ISSN 2005-8063
2017. 9. 30.
Vol.9 No.3
pp. 17-23

말소리와 음성과학
Phonetics and Speech Sciences

한국음성학회지

http://dx.doi.org/10.13064/KSSS.2017.9.3.017

Growth curve modeling of nucleus F0 on Korean accentual phrase*

Tae-Jin Yoon**

Abstract

The present study investigates the effect of Accentual Phrase on F0 using a subset of large-scale corpus of Seoul Korean. Four syllable words which were neither preceded nor followed by silent pauses were presumed to be canonical exemplars of Accentual Phrases in Korean. These four syllable words were extracted from female speakers' speech samples. Growth curve analyses, combination of regression and polynomial curve fitting, were applied to the four syllable words. Four syllable words were divided into four groups depending on the categorical status of the initial segment: voiceless obstruents, voiced obstruents, sonorants, and vowels. Results of growth curve analyses indicate that initial segment types have an effect on the F0 (in semitone) in the nucleus of the initial syllable, and the cubic polynomial term revealed that some of the medial low tones in the 4 syllable words may be guided by the principle of contrast maximization, while others may be governed by the principle of ease of articulation.

Keywords: Korean, Growth Curve Analysis, Accentual Phrase, Initial segment, F0, Contrast maximization, Corpus phonetics

1. Introduction

According to Jun (1996, 1998), Korean is classified as a language without any head prominence related to lexical stress, pitch accent or tone, but a language with edge prominence. As a language with edge prominence, Seoul Korean has the hierarchically organized intonational structure with more than one word forming an Accentual Phrase, and more than one Accentual Phrase, in turn, forming an Intonational Phrase. In Jun's model, the Accentual Phrase (AP) is the smallest unit demarcated by the pitch contour. An AP tends to be approximately word-sized (Jun, 1998) and has 3-4 syllables and 1.14-1.2 content words on average (Jun & Fugeron, 2000; Jeon, 2011).

As a distinct property, the Accentual Phrase in Seoul Korean has drawn wide discussion in the previous literature, including the phonetic effect of the AP-initial segment. Jun (1996) observed that F0 after aspirated and tense stops were significantly higher (in average 50-80Hz) than that after lenis stops and sonorant consonants, and these f0 differences persisted until the end of the vowel. That is, segmental effects on f0, or microprosody, have been well attested over various languages (e.g., Gandour, 1974; Hombert, 1978; Hombert *et al.*, 1979) in such a way that the f0 at vowel onset is higher after a voiceless consonant but is lower after a voiced consonant. Such microprosodic effects are known to occur during the transition between segments, lasting for about 20–40 ms after the vowel onset. However, in Korean, the segmental effects persist over the whole syllable, and the degree of f0 change is much greater (on average about 80 Hz in female speech) than those found in other languages (Jun, 1996). Based on the findings, Jun (1996) drew a conclusion that the effects of laryngeal status on F0 is not phonetically driven but prosodically controlled in Korean. Moreover, she further suggests that the realization of the initial tone tends to depend on the laryngeal feature of the phrase initial segment. That is, the initial tone is H when the initial segment is a

* This work was supported by the Sungshin University Research Grant of 2016.
** Sungshin University, tyoon@sungshin.ac.kr
Received 5 June 2017; Revised 12 August 2017; Accepted 18 August 2017

tense or aspirated consonant, or a fricative (/s, s*, h/), and it is L in other cases (i.e. a plain consonant, a sonorant, or a vowel).

In this paper, we pose the first research question that concerns the effect of an initial segment on the F0 of the AP initial syllable. Even though the previous literature divides the prosodic category into two levels of H and L in Seoul Korean by the laryngeal features of the first segment, phonetic consideration makes us doubt regarding the relative phonetic (i.e., F0) realization among the laryngeal categories. Given findings in the previous literature, a possible interaction exists between segmental duration and F0, and the interaction may influence segments of different laryngeal features for their F0 realization. Yun (1998) found that there were effects of initial segment type on the duration of the domain-initial syllable. When the AP-initial segment was voiced (i.e., vowels and nasals), the initial syllable of an AP was noticeably lengthened, whereas when it was led by a voiceless consonant, the initial syllable underwent shortening. It is also reported that if all things are equal, vowels on rising tones or contour tones are longer than those on falling tones, and vowels on low tones are longer than those of high tones (Cho & Flemming, 2011). In other words, longer vowels are more likely to be observed on rising or contour tones than on falling tones. Also, longer vowels are to be seen on the low tones than on the high tone.

One of the consequences this observation between duration and tone makes is as follows: If we assume that an AP is four-syllable in size, and the initial and final syllables of the AP consist of a vowel, then we may expect that the duration of the initial vowel will be longer than the final vowel. This asymmetry of vowel length is because the canonical tonal pattern of AP is LHLH. Low tones are longer than high tones. This hypothesis seems to be borne out, given the finding that the first segment of the Accentual Phrase is lengthened (Jun, 1996), but the last syllable of the Accentual Phrase is not lengthened or lengthened to a negligible extent (Jun, 2000; Yoon, 2015). The absence or negligible effect of lengthening is a unique characteristic of the Accentual Phrase in Korean, even though it is a general finding that final lengthening of large prosodic domains such as the IP or the Utterance is consistently observed across a wide range of languages including Seoul Korean. Given the discussion so far, the phonetic remnants resulting from the interaction of duration and F0 may be phonologized in the shaping of the Accentual Phrase in Seoul Korean.

Let us consider when the initial segment of an AP is a lenis stop vs. a vowel. The vowel following a lenis stop is expected to be shorter than the initial vowel. Thus, we can expect that the initial syllable with lenis obstruents as the initial segment will be realized as an L tone, and the L tone may have lower F0 values than vowels realized as the same L tone. Therefore, the first research question is to examine the F0 when it occurs in the initial segment of the initial syllable of a canonical AP in Seoul Korean with different laryngeal features.

The second question is whether a medial L in the canonical AP is affected by the principle of contrast maximization or ease of articulation (Lindblom, 1990). Jun (1996) suggests that the AP has the underlying tonal pattern of THLH, and the realization of the initial tone (T) tends to depend on the laryngeal feature of the phrase initial segment. Jun (2000) reported 14 pitch contours of the AP (i.e., LH, LHH, LLH, HLH, HH, HL, LHL, HHL, HLL, LL, HHLH, LHLH, LHLL, and HHLL). The last tone, either L or H, is associated with the AP-final syllable. Even though it is canonical to end the AP with H, the possibility of having L on the AP final syllable is not excluded. Variation of AP contours is determined by the number of syllables in an AP. Jun (1998) states that all four tones are realized when there are four or more syllables in an AP, but if there are fewer than four syllables, some of the tones are undershot. When an AP is shorter than four syllables, the tones in the middle are deleted. To make matters simple, let us consider the canonical LHLH pattern of AP. In such a case, we observe that the medial L tone is surrounded by H tones. Previous studies state that the last two tones (i.e. LH) are realized on the last two syllables, and the F0 contour is interpolated across the syllables between the first two syllables and the last two syllables, when the AP has 4 or more syllables (Jun, 2000; Jun & Lee, 1998). However it is not clear how low the medial L can be. If the tonal contrast between medial L and the surrounding H is to be maximized, the medial tone can be as low as or lower than the initial L tone. If the medial tone is not realized in such a way, but is influenced by the flanking H tones, it can be expected that the F0 of the medial tone may be higher than the initial L tone, probably due to the principle of articulatory ease. Thus, the second goal of the paper is to examine the F0 of the L tone in the presumed AP patterns of LHLH, by applying to a large-scale speech corpus growth curve analyses, which are ways to quantify and assess the shapes of time dependent curves of nested data (Mirman, 2014: 3).

In this paper, we will address these two research questions using a large corpus of speech samples taken from female speakers commanding Seoul Korean. Because F0 values are known to differ between male and female speakers, speech samples of only one gender are used.

2. Data Analysis

2.1. Data

The data for this study come from a subset of The Reading-Style Speech Corpus of Standard Korean (The National Institute of the Korean Language 2005), a corpus collected from speakers of Korean residing in the Seoul metropolitan area. For the acoustic analysis, we chose a subset of the data which contains 17 speakers whose age is in their 20's at the time of recording in 2003. Each of the speakers read nine well-known fairy tales, short stories, or essays listed in <Table 1>. In <Table 1>, the third column lists the number of sentences (S. num.), and the last column lists the articulation duration, that is, the duration of speech samples excluding silent pauses, read by one speaker.

The acoustic analysis of speech segments was aided by the forced alignment system for Korean (Yoon & Kang, 2012; Yoon, 2015). The forced aligner takes a sound file and its Korean transcription as input and automatically generates word and phone annotations with time information. With the phone- and word-aligned data, we extracted features of F0 (in semitone) that can aid us in finding answers to our research questions. Semitone scale is a logarithmic transformation of the physical Hertz. Semitone instead of Hertz is used an a unit of measurement, based on Nolan's (2003) experimental report that among a number of widely used psycho-acoustic scales (i.e., semitone, mels, Bark and ERB-rate) and Hertz, semitone best reflects subjects' intuitions about melodic equivalence.

Table 1. Contents of the Reading Style Corpus of Standard Korean (Kang *et al.*, 2015).

	Reading materials	S. num	Duration
1	호랑이와 곶감(The tiger and the dried persimmon)	42	3m7s
2	해님 달님(The sun and the moon)	28	2m07s
3	그리운 시내가(Longing for the stream bank)	36	5m1s
4	광화문 지하도 아주머니(An elderly lady in the Kwanghwamun underground passage)	27	2m47s
5	막 지은 밥(Rice just cooked)	17	2m17s
6	눈 오던 날(The day when it snowed)	35	2m46s
7	숭늉의 지혜(The wisdom of sungnyung, scorched-rice water)	19	1m58s
8	까만 눈동자들 앞에서(In front of dark eyes)	27	3m43s
9.	내 고향 개울 물에서(In the brook in my hometown)	40	3m12s
	합계	930	26m59s

The approach we adopted in this paper is to single out APs from the corpus using findings from existing literature on the acoustic nature of APs. The approach is taken due to two main reasons. The first reason is that prosodic annotation demands subjective and manual methods. But the sheer amount of data makes it hard to conduct such a subjective and manual annotation of intonational structure (Xu, 2011). The second reason is to minimize subjective influences on prosodic annotation. That is, despite actual durational difference, listeners may perceive durational differences due to tonal differences. For example, in experiments in which sequences of digits were used, English, German, or French listeners perceived the syllable duration longer when the syllable had dynamic pitch compared to constant pitch in the syllable. The perceptual difference occurred despite the fact that there was no actual durational differences between the syllables (Cumming, 2010; Kohler, 2008). These may influence annotators on their decision on prosodic events.

To minimize the subjective encoding in the speech signal, we assumed that four syllable words which were both preceded and followed by non-silent speech event were canonical Accentual Phrases in Seoul Korean. If we extract these presumably prototypical Accentual Phrases, we would expect that the tonal contour would be THLH. The AP is mainly affected by the tonal patterns of THLH and silent pauses are rarely reported to be a cue signaling APs. Furthermore, even the same words in the same sentential position may be an AP in one instance, but not an AP in other instances. This observation is borne out when we examined the number of tokens extracted with the above-mentioned criteria. For example, the four-syllable word *horangineun* 'tiger-TOP' was extracted 363 times, the word *eomeoniga* 'mother-NOM' was extracted 41 times. Since 17 speakers uttered the same token, the total number of tokens must be multiples of 17. The fact that *eomeoniga* 'mother-NOM' was extracted 41 times (cf. 17*2=36, 17*3=51) means that some speakers realized the utterance as an intonational phrase rather than an Accentual Phrase by putting silence pause either before or after it. This indicates that the choice of prosodic boundary can differ from speaker to speaker.

2.2. Basic statistical description

Before we proceed in Section 3 to the analysis and modeling of the Accentual Phrase with regard to the acoustic features of F0 and the initial segment types, we will present some basic statistical description of the data used in the current study.

First of all, the number of vowels extracted from the corpus is 12,606, as summarized in <Table 2>.

Table 2. Number of vowel tokens in the corpus (n=12, 606)

Vowel	a	ae	e	eo	eu	i	o	u
Number of Tokens	3502	574	447	1425	2027	2396	1439	796

The number of syllables per word which is not preceded nor followed by silent pauses in the extracted data is also calculated, and listed in <Table 3>. The two most frequent number of syllables per word is 3 (n=3,588), and 4 is the second most frequent number of syllables a word consists of (n=2,918). The F0 contour shapes over the Accentual Phrases in Korean show a wide range of variation (including LH, LHH, LLH, HLH, HH, HL, LHL, HHL, HLL, LL, HHLH, LHLH, LHLL, and HHLL). This variation poses complications in speech data analyses and the majority of previous studies limited their scope of analysis to F0 contour of the Accentual Phrase (AP) that manifests LHLH or HHLH (Jun & Lee, 1998). In this study, we also singled out only words whose syllable length is 4 out of the tokens in <Table 3>.

Table 3. Number of tokens per the number of syllables for each word

# Syllable	1	2	3	**4**	5	6	7
# Tokens	370	1,681	3,588	**2,918**	1,799	629	148

In order to exclude potential outliers, the tokens whose duration exceeds 2.5 times the standard deviation were excluded, reducing the number to the ones listed in <Table 4>.

Table 4. Number of vowel tokens for the 4 syllable words (n=2828)

Vowel	a	ae	e	eo	eu	i	o	u
Number of Tokens	738	216	100	249	296	697	454	78

The extracted four-syllable words were divided into four groups on the basis of the status of initial segments: vowels, sonorants, lenis obstruents, fortis obstruents. Tense obstruents and /s/ and /h/ are grouped together with fortis stops under the label of fortis obstruents. Even though the F0 is known to be affected by the height of vowel, further division of vowels into high and low vowels is not considered in this study due to insufficient token numbers of each vowel type if vowels were divided into types. The categorical status of initial segments of these 4-syllable words and the number of tokens are listed in <Table 5>.

Table 5. Categorical status of initial segments (Total: 2,828)

Initial segment	Vowel	lenis Obs.	fortis Obs.	Sonorant
# Tokens	996	627	785	420

3. Results

3.1. F0 of the initial segments

The first question is to find out whether the initial segment category influences the F0 on the nucleus of the target syllable. <Figure 1> illustrates the F0 (measured in semitone) for each the initial segment category.

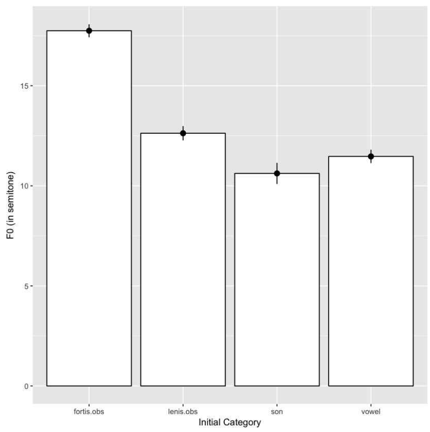

Figure 1. Box plot for the F0(semitone) for each category of initial segments

<Table 6> lists descriptive statistics of the four broadly-classified initial segments.

Table 6. Descriptive statistics of F0 (in semitone) for each initial segment category

	# tokens	mean	std. dev.
fortis obstruents	208	17.74	2.36
lenis obstruents	169	12.62	2.35
sonorants	135	10.62	3.09
vowels	244	11.46	2.65

One-way ANOVA was conducted with the dependent variable of F0 (in semitone) and the four levels of initial segment category (i.e., fortis obstruents, lenis obstruents, sonorant, and vowels) as the independent variable. The univariate analysis revealed a main effect of the initial segment categories on F0 $F(3, 752) = 294.9$, $p < .001$). Post-hoc test with Bonferroni correction showed that each pair of the initial categories was statistically different from each other (with all $p < .001$ except $p < .05$ between sonorants and vowels). The result indicates that F0 (in semitone) is in fact affected by the type of initial categories. This poses a problem in a phonological theory that collapses vowels, lenis obstruents, and sonorants to a single tonal category of L.

3.2. Growth curve analysis for underlying tonal pattern

To find out what underlying tonal pattern is revealed, the four-syllable words are analyzed using Growth curve analysis (Mirman, 2014). Growth curve analysis is used because groups of observations all come from the same source, that is, the same four-syllable word produced by the same speaker in our case. The acoustic properties of two randomly selected speakers may be uncorrelated, but the properties of a speaker at syllable *t* may be correlated with that speaker's phonetic values at syllable *t-1*. These nested observations are not independent and this non-independence needs to be taken into account in data analysis. Growth curve analysis is "a way to analyze nested data that takes the grouping into account and provides a way to quantify and assess the shapes of time dependent curves (Mirman, 2014: 3)." Due to the adequacy of growth curve analysis to our data that consist of time-dependent 4-syllable AP-like units, we carried out growth curve analyses in R version 3.0.2 using the lme4 package (version 1.1-10) (Bates *et al*., 2015). Results of growth curve analysis will be presented in the order of overall F0 patterns of the four syllable word, and then the F0 patterns conditioned by the type of initial segments.

The first step in statistical analysis is to assess the need to do it in the first place. If there is no significant variation across contexts in the first place, then there will be no need for including the context in the modeling. <Figure 2> illustrates the stylized F0 contour of the presumed AP patterns for each of the 17 speakers. It seems that there is speaker dependent variation, and also that the general pattern seems to be LHLH, nevertheless.

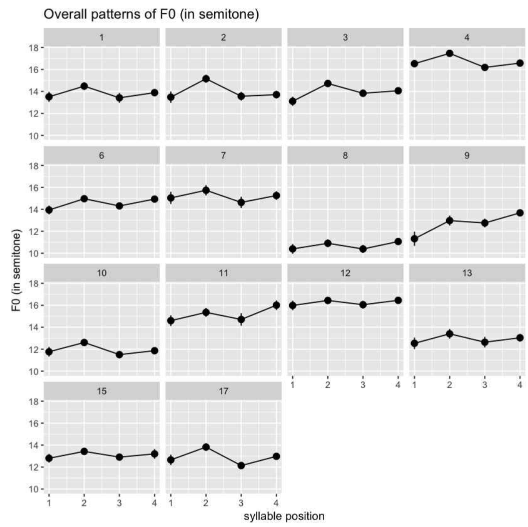

Figure 2. Stylized F0 contour of the presumed APs for each speaker

The following steps are taken in order to ascertain whether there is variation over contexts (i.e., speakers in our case): First, we fit a baseline model in which we include only the intercept; next, we fit a model that allows intercepts to vary over contexts (in the present, speakers); finally we compared these two models to see whether the fit had improved as a result of allowing intercepts to vary. The results indicate that the intercepts vary significantly across the different speakers($\chi^2(1)=691.554$, $p < .0001$). Because our model significantly improves, we modeled the variability in intercepts.

The next consideration in growth curve analysis is that the functional form must be adequate for the shape of the data. For a functional form to be adequate, it needs to be able to capture the shape of the observed data. As mentioned in the introduction, AP has the underlying tonal pattern THLH, with the realization of the initial tone (T) being dependent on the laryngeal feature of the phrase initial segment. The investigation of the four-syllable words in our corpus reveals that majority cases of fortis obstruents has /h/ as its initial segment (1463/1570), and this extremely skewed distribution of data may contribute to the general underlying tonal pattern of LHLH. Overall tonal patterns were modeled with third-order orthogonal polynomials. <Figure 3> illustrates the model fit to the F0 data. The (gray) dotted line reflects the fitted model of baseline. The (blue) dotted dashed line is for the fitted model of quadratic change. The (red) dotted line reflects the fitted model of cubic change. <Figure 3> demonstrates that the overall F0 contours corroborate the previous studies that the default intonational contour of the Accentual Phrase in Seoul Korean intonational phonology is LHLH. Note again that fortis obstruents are very infrequently observed in the corpus at hand.

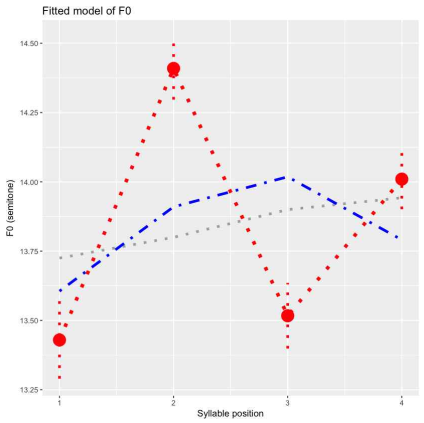

Figure 3. Observed data (red dots) and growth curve model fits (gray for linear, blue for quadratic, and red for cubic) for tonal patterns

The observation is verified by the following results of growth curve analysis. <Table 7> shows the fixed effect parameter estimates and their standard errors along with p-values estimated using the normal approximation for the t-values.

Table 7. Parameter Estimates for Analysis of F0 on Tonal Patterns

	Estimate	Std. Error	*t*	p
Intercept	13.866	0.424	32.696	<.0001***
Linear	0.151	0.131	1.146	.268
Quadratic	-0.166	0.109	-1.518	.129
Cubic	0.718	0.105	6.827	<.001***

The precise point of the peak alignment may vary. According to Jun (1996), the first tone (i.e. T in THLH) is realized on the first syllable of the AP, and the second tone (i.e. H) is loosely associated with the second syllable when an AP has more than four syllables (cf. Kim, 2013; Jeon & Nolan, 2017). On the other hand, de Jong (1994) showed that the AP-initial peak (which correspond to the second tonal target of the (T)HLH pattern) was aligned at the end of the AP-initial syllable. However, given the pattern in figure 4, what de Jong (1994) asserts on the anchoring of the peak on the AP-initial syllable does not seem to hold, as the F0 peak is realized on the second syllable position in <Figure 3>.

3.3. The F0 of medial L in LHLH tonal patterns

The second question is to look at the effect of medial L in the LHLH pattern. As a first step, growth curve analysis (Mirman, 2014) was used to analyze the tonal pattern of the four syllable words. The overall tonal curves were modeled with third-order orthogonal polynomials and fixed effects of initial segment types on all syllable position terms. The base model captures the overall time course with a third-order (cubic) orthogonal polynomial. The base model allows individual participants to vary randomly on any of the three components of the overall time course (intercept, linear, quadratic, and cubic). Model 0 is made by adding to the base model the fixed effect of initial segment category on the intercept. Model 1 is constructed by adding to model 0 fixed effect of initial segment category on the linear term. Model 2 added to model 1 fixed effect of initial segment category on the quadratic term. The final model, model 3 added to model 2 fixed effect of initial segment category on all terms.

Effects on model fit were evaluated using model comparisons. Improvements in model fit were evaluaflted using −2 times the change in log-likelihood, which is distributed as χ^2 with degrees of freedom equal to the number of parameters added. Again, all analyses were carried out in R version 3.02 (R Development Team, 2009) using the lme4 packages (Bates *et al*., 2015). As in <Table 8>, the effect of the type of initial segment did improve model fit ($\chi^2(2) = 209.18$, $p < .0001$), as well as the effect of the initial segment type on the linear, quadratic and cubic terms ($\chi^2(2) = 27.09$, $p < .0001$; $\chi^2(2) = 6.38$, $p < .05$; $\chi^2(2) = 20.16$, $p < .0001$). Even though all the models showed significant improvement from their own previous model, model 3 is chosen as the optimal model because it has the lowest value in log-likelihood ratio test.

Table 8. Model comparison

	df	-2LL	χ^2	χ^2 df	*P*
Base	15	-4336			
Model 0	17	-4231	209.18	2	<.0001
Model 1	19	-4218	27.09	2	<.0001
Model 2	21	-4214	6.38	2	<.05
Model 3	23	-4204	20.16	2	<.0001

To compare the effect of the models, <Figures 4 and 5> are presented below. <Figure 4> shows the model fit to the data using model 2, and <Figure 5> shows the model fit to the data using model 3. Albeit its statistical improvement from the previous model, model 2 does not appear to capture the underlying tonal pattern as good as model 3 does.

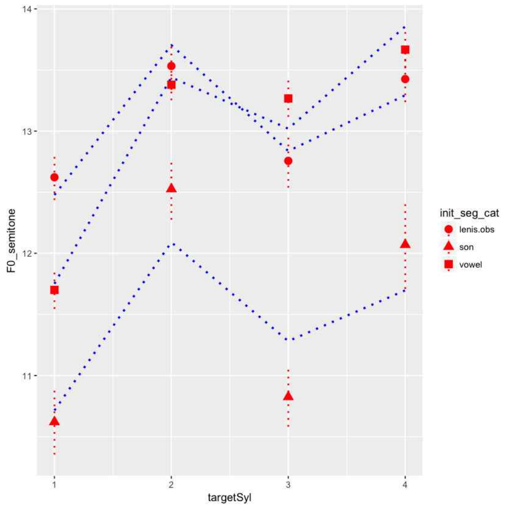

Figure 4. Observed data and growth curve model fits (model 2)

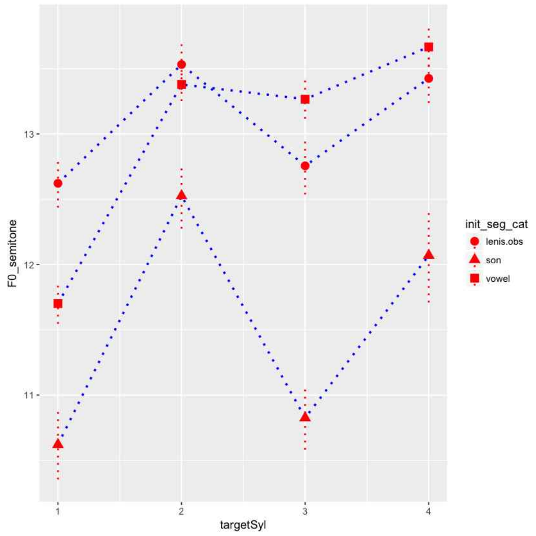

Figure 5. Observed data and growth curve model fits (model 3)

<Table 9> shows the fixed effect parameter estimates and their standard errors along with *p*-values estimated using the normal approximation for the *t*-values. Since we mentioned that the log-likelihood ratio test indicated that model 3 in <Table 8> revealed the best fit to the data, we will discuss the result of <Table 9> focusing on the functional form of the cubic polynomial term, rather than linear or quadratic forms. Polynomial curve shapes differ between the data whose initial segment is lenis obstruents, and the data whose initial segment is sonorant, but the shapes are not statistically different from data with the lenis obstruent initial segments and data with the vowel initial segments. Note that the polynomial analysis does not indicate the difference between lenis obstruent and vowel in the initial position, the difference between the two in the initial position is verified in section 3.1. No statistical significance between the two is related to the polynomial curves. <Figure 5> appears to indicate that the similarity between the two in syllable positions 2 and 4 makes the two not significantly different from each other.

Table 9. Parameter Estimates for Analysis of Effect of Initial Segment Type on tonal changes in the AP.

	Estimate	Std. Error	*t*	*p*
Intercept	13.06	0.43	30.12	<.001***
Linear	0.41	0.18	2.27	<.05***
Quadratic	-0.12	0.14	-0.82	.41
Cubic	0.71	0.14	4.83	<.001***
Son: Intercept	-1.54	0.12	-12.71	<.001***
vowel:Intercept	-0.06	0.09	-0.68	.49
son: linear	0.24	0.25	0.97	.33
vowel:linear	0.92	0.19	4.68	<.001***
son: quadratic	-0.11	0.24	-0.47	.63
vowel: quadratic	-0.47	0.19	-2.50	<.05
son: cubic	0.76	0.23	3.32	<.001
vowel: cubic	-0.17	0.18	-0.94	.34

Going back to the research question of whether the medial L tone will be governed by contrast maximization or ease of articulation, closer look at the figure indicates that results are mixed. In the case of lenis obstruent and sonorant, the medial tone shows the same level as the initial segment, indicating that maximization of contrast might be at work. However, in the case of vowels, the medial tone does not come down as low as the initial vowel but seems to show interpolating behavior between two surrounding high tones.

4. Conclusion

In this paper, the acoustic properties of APs, defined as word-like prosodic units neither preceded by nor followed by silent pauses, were extracted from a large scaled naturally occurring speech samples. Two research questions were posed: (i) whether the initial segment would have more than binary effects on the F0 of the syllable initial position, and (ii) whether the medial L tone in LHLH is governed by the contrastive maximization principle. In order to probe answers to these questions, the F0 values measured in semitone are fed into Growth Curve Analyses with the functional form of the cubic polynomial terms. The result showed that in general the initial segment had an effect on the F0 on the nucleus of the initial syllable. It also showed that the medial F0 in the LHLH form seems to be governed by the contrast maximization when the initial segment is either a lenis obstruent or a sonorant, and by the ease of articulation when it is a vowel.

The design of the paper focused on the initial segments, and the F0 (in semitone) is taken from the nucleus of syllable. The F0 of the syllable is affected by multiple factors, including the type of vowels, and the type of syllable structure. It would be desirable to explore the nature of the APs by designing a more detailed and controlled experiment using F0 as well as duration, and also by taking into consideration the type of syllable structure. These more detailed and extended work will be left for future research.

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.

Cumming, R. (2010). *Speech Rhythm: The Language-Specific Integration of Pitch and Duration*. Ph.D. Dissertation, University of Cambridge.

de Jong, K. (1994). Initial tones and prominence in Seoul Korean. *OSU Working Papers in Linguistics*, 43, 1-14.

Gandour, J. (1974). Consonant types and tone in Siamese. *Journal of Phonetics*, 2, 337-350.

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 form http://doi.org/10.5334/labphon.48 on March 20, 2017.

Jun, S.-A. (1996). *The Phonetics and Phonology of Korean Prosody: Intonational Phonology and Prosodic Structure.* New York: Garland Publishing.

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/K-tobi.html on April 2, 2015.

Jun, S.-A., & Fougeron, C. (2000). A phonological model of French intonation. In A. Botinis (Ed.), *Intonation: Analysis, Modeling and Technology* (pp. 209-242). Dordrecht: Kluwer Academic Publishers.

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.

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

Kim, K. (2013). *Tone, pitch accent and intonation of Korean – A synchronic and diachronic view*. Ph.D. Dissertation, Universität zu Köln.

Kohler, K. (2008). The perception of prominence patterns. *Phonetica*, 65, 257-269.

Lindblom, B. (1990). Explaining phonetic variation: A sketch of the H&H theory. In W. Hardcastle, & A. Marchal (Eds.), *Speech production and speech modeling* (pp. 403-439). Amsterdam: Kluwer.

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.

R Development Core Team (2009). *R: A Language and Environment for Statistical Computing*. R for Statistical Computing, Vienna, Austria.

Xu, Y. (2011). Speech Prosody: A Methodological Review. *Journal of Speech Sciences*, 1(1), 85-115.

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.

• **Tae-Jin Yoon**
Department of English Language and Literature
Sungshin University
34Da-Gil 2, Bomun-ro, Seongbuk-Gu, 02877, Korea
Tel: 02-920-7185
Email: tyoon@sungshin.ac.kr
Fields of Interest: Phonetics, Phonology, Speech and Language Technology