INTRA- AND INTERSEGMENTAL DURATIONAL COMPENSATION OF KOREAN PLOSIVES

Yeongeun Choi

Department of Computational Linguistics, University of Zurich
yeongeun.choi@uzh.ch

ABSTRACT

This study investigates the durational characteristics of three types of word-initial Korean plosives (fortis, lenis, aspirated), with a focus on the correlations between *Closure duration* (CD), *Voice onset time* (VOT), and the following vowel. Read speech of 15 native Korean speakers was recorded. The results show that (1) the VOT of aspirated plosives is as short as that of lenis plosives (even shorter in the case of velars), (2) CD plays a more important role in distinguishing Korean plosives than VOT, and (3) the durations of CD and VOT are dependent on one another in a complementary relationship. (4) In general, plosive duration influences vowel length (the longer the plosive, the shorter the vowel); particularly, (5) VOT is closely correlated with vowel length across place and manner of articulation. In sum, CD and VOT of Korean plosives influence each other and partially also the duration of the following vowel.

Keywords: Korean plosives, Closure duration, Voice onset time, Vowel duration, Durational compensation

1. INTRODUCTION

Plosives (henceforth also called 'stops') are characterized by two articulatory phases, with closure and release occurring consecutively before the onset of the following vowel. The time intervals of stop closure, known as *Closure duration* (CD), and of stop release, known as *Voice onset time* (VOT), vary depending on plosive type and place of articulation. The variation in the durations of CD and VOT arises in a complementary relationship, i.e. a long CD induces a short VOT and *vice versa*, and results in a limitation of the total duration of vocal fold opening [19, 4]. The durational adjustment occurs not only within a plosive, but also with an adjacent vowel. In particular, previous studies on pre-consonantal vowels in different languages have provided evidence that vowel length is determined by the relative length of neighboring consonants, i.e. longer vowels shorten the VOT of the following stops, e.g. in varieties of English [25], and also in the Central Bavarian dialect of German [9]. Conversely, durational compensation between vowels and preceding consonants seems to be less common.

The three types of homorganic plosives in Korean are all voiceless in word-initial position and are differentiated by the amount of [aspiration] as follows: fortis /p*, t*, k*/[1] (unaspirated), lenis /p, t, k/ (slightly aspirated), and aspirated /p^h, t^h, k^h/ (strongly aspirated). These properties lead to the differences in VOT, fortis with short-lag, lenis with rather longer, namely 'intermediate'-lag, and aspirated with long-lag [18, 6]. However, more recent studies (e.g. [23, 11, 3]) have found an overlap, commonly known as 'merger' of VOT between lenis and aspirated plosives, especially in the Seoul dialect, which is generally considered as Standard Korean. Regarding CD duration in Korean word-initial plosives, fortis plosives have the longest CD with a large duration overlap with aspirated plosives, which have the second longest CD, while lenis plosives have the shortest duration [16, 5].

Korean has been posited to lose its distinctive vowel quantity. Although vowel length distinctions are still preserved in a few homonyms, native Korean speakers barely distinguish between long and short vowels [20, 21] (see also [12, 24]). Despite the loss of phonemic vowel length contrasts, vowels before or after plosives appear to vary according to allophonic distribution. An electromyographic study [10] found that a preceding plosive affects the duration of the neighboring vowel more than a following plosive does, in terms of the voicing gesture: the larger the glottal opening area of stops, the larger the degree of devoicing of the following vowel. For instance, vowel devoicing was rarely observed after word-initial fortis stops associated with a very small glottal opening area (see also [7]). The voicing gesture may trigger a reduction in the actual time interval of the post-consonantal vowel, but this aspect was not examined in the current study.

Based on these insights, we hypothesize that the two basic acoustic cues of Korean word-initial plosives (i.e. CD and VOT) adjust their durations in a compensatory behavior, and are also related to the duration of post-consonantal vowels. In order to verify these hypotheses and to gain a better understanding of the relative durational behavior in Korean stops, a carefully designed corpus of read speech was collected from 15 native speakers of Korean (cf. § 2).

2. METHOD

2.1. Participants

15 native Korean speakers were recruited (13 female, 2 male; aged 20-29 years, mean 22). All subjects speak the Seoul dialect and had lived abroad for less than one year.

2.2. Speech materials

A reading task was conducted with a list of Korean interrogative sentences, which all begin with the interrogative adverb *eonje* 'when'. A word pair item 'A and B' was embedded in the carrier sentences, where the target word containing a word-initial plosive always appeared in the A position (e.g. /b/ in *eonje* ***baji****-wa chima-reul ibeoyo*? 'When do you wear **pants** and skirts?'). A total of 72 test sentences were created (3 places of articulation x 3 plosive types x 2 vowels x 4 word pair types). Additionally, 10 warm-up sentences with other initial consonants were presented to the participants at the beginning of the experiment. The warm-up sentences only contained actual Korean words.

The pairs of the target words were mixed with real and pseudo-words. Only disyllabic words (CVCV, CVCVC, CVCCV, and CVCCVC, with C indicating a consonant and V a vowel) were used to elicit a minimal pair-like speaking style with a maximally careful articulation. Only the vowels /a/ and /o/ followed the initial plosives of the target words. Overall, a total of 1,080 tokens (72 words x 15 subjects) were collected for analysis. One recording was excluded due to an excessive pause before the target word.

2.3. Experimental procedure

The experiment began with the warm-up sentences, followed by the sentences containing the target words, which were randomly displayed on a computer screen, following the procedure used by [15]. Normally, speakers read the given sentences aloud only once, but they could repeat them if they mispronounced a part of the sentence.

The production task was recorded using SpeechRecorder software [8] in soundproof recording facilities at the University of Zurich, Goethe-University Frankfurt, and the Acoustics Research Institute in Vienna. The following microphones were used for the recordings: a Røde NT2-A microphone in Zurich; a Røde Lavalier microphone attached to a Zoom H4n audio recorder in Frankfurt; and a AKG C451EB microphone in Vienna. The recordings were all sampled at 44.1 kHz and 16 bits.

2.4. Acoustic analysis

The three main acoustic cues, i.e. the CD and VOT of the initial plosive and the duration of its following vowel, were measured using Praat [2]. The following time intervals were all manually segmented: for CD, the interval from the voicing offset of the preceding vowel to the stop release; for VOT, the interval from the stop release burst to the voicing onset of the following vowel; and for the vowel, the interval from the onset to the offset of vocal cord vibration.

The durations of all target intervals (CD, VOT, vowel) were normalized by calculating z-scores to reduce the impact of the phonotactic context and to control for speech rate in the collected recordings.

2.5. Statistical analysis

Linear mixed-effects models were implemented in R [22] to estimate the complementary durational relationship between the main acoustic cues depending on their segmental context using the R packages *lme4* [1] and *lmerTest* [14]. The analysis of variance models included as dependent variables the z-scored CD duration (CD_{norm}), the z-scored VOT duration (VOT_{norm}), the z-scored plosive (CD plus VOT) duration ($Plosive_{norm}$), and the z-scored duration of the following vowel ($Vowel_{norm}$), respectively. All models used the fixed effects *plosive type* (lenis, fortis, aspirated), *place of articulation* (bilabial, alveolar, velar), *first syllable* (CV, CVC), and *vowel* (/a/, /o/), as well as the random effects *speaker* and *target word*, including both random intercepts and random slopes.

We report the results of type III F-tests and the corresponding p-values computed with the Satterthwaite's method, and corrected with the Tukey method, using the R package *emmeans* [17]. Correlations were calculated with Pearson's product-moment correlation.

3. RESULTS

3.1. Closure duration and Voice onset time

The results for CD_{norm} showed that both *plosive type* ($F[2, 71.4] = 227.3$, $p < .001$) and *place of articulation* (henceforth *PoA*, $F[2, 71.4] = 66.9$, $p < .001$) had highly significant effects on CD duration. The interaction of *plosive type***PoA* ($F[4, 71.4] = 4.3$, $p < .01$) also influenced the CD values. As shown in Figure 1, fortis plosives had the longest CD duration, followed by aspirated plosives, with lenis plosives having the shortest CD duration. Regarding *PoA*, CD values consistently decreased from bilabial to velar (a pattern which is well-known for VOT as well; cf. [9]).

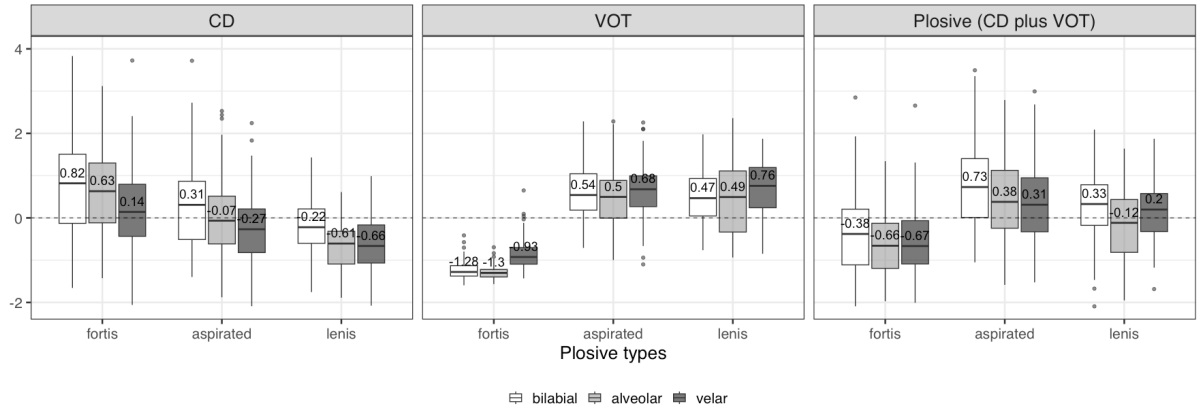

Figure 1: *CD*$_{norm}$, *VOT*$_{norm}$, and *Plosive*$_{norm}$, for fortis, aspirated and lenis plosives according to *PoA* (bilabial, alveolar, velar).

As concerns the three types of plosives, the VOT values showed exactly the opposite results. Fortis stops had very short absolute VOT values and lenis stops had relatively longer VOT values (see Table 1), consistently with previous findings (e.g. [13, 6, 23]). Aspirated stops, however, exhibited much shorter VOT values, which is inconsistent with previous reports (cf. ibid.).

	Absolute VOT (ms)	**Z-scored** VOT
fortis	18 ± 8	−1.13 ± 0.31
aspirated	61 ± 17	0.58 ± 0.67
lenis	61 ± 19	0.55 ± 0.73

Table 1: Mean and standard deviation for plosive types calculated from absolute VOT values (in milliseconds) and z-scored VOT values.

The results of pairwise comparisons in the model of *VOT*$_{norm}$ supported that the 'merger' of VOT between lenis and aspirated plosives was clearly present in all *PoA*s (see Table 2 and light gray boxplot in Figure 1). Additionally, lenis plosives had a relatively greater standard deviation than aspirated plosives, suggesting that some lenis obstruents were even longer than aspirated ones. This occurred primarily in the case of alveolar and velar plosives.

	estimate	SE	df	t.ratio	p-value
bilabial	0.09	0.13	72.0	0.70	0.76
alveolar	−0.03	0.13	73.4	−0.24	0.97
velar	−0.07	0.13	71.9	−0.57	0.84

Table 2: Pairwise comparison evaluation of *VOT*$_{norm}$ between aspirated and lenis plosives according to *PoA* (bilabial, alveolar, velar). Results are averaged for the levels *first syllable* and *vowel*.

Regarding *PoA*, VOT durations were longest for velar plosives compared to the other two more anterior places (bilabial-velar: estimate = −0.22, Standard error, henceforth SE = 0.08, df = 72.4, t = –2.89, $p < 0.05$; alveolar-velar: estimate = −0.29, SE = 0.07, df = 73.0, t = –3.91, $p < .001$) for all plosive types. The difference between bilabial and alveolar plosives was not significant (estimate = −0.07, SE = 0.07, df = 72.5, t = −0.96, p = 0.60; see also Figure 1).

3.2. Intrasegmental interaction within a plosive

The duration of the plosive phases changed steadily in an inverse relationship: a longer CD was followed by a shorter VOT and *vice versa* (see Figure 1). The result of the Pearson's correlation test supported the negative correlation between *CD*$_{norm}$ and *VOT*$_{norm}$ (r = −0.32, $p < .001$). This duration correlation was observed across *plosive type* and *PoA*, except for fortis and aspirated plosives at the alveolar place of articulation, where the shorter CD duration was not accompanied by longer VOT duration. Velar lenis plosives also displayed a marginally reduced CD duration compared to the alveolar lenis plosives, but they consistently maintained the 'short CD plus long VOT' pattern.

3.3. Intersegmental interaction between the target plosive and its following vowel

The target words were classified into two groups based on the first syllable: the CV group, embedding an initial consonant and a vowel, and the CVC group, containing an additional consonant in coda position. In both groups, the results of the regression analysis showed that *plosive types* had a significant effect on *Plosive*$_{norm}$ ($F[2, 71.9] = 92.8$, $p < .001$) and *Vowel*$_{norm}$ ($F[2, 70.6] = 132.3$, $p < .001$). Vowel duration was reduced in the order of fortis > lenis > aspirated, whereas plosive duration was marginally negatively correlated (r = −0.14, $p < .001$; see also Figure 2).

Figure 2 shows that the plosive and vowel durations were both shorter in the CVC syllables than in the CV syllables. In particular, the vowel duration was drastically reduced.

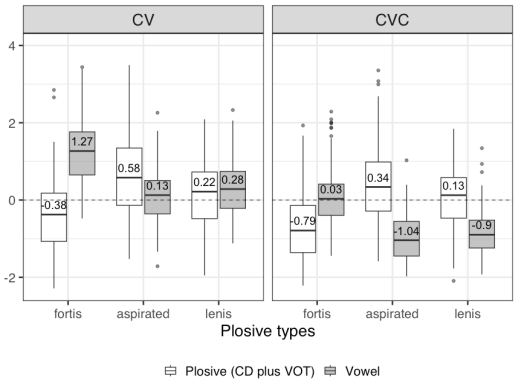

Figure 2: *Plosive*$_{norm}$ and *Vowel*$_{norm}$ for fortis, aspirated, and lenis plosives in the two different syllable types.

In contrast to the plosive type, the effect of *PoA* was less relevant to the interplay between *Plosive*$_{norm}$ and *Vowel*$_{norm}$. For aspirated and lenis plosives, the decrease in vowel duration values was proportionate to the reduced consonant length in both CV and CVC syllables (see Figure 3). Conversely, for fortis plosives, the variation in the duration of the two segments was irregular for both types of syllables on the basis of *PoA*.

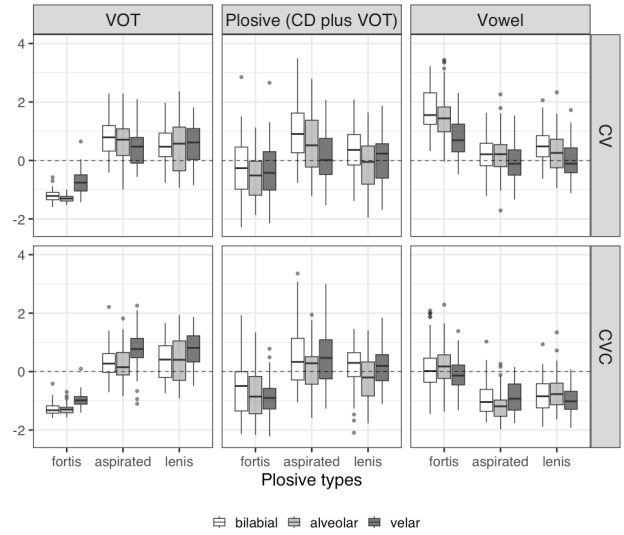

Figure 3: Durational variation of *VOT*$_{norm}$, *Plosive*$_{norm}$, and *Vowel*$_{norm}$, for fortis, aspirated, and lenis plosives according to *PoA* and syllable type.

Nevertheless, vowels gradually decreased in duration from the bilabial to the velar place of articulation of the preceding plosive (*PoA* effect on *Vowel*$_{norm}$: $F[2, 70.7] = 11.8$, $p < .001$), with the exception of alveolar fortis, velar aspirated, and alveolar lenis plosives in the CVC syllables. Similarly to the complementary relationship between CD and VOT (see § 3.2), this pattern was closely related to VOT alone in terms of durational compensation, as indicated above.

4. DISCUSSION

The present study investigated durational correlations within a plosive (intrasegmental) and between a plosive and a vowel (intersegmental) in word-initial plosives in Standard Korean. The findings on the acoustic cues of plosives and their internal and external interactions with adjacent segments can help shed more light on the acoustic properties of Korean plosives.

The results showed that (young) Korean speakers tend to produce aspirated obstruents with shorter VOT than reported in previous studies (e.g. [23, 11, 3]), especially for velar place of articulation. This points towards a merger between lenis and aspirated plosives and suggests that VOT alone does not seem to distinguish the different amount of aspiration any longer; rather the presence/absence of aspiration does, particularly when plosives are articulated in word-initial position (for word-medial plosives, see [10]). Instead, the lack of distinctiveness of VOT seems to be compensated by CD as an additional acoustic cue. This coherent pattern of variation in closure duration indicates that CD may play an essential role in discriminating lenis and aspirated plosives, regardless of place of articulation. As hypothesized, the lengthening of CD may serve as a means of adjusting the durations of CD and VOT in terms of a complementary relationship (see also [5, 6]).

Moreover, total plosive duration and vowel length were interdependent in that they were inversely related across plosive types. In particular, VOT considerably affected the duration of the following vowel: vowels were longer after unaspirated stops than after (slightly/strongly) aspirated stops. This variation in vowel duration is in line with a universal pattern of vowel length in CV or VC sequences [19, 4]. Furthermore, both directions of the transition from C to V or from V to C did not cause an unexpected pattern of vowel length variation. In other words, the adjacency between plosives and vowels itself plays a key role in adjusting the duration of target segments in a complementary relation.

5. CONCLUSION

In the present study, the three types of Korean plosives were clearly distinguishable in CD and VOT. These acoustic cues were closely interdependent in terms of a complementary adjustment of their durations, and this durational compensation showed a rather regular pattern across places of articulation. Moreover, vowel length distinction was highly interrelated with the presence/absence of aspiration.

6. ACKNOWLEDGMENTS

My sincere appreciation goes to my supervisors Stephan Schmid and Eva Maria Luef, as well as the two anonymous reviewers, for their valuable feedback. I am also grateful to Carolina Lins Machado and Franka Zebe for their assistance with statistical analyses. Any remaining errors are solely my own.

7. REFERENCES

[1] Bates, D., Maechler, M., Bolker, B., Walker, S. 2015. Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software* 67(1), 1–48.

[2] Boersma, P., Weenink, D. 2021. Praat: doing phonetics by computer [Computer program], version 6.1.52. http://www.praat.org/.

[3] Byun, H.-G. 2016. F0 as a primary cue for signaling word-initial stops of Seoul Korean. *Phonetics and Speech Sciences* 8(1), 25-36.

[4] Cho, T., Ladefoged, P. 1999. Variation and universals in VOT: evidence from 18 languages. *Journal of Phonetics* 27(2), 207-229.

[5] Cho, T., Keating, P. A. 2001. Articulatory and acoustic studies on domain-initial strengthening in Korean. *Journal of Phonetics* 29(2), 155-190.

[6] Cho, T., Jun, S.-A., Ladefoged, P. 2002. Acoustic and aerodynamic correlates of Korean stops and fricatives. *Journal of Phonetics* 30(2), 193-228.

[7] Choi, H. 2011. Vowel Duration as a Perceptual Cue for Preceding Stop Laryngeal Contrast in Korean. *Proceedings of the 17th International Congress of Phonetic Sciences* Hong Kong, 17-21.

[8] Draxler, C., Jänsch, K. 2004. SpeechRecorder: A universal platform independent multi-channel audio recording software. In: Lino, M. T., Xavier, M. F., Ferreira, F., Costa, R., Silva, R. (eds), *Proceedings of the IV international conference on language resources and evaluation* Lisbon, 559–562.

[9] Jochim, M., Kleber, F. 2022. Fast-Speech-Induced Hypoarticulation Does Not Considerably Affect the Diachronic Reversal of Complementary Length in Central Bavarian. *Language and Speech*, 1-35.

[10] Jun, S.-A., Beckman, M., Niimi, S., Tiede, M. 1998. Electromyographic Evidence for a Gestural-Overlap Analysis of Vowel Devoicing in Korean. *UCLA Working Papers in Phonetics*, 1-42.

[11] 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.

[12] Kang, Y., Yoon, T.-J., Han, S. 2015. Frequency effects on the vowel length contrast merger in Seoul Korean. *Laboratory Phonology* 6(3-4), 469–503.

[13] Kim, M. R. C. 1994. *Acoustic characteristics of Korean stops and perception of English stop consonants* (Ph.D. dissertation). Madison: University of Wisconsin-Madison.

[14] Kuznetsova, A., Brockhoff, P. B., Christensen, R. H. B. 2018. lmerTest: Tests in linear mixed effects models. R Package, version 3.1.3.

[15] Ladd, D. R., Schmid, S. 2018. Obstruent voicing effects on F0, but without voicing: Phonetic correlates of Swiss German lenis, fortis, and aspirated stops. *Journal of Phonetics* 71, 229-248.

[16] Lee, S. H. 1998. On the Effects of Places of Articulation of Stops on Their Closure Duration in Korean. *The Journal of the Acoustical Society of Korea* 17(6), 8-13.

[17] Lenth, R. V., Buerkner, P., Giné-Vázquez, I., Herve, M., Jung, M., Love, J., Miguez, F., Riebl, H., Singmann, H. 2022. emmeans: Estimated marginal means, aka least-squares means. R package, version 1.8.3.

[18] Lisker, L., Abramson, A. S. 1964. A cross-language study of voicing in initial stops: Acoustical measurements. *Word* 20(3), 384-422.

[19] Maddieson, I. 1997. Phonetic Universals. In: Laver, J., Hardcastle, W.J. (eds), *The handbook of phonetic sciences*. Oxford: Blackwell, 619-639.

[20] Magen, H. S., Blumstein, S. E. 1993. Effects of speaking rate on the vowel length distinction in Korean. *Journal of Phonetics* 21, 387-409.

[21] Park, J. W. 1994. Variation of vowel length in Korean. In: Y. K. Kim-Renaud (ed), *Theoretical issues in Korean linguistics*. Stanford, CA: Center for the Study of Language and Information, 175-187.

[22] R Core Team. 2022. R: A language and environment for statistical computing, version 4.2.2. http://www.R-project.org/.

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

[24] Shin, J., Kiaer, J., Cha, J. 2012. *The sounds of Korean*. Cambridge University Press.

[25] Tanner, J., Sonderegger, M., Stuart-Smith, J., Fruehwald, J. 2020. Toward "English" Phonetics: Variability in the Pre-consonantal Voicing Effect across English Dialects and Speakers. *Frontiers in Artificial Intelligence* 3, 38. https://doi.org/10.3389/frai.2020.00038

[1] The [+tense] feature of Korean fortis consonants is often marked with the diacritic /*/ (e.g. in [6]).