Shortening of Voice Onset Time in Korean Aspirates: Phonetic OCP or Timing Regulation?

Mira Oh and Dahee Kim

Chonnam National University and Kyungbuk National University
mroh@jnu.ac.kr and inureyes99@naver.com

ABSTRACT

The Voice Onset Times of word-initial Korean aspirated stops vary systematically as a function of the laryngeal feature of the onset of the following syllable. Precisely, VOT of a word-initial aspirated stop is shorter when the next syllable begins in an aspirated stop than when the next syllable begins in a lenis stop. The current study is aimed at testing two competing hypotheses to account for such VOT shortening in Korean aspirated stops, namely phonetic OCP[1, 2] or conforming to language-specific isochrony[3, 4]. VOT, f0, and spectral tilt were measured from productions of C_1 in /C_1VC_2Vta/, where the laryngeal properties of C_1 and C_2 were orthogonally varied among lenis, tense, and aspirated. Results suggest that isochrony can account for VOT shortening, while no phonetic dissimilation was found from phonetic measures other than VOT.

Keywords: VOT, Korean three-way contrast, phonetic OCP, interval, speech timing.

1. INTRODUCTION

Avoidance of similarity (e.g., OCP) and its relevance has been an important focus in the phonological literature, in particular, to account for phonotactic constraints against identical phonemes occurring within a certain phonological domain [1]. In monosyllabic native-Korean, aspirated consonants may not co-occur within a stem [5], and the author accounted for this laryngeal cooccurrence restriction as aspirated C_2 blocking aspirated C_1 in C_1VC_2 stems. Laryngeal co-occurrence restrictions are widely attested across languages [6, 7]. Gallagher [7] identified three patterns of laryngeal cooccurrence restrictions, namely assimilatory (tautomorphemic stops must have the same laryngeal feature), dissimilatory (stops of the same laryngeal feature do not occur within a root morpheme), and mixed patterns (homorganic stops must have the same laryngeal feature but non-homorganic stops must not have the same laryngeal feature). Avoidance of Asp-V-Asp sequences in native Korean monosyllabic stems would be one case of dissimilatory laryngeal cooccurrence.

However, avoidance of Aspirated stop(henceforth Asp)-V-Asp in native Korean stems may qualitatively differ from laryngeal cooccurrence restrictions found from other languages. First, the constraint against two consecutive aspirated stops is only found in one specific stratum of the lexicon. In other strata of the lexicon, a sequence of aspirate onsets within a monomorphemic word is widely found from other strata of the lexicon, such as loanwords (e.g., /k^hʌmp^hjut^hʌ/ "computer") or Sino-Korean words (e.g., /p^hok^hæŋ/ "violence"). In addition, the constraint against Asp-V-Asp is valid only for monosyllabic stems thus morphologically complex words that are of native Korean origin may still have an Asp-V-Asp sequence.

A recent study by Oh and Kim [2] addressed the question of how Asp-V-Asp sequences are phonetically realized. In particular, they focused on the marked status of Asp-V-Asp sequences in the native stratum of the Korean lexicon [5] and examined whether such sequences are "phonetically repaired." They found that Oejeol-initial Asps have shorter VOTs when followed by another Asp, suggesting that the shortening of VOT might suggest dissimilation or phonetically repairing a marked structure. In the current study, we argue that the shortening of VOTs in the first Asps in the Asp-V-Asp sequences can be better explained by an account appealing to the timing regularity in the language, rather than phonetic OCP.

The rest of the paper is organized as follows. In section 2, we reinterpret the data in [2] and provide pieces of evidence suggesting that dissimilation does not fully account for the phonetic variability of Korean stop production. In section 3, we provide an alternative account for the observed variability, and argue that the timing account is simpler and more comprehensive. Section 4 concludes the paper.

2. REVIEW OF OH AND KIM (2016)

Oh and Kim [2] had 18 Korean speakers (10 Seoul speakers and 8 Gwangju speakers) produce three-syllable nonsense words in carrier phrases, /iketto C_1aC_2ata/ or /iketto C_1alata/. C_1, and C_2 had same place and laryngeal features. They measured VOT of C_1, VOT of C_2 (Not discussed in [2]), f0 and spectral tilt of /a/ between C_1 and C_2. Both measures were

made at 10% point of the vowel. VOT of C_1 is summarized as Table 1 below.

Table 1: VOT of C_1 stops, data from [2].

Dialect	C_1	C_2	C_1 VOT (ms)
Seoul	Lenis	Lenis	62.73
		Lateral	76.36
	Aspirated	Aspirated	43.10
		Lateral	81.06
	Tense	Tense	20.04
		Lateral	25.12
Gwangju	Lenis	Lenis	59.37
		Lateral	76.01
	Aspirated	Aspirated	42.33
		Lateral	73.93
	Tense	Tense	20.88
		Lateral	28.02

As can be seen from Table 1 above, VOT of aspirated C_1 is shorter when C_2 was an aspirated stop than when C_2 was a lateral stop in both dialects. VOTs were submitted to a mixed-effects linear models where C_1 and C_2 were fixed effects and talker and place of articulation were random effects. A series of pairwise comparisons using TukeyHSD tests revealed that the ~33ms difference in VOTs of Asp C_1 between Asp(C_1)-V-Asp(C_2) and Asp(C_1)-lateral(C_2) was statistically significant at alpha level of 0.05 (β = -33.90; z = -4.25; $p < 0.001$), confirming that the VOT of Asp(C_1) is shorter before Asp(C_2) compared to lateral (C_2). In contrast, for lenis and tense stops, VOTs did not differ as a function of C_2.

Readers may wonder whether the observed pattern is due to the fact that C_1 and C_2 were identical and the repetition of the same aspirated stop (i.e., repetition of the identical segment, as opposed to the same laryngeal feature) was causing VOT shortening in Asp(C_1)s. In [2], the authors reported that they found a similar magnitude of VOT shortening from a spontaneous speech corpus of Seoul Korean, even C_1 and C_2 were not identical, suggesting that the effect is likely to be due to the repetition of laryngeal feature, not the repetition of an identical segment. Apparently, the "shortening" of C_1 VOT in Asp(C_1)-V-Asp(C_2) sequences seems to suggest that the dissimilation account in [2] is accurate. However, f0 and H2-H1 measured at 10% point of the vowel between C_1 and C_2 summarized in Table 2, may suggest otherwise.

A series of comparable mixed-effects models tested whether and to what extent f0 and H2-H1 of C_1 varied as a function of C_2. Interestingly, none of the two phonetic properties differed as a function of C_2, suggesting that no dissimilative effects were found. In other words, an OCP-style account (i.e., dissimilation) is burdened to explain why only VOT was affected by a nonlocal C_2 but no other phonetic properties were affected by C_2.

Table 2: f0 and spectral tilt of C_1 stops, data from [2]

Dialect	C_1	C_2	f0 (Hz)	H2-H1
Seoul	Lenis	Lenis	170	11.55
		Lateral	170	10.91
	Aspirated	Aspirated	227	11.36
		Lateral	232	11.10
	Tense	Tense	216	8.39
		Lateral	213	7.04
Gwangju	Lenis	Lenis	130	8.26
		Lateral	125	8.43
	Aspirated	Aspirated	183	9.49
		Lateral	186	10.65
	Tense	Tense	158	4.65
		Lateral	171	2.69

Yet an alternative explanation is possible. One possible explanation for this pattern is that the primary cue—f0—to laryngeal distinction in Korean [8] stays constant, while secondary cues vary as a function of other properties. Still, such an account does not predict why H2-H1, which is a secondary cue compared to f0, stays stable across the C_2 conditions. For these reasons, we argue that dissimilation does not provide a simple and comprehensive account for VOTs in Korean. In Section 3, we provide an alternative explanation for the reported variability.

3. ISOCHRONY IN SPOKEN LANGUAGES

In this section, we argue that an account based on the speech timing can better explain the VOT modulation found in [2] and summarized in section 2. First of all, the appeal to timing removes the theoretical burden to explain why VOT of C_1 was the only phonetic property affected by C_2. Second, independent evidence suggests that C_1 affects V_1 duration [9, 10] and C_2 affects V_1 duration in C_1-V_1-C_2-V_2-sequences [11]. If this is the case, C_2 may influence duration of V_1, and possibly, duration of C_1 as well. In order to examine whether C_1 and C_2 predicts duration of neighboring segments, we measured durations of /a/ between C_1 and C_2, closure duration of C_2, and VOT of C_2 from data generated in [2], Experiment 2 of which the participants were 10 Gwangju Korean speakers and the task was to read /iketto C_1aC_2ata/ or /iketto C_1alata/ three times in a random order, where C_1 and C_2 had same or different laryngeal features but had identical place feature. (i.e., thus, when C_1 was

/k/, there were four different types of stimuli: /kakata/,/kak'ata/,/kak^{h}ata/, and /kalata/.) Table 3 summarizes the duration of V1 and C2 closure; Table 4 summarizes VOTs of C2.

Table 3: Duration of V_1 and C_2 closure

C_1	C_2	V_1 (ms)	C_2 closure (ms)
Lenis	Lateral	89.27	NA
	Lenis	67.42	45.90
	Aspirated	52.86	111.29
	Tense	52.32	141.70
Aspirated	Lateral	85.59	NA
	Lenis	82.38	47.52
	Aspirated	46.45	103.71
	Tense	67.19	135.47
Tense	Lateral	140.10	NA
	Lenis	124.40	45.85
	Aspirated	108.44	111.34
	Tense	76.47	126.40

Duration of V_1 was found to differ systematically as a function of C_1. Vowels following lenis (M = 65.39 ms) and aspirated (M = 70.58 ms) stops were shorter than vowels following tense stops (M = 112.18 ms) and pairwise comparison revealed that all three levels were significantly different from each other at alpha level of 0.05.

V_1 duration also differed as a function of C_2. Vowels followed by laterals were, overall, longest (M = 104.97 ms); vowels followed by a lenis stop (M = 90.52 ms) were longer than vowels followed by aspirated (M = 69.12 ms) or tense (M = 65.21 ms) stops. Pairwise comparisons revealed that V_1 duration did not differ between when C_2 was aspirated and when C_2 was a tense stop, but all other differences were found to be statistically significant. These results confirm that aspirated and tense C_2s shorten the duration of preceding V_1 [11].

Variation in V_1 duration caused by neighboring segments, C_1 and C_2, suggests that the motivation for duration modulation in Korean is not limited the domain of a syllable. In the stimuli, C_1 and C_2 are onsets of two different syllables yet C_2 affected the duration of the vowel, which is the nucleus of a syllable preceding C_2. Moreover, limiting the domain to a syllable cannot account for why C_2 conditions VOT of C_1. Considering these observations, we also measured closure duration and VOTs of C_2.

Closure duration of C_2 was found to be conditioned by C_2 (M = 46.43 ms for lenis C_2; M = 108.80 ms for aspirated C_2; M = 134.57 ms for tense C_2; pairwise comparisons revealed significant differences across all levels) as well as C_1 (M = 100.15 ms for lenis C_1; M = 95.41 ms for aspirated C_1; M = 95.93 ms for tense C_1; pairwise comparisons revealed significant differences between tense and lenis C_1 conditions), suggesting that duration of C_1 influences that of C_2 vice versa. To further examine durational properties of C_2, as a function of C_1 and C_2,

Table 4: VOT of C_2

C_1	C_2	C_2 VOT (ms)
Lenis	Lenis	15.69
	Aspirated	62.30
	Tense	14.44
Aspirated	Lenis	15.97
	Aspirated	54.94
	Tense	15.74
Tense	Lenis	16.02
	Aspirated	69.01
	Tense	14.78

As can be seen from Table 4, VOT of aspirated C_2s differ as a function of C_1. Similarly to the shortening of aspirated C_1 VOT reported in [2], VOT of aspirated C_2 was also shorter after an aspirated C_1, suggesting that the influence of non-local Cs is bidirectional. Statistical analyses revealed that the VOTs of aspirated C_2 after an aspirated C_1 was significantly shorter than VOTs of aspirated C_2 after a lax C_1 (β = -7.20; z = -3.19; $p < 0.01$) and VOTs of aspirated C_2 after a tense C_1 (β = 13.82; z = 6.07; $p < 0.001$).

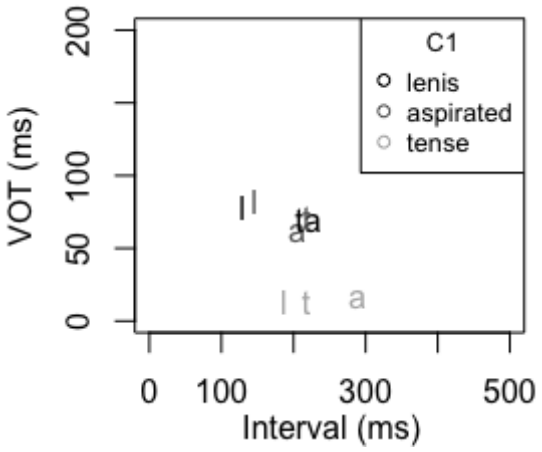
An account based on dissimilation cannot account for this bidirectional "shortening" of VOTs of aspirated stops without stipulation, since VOT of C_1 and C_2 eventually became similar to each other via shortening. One could potentially theorize the VOT patterns by first positing a VOT shortening, dissimilation rule of which the trigger is aspirated C2 and target is aspirated C1, which feeds assimilatory shortening of C2 VOT. However, such an account lacks a consistent motivation for duration modulation.

To summarize, we found that duration modulation in Korean stops and intervening vowels. To account for such modulation, a domain that is longer than a syllable is required, since onsets of one syllable condition duration of onset of a following syllable.

There have been proposals suggesting that vowel to vowel interval forms an integral unit in phonetics and phonology [3, 4]. For instance, [4] suggested that phonetic duration of vowel to vowel interval predicts stress placement in American English. Similarly, interval was found to be a better predictor of stress in Portuguese than syllables [12]. Although not much is known regarding the rhythmic unit of spoken Korean, there have been debates as to whether Korean is stress-time, mora-timed, or

syllable timed. Results are inconclusive, as different measures led to different conclusions [13, 14]. We focused on the shortening of aspirated C_1 VOT before tense and aspirated C_2, and hypothesized that V to V interval, which is the sum of V_1 duration, C_2 closure, and C_2 VOT may predict VOT of C_1. Figure 1 summarizes how V-to-V interval and VOT are related.

Figure 1: V-to-V interval and VOT of C_1. Letters on the plot corresponds to C2.



As can be seen from Figure 1, we find that the longer the V-to-V interval is, the shorter C_1 VOT is within the same C_1, suggesting that the inverse relationship might help V-to-V interval predict VOT of C_1. However, in order to evaluate whether this trend is a reliable predictor of C_1 VOT, a future study focusing on the variation of V-to-V interval (e.g., VOT of C_1 in C_1VCCV vs. C_1VCV) is needed.

4. CONCLUSIONS

In this study, we reexamined VOT shortening of Korean aspirated stops. Aspirated stops are typically considered as a stop category that is distinguished by long VOT and the long VOT of aspirated stops have been confirmed in productions of onset of monosyllables. However, in this study, we showed that in word initial positions, due to non-local C to C coarticulation or timing modulation, VOT of aspirated stops can be shorter than VOT of lenis stops. An account that is based on dissimilation cannot account for bidirectional shortening of VOTs of aspirated stops and does not predict why VOT, but not other phonetic properties, are modulated by laryngeal features of C_1 and C_2. In contrast, appealing to timing such as V-to-V interval predicts why VOT is modulated by C_1 and C_2. However, further research is needed to determine why such relationship between timing units (or some type of isochrony) is found from speech. We leave this question for future research. (We are in the process of examining the extent to which V-to-V interval predicts segment duration in read speech as well as in spontaneously spoken present day Seoul Korean.)

7. REFERENCES

[1] Liberman, M., Schultz, M., Hong, S., & Okeke, V. 1993. The phonetic interpretation of tone in Igbo. *Phonetica* 50(3), 147-160.

[2] Oh, M., & Kim, D. 2018. Long-distance effects on the VOT variation of Korean stops. Poster presented at the 16th Conference on Laboratory Phonology.

[3] Steriade, D. 2012. Intervals vs. syllables as units of linguistic rhythm. Handouts, EALING, Paris.

[4] Hirsch, A. 2014. What is the domain for weight computation: the syllable or the interval? *Proc. 2013 Annual Meetings on Phonology*, 1-12.

[5] Ito, C. 2007. Morpheme structure and co-occurrence restrictions in Korean monosyllabic stems. *Studies in Phonetics, Phonology and Morphology 13*, 373-394.

[6] MacEachern, M. 1999. *Laryngeal Cooccurrence Restrictions*. New York: Garland.

[7] Gallagher, G. 2010. Perceptual distinctness and long distance laryngeal restrictions. *Phonology* 27, 435-480.

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

[9] Pae, J., Shin, J., & Ko, D-H. 1999. Some acoustical aspects of Korean stops in various utterance positions, *Speech Science* 5, 139-159.

[10] Chung, H., Kim, K., & Huckvale, M. 1999. Consonantal and prosodic influences on Korean vowel duration. *Proc. of Eurospeech*, 1-4.

[11] Sohn, H S., & Ahn, M A. 2011. Voice onset time differences in Daegu Korean stops as a function of their position in words. *The Journal of Linguistics Science* 59, 149-180.

[12] Garcia, G. D. 2016. The computation of weight in Portuguese: syllables and intervals. *Proc. of the 33rd West Coast Conference on Formal Linguistics*, 137-145.

[13] Arvaniti, A. 2009. Rhythm, timing and the timing of rhythm. *Phonetica* 66, 46-63.

[14] Cho, M-H. 2004. Rhythm typology of Korean speech. *Cognitive Processing* 5, 249-253.