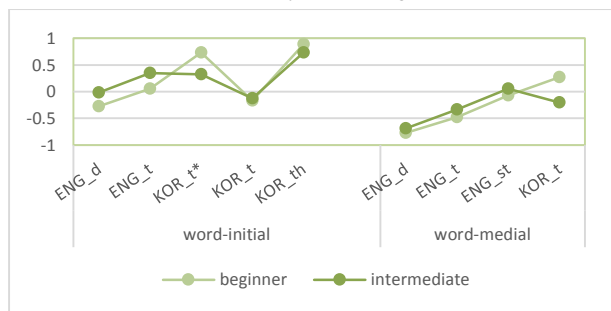


Figure 4: Mean values of normalized f0

3.2.2 VOT

Figure 5 shows that the VOT values used for the stop contrasts in the three languages did not differ much from the beginners to the intermediate learners. It seems that the word-initial English /d/, Korean /t*/ and the word-medial English /d/ and Korean /t/ were all mapped onto the L1 Cantonese unaspirated /t/, whereas the word-initial English /t/ and Korean /t^h/ as well as the word-medial English /t/ were all mapped on the Cantonese aspirated /t^h/. Moreover, the word-initial Korean lax stop /t/ and the medial English /t/ after /s/ were newly established categories in the phonetic space with an intermediate value between the Cantonese /t/ and /t^h/.

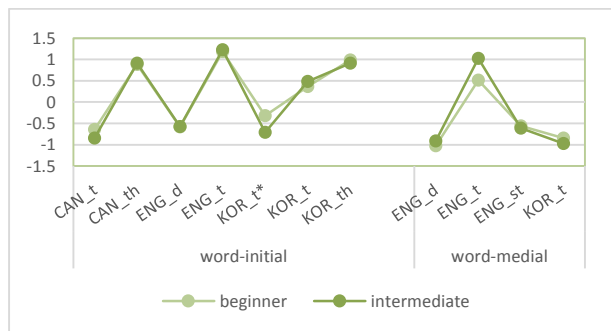


Figure 5: Mean values of normalized VOT

Concerning the influence of tonogenesis in Korean, the results here do not show a VOT merger between the Korean lax and aspirated stops, but rather a traditionally described short, intermediate and long lag VOT for tense, lax and aspirated stops respectively in the learners of L3 Korean.

4 Discussion

This study examines the influence of tonogenesis in Seoul Korean on L3 acquisition of Korean stop contrasts. Unlike the prediction, all the learners, even the beginners consistently used different VOT values, i.e. intermediate lag and long lag VOT, for the Korean initial lax and aspirated stops. This result is somewhat surprising because we know that three-way stop contrast on VOTs is difficult for native speakers of languages with only a two-way contrast to acquire. Considering both such acquisition difficulties and the changing input, it is reasonable to expect the VOT merger to happen. Possible explanations for this can be the classroom learning environment where teachers tend to adopt an 'authoritative' pronunciation. Following some

textbooks, the Korean teachers may still keep a distinctive VOT for each phonation type. Additionally, concerning the Korean teachers' demographics, the subjects have three full-time Korean teachers who are all females with one born between 1950 and 1970, one between 1970 and 1975, and the youngest one born after 1970. Accordingly, at least one of their teachers should still use different VOTs for the lax-aspirated stop contrast - that is to say, the learners were exposed to a mixed input in terms of the VOT merger.

Actually, the subjects did show the tendency for tonogenesis, such as their heavy use of f0 in producing different Korean stops compared with the English ones. Other evidence of tonogenesis comes from the cross-linguistic influences.

Cross-linguistic influences were observed with different sources and directions in this study. In terms of f0, firstly, in session 3.1, the intermediate Korean learners relied on f0 more than the beginners to produce the English word-initial stops. This indicates a regressive transfer from L3 Korean to L2 English because the subjects' L3 productions used f0 to categorize stops consistently and to a large extent. Besides, both groups produced a higher onset f0 for Cantonese aspirated stops. Such a reverse pattern compared with native Cantonese speakers not knowing Korean also suggests a regressive transfer from L3, because aspirated stops are realized with a higher onset f0 in Korean. Secondly, in session 3.2.1, bi-directional interaction is indicated by the pattern that f0 values of the word-initial Korean /t*/ and English /t/ moved closer to each other to their intermediate value as the subjects' language proficiency increased, which suggests an assimilatory process between their L2 and L3. In terms of VOT, all foreign categories were assimilated to either the Cantonese aspirated or unaspirated stops except that new categories were created for the Korean initial /t/ and the English medial /st/.

Comparing the L3 Korean beginning and intermediate learners, both groups performed well in categorizing different stop contrasts using both VOT and f0. However, the intermediate learners produced a more native-like Korean word-medial lax stop /t/ in that the f0 values they used were concentrated near the lower end of their individual pitch range, which corresponds to the low f0 of Korean word-medial lax stops. Conversely, the beginners were more native-like in the English initial stop production in terms of f0 use due to the intermediate learners' overuse of f0 in differentiating the two English stop categories, which deviates from native English speakers since f0 is only a very weak cue to the English stop contrast. Based on the above-mentioned differences between the two proficiency groups, it can be concluded that f0 learning happened in stop categorization as the speakers with L1 Cantonese and L2 English acquired a third language Korean which shows amplified f0 distinction as a result of tonogenesis.

To conclude, in examining the influence of tonogenesis in Seoul Korean on the L3 learners' production patterns, the results show that on the one hand, the learners were not largely influenced by this process since they still maintained VOT differences for the Korean lax and aspirated stop contrast. On the other hand, tonogenesis is reflected by their heavy use of f0 for this contrast and its transfer to the L2 English. The use of both VOT and f0 cues for the Korean lax-aspirated contrast seems to be a compromise response to the mixed input, which needs to be further investigated. Importantly, this study of L3 phonological acquisition demonstrates that even two phonetic cues for the same property can have different transfer patterns, arguing for a cue-by-cue L3 transfer.

5 References

- [1] Y. Kang, "Voice Onset Time merger and development of tonal contrast in Seoul Korean stops: A corpus study," *Journal of Phonetics*, vol. 45, pp. 76-90, 2014.
- [2] S. Cho, *DEVELOPMENT OF PITCH CONTRAST AND SEOUL KOREAN INTONATION* (Doctoral dissertation, University of Pennsylvania), p. 158, 2017
- [3] A. M. Schmidt, "Cross-language consonant identification: English and Korean," *Language experience in second language speech learning: In honor of James Emil Flege*, vol. 17, p. 185, 2007.
- [4] J. J. Holliday, "The perceptual assimilation of Korean obstruents by native Mandarin listeners," *The Journal of the Acoustical Society of America*, vol. 135, no. 3, pp. 1585-1595, 2014.
- [5] J. J. Holliday, "A longitudinal study of the second language acquisition of a three-way stop contrast," *Journal of Phonetics*, vol. 50, pp. 1-14, 2015.
- [6] A. L. Francis, V. Ciocca, V. K. M. Wong and J. K. L. Chan, "Is fundamental frequency a cue to aspiration in initial stops?," *The Journal of the Acoustical Society of America*, vol. 120, no. 5, pp. 2884-2895, 2006.
- [7] L. Davidson, "Variability in the implementation of voicing in American English obstruents," *Journal of Phonetics*, vol. 54, pp. 35-50, 2016.
- [8] A. S. Abramson and D. H. Whalen, "Voice Onset Time (VOT) at 50: Theoretical and practical issues in measuring voicing distinctions," *Journal of phonetics*, vol. 63, pp. 75-86, 2017.
- [9] A. P. Vogel, P. Maruff, P. J. Snyder and J. C. Mundt, "Standardization of pitch-range settings in voice acoustic analysis," *Behavior research methods*, vol. 41, no. 2, pp. 318-324, 2009.
- [10] Q. Luo, K. Durvasula and Y. H. Lin, "Inconsistent consonantal effects on f0 in Cantonese and Mandarin," in *Tonal Aspects of Languages 2016, May 24-27, Buffalo, New York, Proceedings*, 2016, pp. 52-55.