

Segmental Effects on the Accentual Phrase in Korean

Tae-Jin Yoon

Sungshin Women's University
tyoon@sungshin.ac.kr

ABSTRACT

The Accentual Phrase (AP) in Korean is said to be conditioned by the glottal status in the AP-initial segments, with the canonical shape of THLH. T is realized a H when the initial segment of the AP bears [+stiff vocal cords] and L when it bears [-stiff vocal cords]. The paper raises the question of whether the binary categorical distinction for the AP initial T is reflected upon phonetic realization of F0 (in semitone). The question is pursued by analyzing pitch contour shapes of canonical APs over word-like chunks in Korean reading-style corpus. The initial segments are divided into three groups of fortis obstruents, lenis obstruents, and sonorants. The results showed that the F0 is affected not only by the laryngeal status but also by the sonority difference between lenis obstruents and sonorants. The research raises an empirical question regarding the binary distinction of categorical prosodic events in the Tone and Break Indices (ToBI) system.

Keywords: Korean, Accentual Phrase, sonorants, F0(semitone), duration, K-ToBI.

1. INTRODUCTION

The glottal status of the initial segments is said to condition the tonal realization of Accentual Phrase (AP) in Korean. That is, the AP is demarcated by its tonal pattern THLH, where T represents either H or L depending on the status of the segmental feature [stiff vocal cords] in the syllable onset (Jun, 1996, 2000, among others). According to the theory of K-ToBI (Korean Tone and Break Indices), presence of the feature [stiff vocal cords] places fortis, aspirated obstruents, and fricatives (i.e., /p', t', k', ts', p^h, t^h, k^h, ts^h, s, s^h/) makes the initial tone in AP realized as H. The absence of the feature, on the other hand, groups the lenis obstruents /p, t, k, ts/ and sonorants together and makes the AP-initial tone L.

The codification of the tonal pattern for Accentual Phrases was originally proposed by Jun (1993). Jun (1993) noted asymmetric behaviour of F0, in that the F0 values after stops with [+stiff vocal cords] stops were significantly higher than those after lenis stops and sonorant consonants. Interestingly, the F0 difference was not due to the well-attested micro-prosodic effect. What distinguishes the pitch

perturbation effect in Korean from other languages was that while such micro-prosodic effect in other languages occurred during the transition between segments, lasting only for about 20-40ms after the vowel onset (Gandour 1974; Hombert 1978; Hombert et al., 1979), the f0 differences in Korean persisted until the end of the vowel (Jun, 1993). This observation led Jun (1993) to propose that the laryngeal effects was phonologically determined in Korean.

The paper aims to test whether the AP-initial segment brings about dichotomous F0 realization. Following the standard assumption in the ToBI tradition due to the influential work of Pierrehumbert (1990), previous literature divides the prosodic category into binary levels of H and L (Jun, 1993). Phonetic studies of duration and F0, on the other hand, cast doubt on binary AP-initial tonal patterns. A question of whether binarity of tones in the AP is empirically valid or not needs to be raised if we consider the interaction of F0 and segmental duration. It is possible that segmental duration and F0 may interact with each other. And the interaction may influence segments of different laryngeal features for their F0 realization. For example, the duration of domain-initial syllable is affected by the type of initial segments (Yun, 1998). When the initial segment is sonorants such as nasals, the initial syllable is noticeably lengthened. If the initial segment is, on the other hand, a voiceless consonant, the initial syllable is shortened. Laryngeally marked segments are expected to have shorter duration along with higher F0. Likewise, sonorant consonants are expected to have longer duration along with lower F0. It is also known that low-toned vowels are longer than high-toned vowels (Gandour, 1977), and vowels on rising or contour tones are longer than vowels on falling tones (Cho & Flemming, 2011, Grice Savino & Roettger, 2018). Similarly, Erikson & Alstermark (1972) observed that the pitch movement is often reduced when vowel length is decreased. Prieto & Ortega-Llebaria (2009) found that even among syllables in final-stress words, rise-falls induce longer duration than simple falls. This seems to be in line with the observation that rising or contour tones are more likely to result in longer vowels than falling tones.

Given the observation, we can think of a scenario in which a four-syllable AP with an initial vowel is

compared to a quadrisyllabic AP with a lenis stop. If the binary distinction is maintained, then we may not see any difference in the realization of F0 between the two initial segment types. On the other hand, if we assume the F0 is affected by the duration of the nucleus of an AP, then AP initial vowels will have lower F0 values than the vowel that occurs after a lenis stop, because the AP initial vowels are longer than the AP initial vowels occurring after a lenis stop. If the duration of tone-bearing unit affects the F0, then there is a possibility that the tone may not be restricted to a binary distinction. Thus, a question can be casted whether the initial tone is best be divided into a binary category, or it can have more detailed categories. In this paper, these two possibilities will be tested using a speech corpus.

2. DATA ANALYSIS

2.1. Data

Speech samples of female speakers are extracted from “The Reading-Style Speech Corpus of Standard Korean” (NIKL, 2005) and then used to test the hypothesis delineated in the previous section. Speakers in their 20’s are chosen for the analysis. To minimize the subjective encoding of prosody in the speech signal, four syllable words which were flanked by non-silent speech event are decided to be canonical Accentual Phrases in Seoul Korean. The decision is made based on the following observations: The shape of the AP tonal contour manifests a wide range of variation, with truncated two-tonal contours (e.g., LL, LH, HL, HH) to four-tonal contours (e.g., HHLH, LHLH, LHLL, and HHLL). This variation poses complications in speech data analyses. Thus, as is the case with the majority of previous studies (Jun & Lee, 1998, Yoon 2017, among others), a subset is chosen to limit the scope of analysis to F0 contours of the AP that manifests LHLH or HHLH. It is found that these two canonical tonal contours are observed in case the number of syllables is at least four. In the case of four syllable Accentual Phrases, the first tone is realized on the first syllable and the second tone is associated with the second syllable of the AP. And the last two tones are linked to the last two syllables (Jun, 1993; Kim, 2013; Jeon, 2011; Jeon & Nolan, 2017).

Phonetic analyses of speech samples were aided by the forced alignment system for Korean (Yoon & Kang, 2012; Yoon, 2015). The forced aligner takes as an input a pair of a sound file and its transcription in Korean, and generates word-like chunks (called *eojeol*) and phone annotations with time information as an output. With the phone- and word-aligned data, I calculated the number of phones in each word-like chunk, as well as the duration F0 (in semitone) of

each phone. Semitone is used instead of Hertz because Nolan (2003) reported that the psycho-acoustic scale better reflects listeners’ intuitions about melodic equivalence than Hertz.

In the corpus used in this experiment, three-syllable and four-syllable words are the two most frequent words, as in Yoon (2017). The token number for the trisyllabic words is 3,588, and the number for the quadrisyllabic words is 2,918. In this study, only four-syllable words were used, on the justification from the previous study that the canonical phrasal accent of THLH is realized on the quadrisyllabic words. In order to test the effect of segmental types on the initial shape, the extracted 4-syllable words were further divided into 4 groups (i.e., vowels, sonorants, lenis obstruents, and fortis obstruents). Fortis obstruents include tense and aspirated stops as well as /s/ and /h/, whereas lenis obstruents consist of plain stops. The number of tokens for each of the categories are as follows:

Table 1: Number of tokens per initial segmental type

Initial segmental type	Number of tokens
Vowels	996
Lenis Obstruents	627
Fortis Obstruents	785
Sonorants	420

3. RESULTS

In this section, I will present results that illustrate that individual speakers manifest their tonal patterns as LHLH when the first segment of the APs is not controlled. And then I will present the visual demonstration that the duration and F0 (in semitone) of the vocalic portion show an inverse relationship for different initial segment types. Finally, results of a statistical analysis will be shown that states that the initial segment types indeed contribute to the modelling of the AP tonal contours.

Visual assessment was first done to see whether there is speaker-dependent variation. Figure 1, which is replicated from Yoon (2017), illustrates mean F0 contours of the presumed AP patterns plotted for each of the 17 speakers. The overall tonal pattern of the four syllable words appeared to confirm that the canonical AP contour was LHLH, despite inter-speaker variations. Note that the type of initial segments is not controlled.

The tonal pattern of the quadrisyllabic words were analysed using third-order orthogonal polynomials. The analysis was carried out in R version 3.02 (R Development Team, 2009), using the lme4 packages (Bates et al., 2015). The results showed that even though linear and quadratic polynomials did not show any statistical significance ($p > 0.26$ and $p > .12$,

respectively), the cubic polynomial term turned out to be significant ($p < .001$), verifying that the underlying tonal pattern could be better modelled including terms for two tonal turning points.

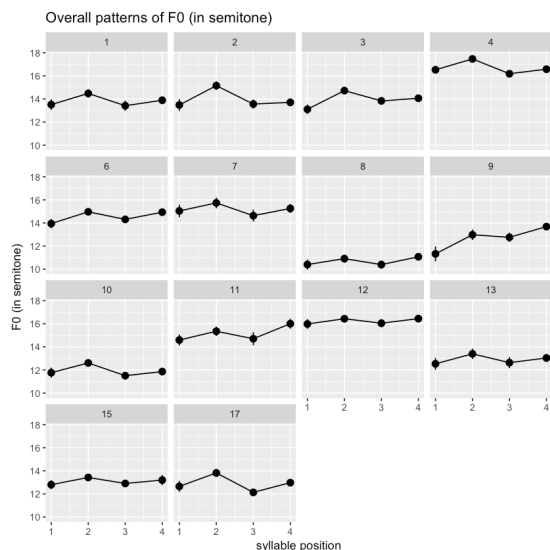


Figure 1: Segmental effect of F0 (Read speech)

With the visual and statistical confirmation that the quadrisyllabic words in the data at hand showed the canonical AP shape of LHLH, I now turn to the question of whether the initial segment type will be best modelled using only binary distinction of L and H. Visually check was made to see the relationship between duration and F0 (in semitone) for initial segmental types. If binary distinction is to be maintained, the tonal shape won't be affected by the duration of the AP-initial nucleus. On the other hand, if the longer nucleus duration is to have lower F0, and shorter nucleus duration higher F0, the height of F0 may be affected by the duration of the AP initial nucleus conditioned by the initial segmental type.

Figure 2 and Figure 3 illustrate that the duration and F0 of the AP-initial nucleus conditioned by the AP-initial segmental types. As apparent in the two figures, the duration of nucleus is inversely related to the F0 depending on the type of the initial segment categories.

Growth curve analyses (Mirman, 2014) were again conducted to analyse the tonal pattern of the four syllable words conditioned by initial segmental type. The overall tonal curves were modelled with third-order orthogonal polynomials. Initial segment types on all syllable position terms were set to be fixed effects. The cubic orthogonal polynomial captured the overall time-dependent syllable position. Results indicated that the effect of the type of initial segment improved model fit ($\chi^2(2) = 209.18$, $p < 0.001$). The effect of the initial segment type on the cubic terms also improved model fit ($\chi^2(2) = 20.16$, $p < 0.001$). The analysis leads us to conclude that the

tonal pattern is affected by the segmental types that occupies the AP-initial position.

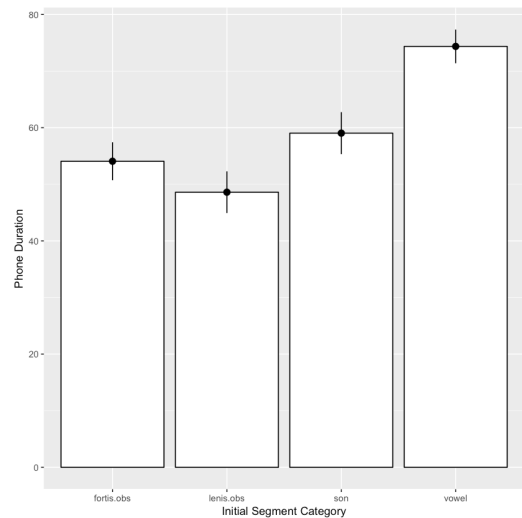


Figure 2: Segmental effects of Duration

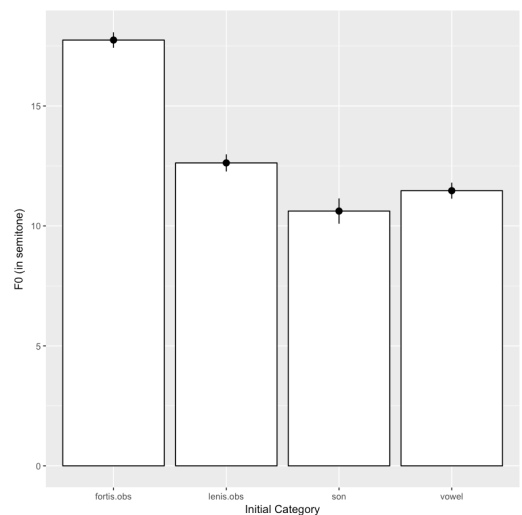


Figure 3: Segmental effects of F0 (in semitone)

4. DISCUSSIONS AND CONCLUSION

In this paper, a question was posed whether the type of the initial segments was shaped by a binary phonological distinction, using a corpus of naturally occurring speech corpora. Acoustic properties of canonical four-syllable APs were extracted from connected speech samples. The F0 values measured in semitone were extracted from the middle of each vocalic portion in the quadrisyllabic words. It is found that the initial segment type has an effect on the realization of both the duration and F0 on the vocalic portion of the AP-initial syllable.

The initial segments are divided into four groups of fortis obstruents, lenis obstruents, sonorants, and vowels. The results showed that the AP-initial segments are not only conditioned by the status of

[stiff vocal cords], but also by the coarsely-defined sonority (Clements, 1990). Vowels which occupy the highest in the sonority scale initiate the AP at the lowest. The fortis obstruents which have the least sonority value show the highest F0 points in the AP initial position. Sonorants and lenis obstruents position in the middle in the sonority scale, and also in the realization of F0 in the AP initial position. Thus, the analysis in the paper strongly implies that even if plain stops and sonorants belong to the same [-stiff vocal cords] group, the sonority difference between these two plays a role in giving rise to different F0 realization at the beginning of the APs in Korean.

Differences in initial tonal realization needs to be verified from perceptual domain and need to find its way to link with binarity-based prosodic modeling.

7. ACKNOWLEDGMENT

I thank the anonymous reviewers for constructive comments. This work was supported by the Ministry of Education of the Republic of Korea and the National Research Foundation of Korea (NRF-2018S1A5A2A011036570).

8. REFERENCES

- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, 67(1), 1-48. Retrieved from <https://www.jstatsoft.org/article/view/v067i01> on October 1, 2016.
- Cho, H., & Flemming, E. (2011). The phonetic specification of contour tones: the rising tone in Mandarin. *Proceedings of the 17th International Congress of Phonetic Sciences*. Hong Kong.
- Clements, G.N. (1990). The role of sonority in core syllabification. Kingston, John & Mary Beckman (eds.), *Papers in Laboratory Phonology I: Between the grammar and physics of speech*, Cambridge University Press, 245–306.
- Erikson, Y. & M. Alstermark. (1972). Fundamental frequency correlates of the grave word accent in Swedish: The effect of vowel duration. *STL-QPSR* 13, 53-60.
- Gandour, J. (1974). Consonant types and tone in Siamese. *Journal of Phonetics*, 2, 337-350.
- Grice, M., M. Savino, & T. Roettger. (2018). Word final schwa is driven by intonation – The case of Bari Italian. *Journal of the Acoustical Society of America*, 143, 2474-2486.
- Hombert, J. (1978). Consonant types, vowel quality, and tone. In V. Fromkin (Ed.), *Tone: A Linguistic Survey* (pp. 77-111). New York: Academic Press.
- Hombert, J., Ohala, J., & Ewan, W. (1979). Phonetic explanations for the development of tones. *Language*, 55(1), 37-58.
- Jeon, H. (2011). *Prosodic Phrasing in Seoul Korean: The Role of Pitch and Timing Cues*. Ph.D. Dissertation, University of Cambridge.
- Jeon, H., & Nolan, F. (2017). Prosodic Marking of Narrow Focus in Seoul Korean. *Laboratory Phonology: Journal of the Association for Laboratory Phonology*, 8(1), 2. Retrieved from <http://doi.org/10.5334/labphon.48> on March 20, 2017.
- Jun, S.-A. (1993). The phonetics and phonology of Korean prosody. Doctoral dissertation. The Ohio State University.
- Jun, S.-A. (1998). The Accentual Phrase in the Korean Prosodic Hierarchy. *Phonology*, 15(2), 189-226.
- Jun, S.-A. (2000). K-ToBI (Korean ToBI) labelling conventions (Ver. 3.1). *UCLA Working Papers in Phonetics*, 99, 149-173. Retrieved from <http://www.linguistics.ucla.edu/people/jun/ktobi/Ktobi.html> on April 2, 2015.
- Jun, S.-A., & Lee, H. (1998). Phonetic and phonological markers of contrastive focus in Korean. *Proceedings of the 5th International Conference on Spoken Language Processing* (pp. 1295-1298). Sydney, Australia.
- Kim, K. (2013). *Tone, pitch accent and intonation of Korean – A synchronic and diachronic view*. Ph.D. Dissertation, Universität zu Köln.
- Mirman, D. (2014). *Growth Curve Analysis and Visualization Using R*. Boca Raton, FL: CRC Press.
- NIKL (National Institute of the Korean Language) (2005). *A Corpus of Reading Style Seoul Korean [DVDs]*.
- Nolan, F. (2003). Intonational equivalence: an experimental evaluation of pitch scales. *Proceedings of the 15th International Congress of Phonetic Sciences*. Barcelona, Spain.
- Pierrehumbert, J. (1980). The phonology and phonetics of English Intonation. Doctoral Dissertation, MIT, Cambridge, MA.
- Prieto, P. & M. Ortega-Llebaria (2009). Do complex pitch gestures induce syllable lengthening in Catalan and Spanish. In *Phonetics and Phonology: Interactions and Interrelations*, eds. By M. Vigarrio, S. Frota, and M. Freitas. pp. 51-70, Benjamins, Amsterdam.
- R Development Core Team (2009). *R: A Language and Environment for Statistical Computing*. R for Statistical Computing, Vienna, Austria.
- Yoon, T. (2017). Growth curve modelling of nucleus F0 on Korean accentual phrase. *Malsori and Phonetic Sciences*, 9(3), 17-23.
- Yoon, T. (2015). Corpus-based study of duration adjustment in Korean. *Studies in Phonetics, Phonology and Morphology*, 21(2), 279-295.
- Yoon, T., & Kang, Y. (2012). A forced-alignment-based study of declarative sentence-ending ‘da’ in Korean. *The proceedings of 2012 Speech Prosody*.
- Yun, I. (1998). *A Study of Timing in Korean Speech*. Ph.D. Dissertation, University of Reading.