Linguistic Research 33(2), 299-317
DOI: 10.17250/khisli.33.2.201606.005

The effect of the consonant−induced pitch on Seoul Korean intonation*

Sunghye Cho · Yong-cheol Lee
(University of Pennsylvania · Cheongju University)

Cho, Sunghye and Yong-cheol Lee. 2016. The effect of the consonant-induced pitch on Seoul Korean intonation. *Linguistic Research* 33(2), 299-317. This study conducted a production experiment with phone-number strings and natural words as stimuli to evaluate the effect of the consonant-induced pitch on Seoul Korean intonation. The results of the present study show that the pitch range of an Accentual Phrase (AP) is higher when it starts with a high-pitch inducing segment (aspirated or tense) than when it starts with a low-pitch inducing segment (lenis, sonorant, or vowel). That is, later syllables, such as the penultimate and final syllables of APs, are also produced with a high pitch if the AP-initial syllable has a high-pitch inducing onset. We also find that the high-pitch inducing consonants in non-AP-initial syllables raise the pitch values of the syllables but their effect is much smaller than those in the AP-initial syllable. This finding that the high pitch in an AP-initial syllable extends to later syllables of the same AP leads us to hypothesize that the consonant-induced pitch is not merely a local segmental effect; rather, it serves as a potential indication of a tonogenetic sound change. **(University of Pennsylvania · Cheongju University)**

Keywords accentual phrase, intonation, tonogenesis, Seoul Korean, consonant-induced pitch, sound change

1. Introduction

In the prosodic hierarchy of Seoul Korean, there are three levels of prosodic units: Intonational Phrase (IP), Intermediate Phrase (ip), and Accentual Phrase (AP) (Jun 1993, 1996, 1998, 2006). An IP is the highest level in the hierarchy, and it marks the final boundary tone of an utterance.[1] An ip is the domain of prosodic

* We would like to thank Mark Liberman, Eugene Buckley, and Jianjing Kuang for their feedback and comments on this project. Our thanks also go to two anonymous reviewers for their valuable comments on this paper. This study was conducted as part of the first author's forthcoming dissertation at the University of Pennsylvania.

focus and pitch resetting within an utterance. Finally, the smallest prosodic unit is an AP, which usually consists of one or two content words optionally followed by a grammatical marker. The basic melody of an AP in Seoul Korean is TH-LH, where T stands for either high (H) or low (L) pitch. Previous studies have shown that AP-initial consonants induce a pitch contrast on the following vowel depending on their laryngeal features (Jun 1993, 1996; Silva 2006; Oh 2011; Kang 2014; among others). An AP shows a high pitch on the AP-initial syllable when it begins with a [+spread glottis] or [+constricted glottis] segment (i.e., aspirated or tense consonants), thereby resulting in a HH-LH melody.[2] However, an AP beginning with other segments shows a low pitch on its first syllable (a LH-LH melody). According to Jun's theory, the first two tones (TH in **TH**-LH) are aligned with AP-first and second syllables, and the final two tones (LH in TH-**LH**) are aligned with AP-penultimate and final syllables.[3] For APs longer than four syllables, the pitch realizations of syllables between the AP-second and the penultimate syllables are interpolated from the first H to the penultimate L (T**H-L**H).

Korean is known to have a unique three-way contrast among stops (and among affricates): aspirated, tense, and lenis categories. It has been shown that the three stop categories have distinct VOT values (Lisker and Abramson 1964, Han and Weitzman 1970); the tense category has the shortest VOT, whereas the aspirated category has the longest VOT. However, recent studies reveal that the VOT difference between aspirated and lenis is neutralized in the case of younger speakers (Kim 1994; Choi 2002; Silva 2006; Kang and Guion 2008; Wright 2007; Oh 2011; Kang 2014). Most of these studies find that the VOT of the aspirated category is shortened in younger speakers' speech, confirming the reduction of the difference in VOT between the aspirated and the lenis categories. In addition, many previous studies show that the f0 distinction between the lenis and the aspirated categories is robust among younger speakers but not among older speakers (Choi 2002; Kim

---

1 *Tone* in Jun's theory does not refer to lexical tones, but refers to intonational tones as the one in the ToBI (Tone and Break Indices) theory.

2 The laryngeal feature used to distinguish the three-way Korean stops may vary. Some use [+stiff vocal fold] (adopted from Halle and Stevens 1971), and others use [+spread glottis] and [+constricted glottis] for aspirated and tense, respectively. The scope of this paper does not address differences between these features, but those who are interested in this topic may refer to Cho et al. 2002, Silva 2006, and Wright 2007.

3 When there are less than four syllables in an AP, the melody may vary. See Jun (1993) and (1996).

2004; Silva 2006; Kang and Guion 2008; Wright 2007; Kong et al. 2011; Oh 2011; Lee and Jongman 2012; Kang 2014). It has been found that the following vowels after the tense and the aspirated categories are produced with a high pitch, whereas those after the lenis category are produced with a low pitch. For example, Kang (2014) studies the trade-off between the VOT contrast and the f0 distinction in depth using a large-scale speech corpus. She shows that the VOT merger and the development of the f0 distinction in the AP-first syllable have taken place together, not in separate stages and that this sound change is a structural sound change that affects all aspirated and tense consonant categories, not only stops but also fricatives and affricates. Previous studies (Silva 2006; Wright 2007; Oh 2011; Kang 2014; among others) suggest that the trade-off between the VOT contrast and the f0 distinction might be an example of *tonogenesis* (first coined by Matisoff 1973), which is a type of sound change, in which tones emerge (in previously toneless languages) and multiply (in previously tone languages).

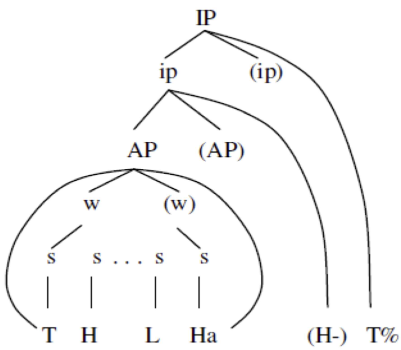

Figure 1. Intonation model of Seoul Korean
(Borrowed from Jun and Cha 2015)

Although many questions about this change in Seoul Korean have been addressed and answered in the literature, previous studies have generally investigated the phonetic characteristics of AP-initial syllables. A few studies (Silva 2006; Kang 2014) have examined AP-second syllables, and their results suggest that pitch values of the second syllables in lenis-initial APs are significantly lower than those in aspirated- or tense-initial APs. This difference indicates that the pitch values of

AP-second syllables are also affected by the laryngeal features of AP-initial onset consonants. What still remains to be explored is what effect the f0 distinction in the AP-initial syllable has on other later syllables within the same AP. Silva (2006) and Kang (2014) reveal that there is a large pitch difference even in the AP-second syllable, which gives a clue that the pitch values of AP-later syllables could also be determined by the AP-initial pitch distinction. However, AP-later syllables and the effect of consonant-induced pitch on those syllables are not investigated in the previous studies, since the main topic of the previous studies is not Korean intonational melody, but the trade-off between VOT and pitch contrasts. Therefore, in order to increase our understanding of the AP pitch patterns in Seoul Korean, systematic and thorough analyses should be conducted to test whether, and in what way, the pitch distinction occurring in the AP-initial syllable spreads to AP-later syllables. Of particular interest is how consonant-induced pitch has an effect on the overall intonational melody of Seoul Korean. While working towards this goal, we conducted a production experiment in which phone-number strings and natural words were employed as test stimuli.

2. Methods

In this experiment, 8 speakers of Seoul Korean participated (4 female and 4 male speakers in their 20s; mean age, 23.88 years). At the time of recording, they were all studying in the U.S. Most of them (7 out of 8) reported their length of residence in the U.S. as less than one year. One male speaker (27 years) reported he had lived in the U.S. for three years, but used Korean daily to talk with his spouse at home.

In the experiment, phone-number strings and natural words were employed as stimuli. For both types, the onset consonants of all syllables of the stimuli were controlled for four contexts (*hh, hl, lh,* and *ll*) to examine the effect of the consonant-induced pitch from onset consonants (*h* or *l*) on Seoul Korean intonation (TH-LH).[4] In the high-high (*hh*) context, all syllables of the target strings/words started with high-pitch inducing consonants (aspirated or tense). In the low-low (*ll*) context, all syllables started with low-pitch inducing segments (lenis, sonorant, or

[4] The segment-induced pitch *h* or *l* is marked with small (italicized) letters to differentiate them from the intonational tones (TH-LH) in Jun's theory, which are represented in capital letters.

vowels). In the other two contexts (*hl* and *lh*), low-pitch inducing types and high-pitch inducing types were alternated with each other so that *h-l-h-l*⋯ or *l-h-l-h*⋯ sequences were constructed.

To rule out any morphological factors, following the method used in Lee (2015), we employed 10 different phone-number strings for each pitch context (*hh, hl, lh,* and *ll*). The positions of the digits were controlled so that each digit (0–9) occurred equally often in each position. In all phone-number strings, the first three numbers were grouped together as an area code, and the last four digits were also grouped together. The number of digits in the middle varied from 2 to 5 to see how the segment-induced pitch and the intonational melody were realized in different AP sizes. The high-pitch inducing digits for the phone-number strings included 3 /sam/, 4 /sa/, 7 /čʰil/, and 8 /pʰal/, whose onset is an aspirated consonant, as well as 1 /il/, which is reported to be pronounced with a high pitch even though it is vowel-initial (Jun and Cha 2015). The other digits, 2 /i/, 5 /o/, 6 /juk/, 9 /ku/, and 0 /koŋ/, were classified as low-pitch inducing digits. The total number of phone-number strings was 1,280 (4 contexts x 4 AP sizes x 10 digit strings x 8 speakers), and the digit strings were presented within a carrier sentence, /ne.pʌn.ho.nɨn 000 − 00⋯ − 0000 (i).ya/ ‘My phone number is 000 − 00⋯ − 0000.’ Schematic forms and examples of the phone-number strings used in the experiment are presented in Table 1.

Table 1. Stimuli examples for the phone-number strings

	2-syllable	5-syllable
hh	*hhh* - *hh* - *hhhh* ex) 384 - 71 - 8148 /sam.pʰal.sa/ - /čʰil.il/ - /pʰal.il.sa.pʰal/	*hhh* - *hhhhh* - *hhhh* ex) 331 - 43371 - 8841 /sam.sam.il/ - /sa.sam.sam.čʰil.il/ - /pʰal.pʰal.sa.il/
hl	*hlh* - *hl* - *hlhl* ex) 804 - 35 - 3585 /pʰal.koŋ.sa/ - /sam.o/ - /sam.o.pʰal.o/	*hlh* - *hlhlh* - *hlhl* ex) 803 - 70157 - 8539 /pʰal.koŋ.sam/ - /čʰil.koŋ.il.o.čʰil/ - /pʰal.o.sam.ku/
lh	*lhl* - *lh* - *lhlh* ex) 230 - 58 - 6154 /i.sam.koŋ/ - /o.pʰal/ - /juk.il.o.sa/	*lhl* - *lhlhl* - *lhlh* ex) 215 - 64049 - 5357 /i.il.o/ - /juk.sa.koŋ.sa.ku/ - /o.sam.o.čʰil/
ll	*lll* - *ll* - *llll* ex) 955 - 26 - 0905 /ku.o.o/ - /i.juk/ - /koŋ.ku.koŋ.o/	*lll* - *lllll* - *llll* ex) 959 - 95929 - 9229 /ku.o.ku/ - /ku.o.ku.i.ku/ - /ku.i.i.ku/

Sixteen natural words employed as the target words are shown in Table 2. They were also controlled in terms of their segment-induced pitch patterns (*hh, hl, lh,* and *ll*) and the number of syllables (2–5 syllables). The target words do not contain low vowels, since there is a cross-linguistic tendency for low vowels to show lower f0 values than high or mid vowels (Whalen and Levitt 1995). All target words of *hh* and *hl* started with aspirated stops to directly compare *hh* to *hl*. The target words were presented in a carrier sentence /i.če.nɨn _______ mal.ha.se.jo/ 'Now say _____.' Each target word was read 10 times by each subject in a randomized order, and the total number of tokens of natural words was also 1,280 (4 contexts x 4 AP sizes x 10 repetitions x 8 speakers).

Table 2. Natural words employed in the experiment

	2-syllable	3-syllable	4-syllable	5-syllable
hh	/kʰʌ.pʰi/ 'coffee'	/kʰʌm.pʰu.tʰʌ/ 'computer'	/kʰʌm.pʰu.tʰʌ.čʰek/ 'computer book'	/pʰi.s'i.kʰʌm.pʰju.tʰʌ/ 'PC computer'
hl	/pʰi.pu/ 'skin'	/kʰe.i.kʰɨ/ 'cake'	/pʰɨ.lʌ.pʰo.čɨ/ 'propose'	/kʰɨ.lim.kʰe.i.kʰɨ/ 'cream cake'
lh	/pu.pʰi/ 'mass'	/in.tʰʌ.net/ 'internet'	/pi.heŋ.ki.pʰjo/ 'flight ticket'	/in.tʰʌ.net.čʰin.ku/ 'online friend'
ll	/u.pi/ 'raincoat'	/ki.lʌ.ki/ 'seagull'	/u.li.ko.mo/ 'my aunt'	/u.li.ʌ.mʌ.ni/ 'my mother'

Recordings were made in a sound-attenuated recording booth at the University of Pennsylvania, and were saved as *.wav* files at the sampling rate of 22.1 kHz, in mono. A headset condenser microphone (Shure WH30) was used, and the recordings were directly digitized into a desktop computer via *Praat*. The subjects signed a consent form before they were recorded, and received 10 dollars for participating in the experiment. The recording took less than 30 minutes per speaker, and the total number of tokens collected was 2,560 (= 1,280 phone-number strings + 1,280 natural words). The onset and offset of syllables were forced-aligned by a Korean forced aligner (Yoon and Kang 2012), and the alignments were manually checked and corrected afterwards. The mean pitch value of each syllable was obtained with a Praat script. The obtained pitch values were double-checked for f0 tracking errors. The obtained f0 values were converted to semitones (St) with 100Hz as a reference f0, using the formula $\log_2(Hz/100) \times 12$.

3. Results

3.1 Phone-number strings

Figure 2 shows the mean pitch values averaged across all 8 speakers. The most noticeable feature is that the pitch range of an entire AP is determined by the AP-initial digit. *hh* and *hl* generally pattern together in terms of their pitch contours, whereas *ll* and *lh* pattern together. There is a large pitch difference between APs starting with a low-pitch inducing digit (*ll* and *lh*) and those starting with a high-pitch inducing digit (*hh* and *hl*). Not only is this difference noticeable in the AP-initial position, but it is also retained until the end of an AP. The high-pitch inducing digit in non-AP-initial syllables of *lh* and *hh* is not as decisive as the digits in the AP-initial position. For example, the pitch contour of the *hh* context in Figure (2a) is slightly higher than that of the *hl* context, and that of the *lh* context is also slightly higher than that of the *ll* context in particular, in the fourth AP of the carrier sentence. However, these differences between *hh* and *hl* (or between *ll* and *lh*) are much smaller than the difference between the *h*-initial and *l*-initial APs (*hh*, *hl* vs. *ll*, *lh*).

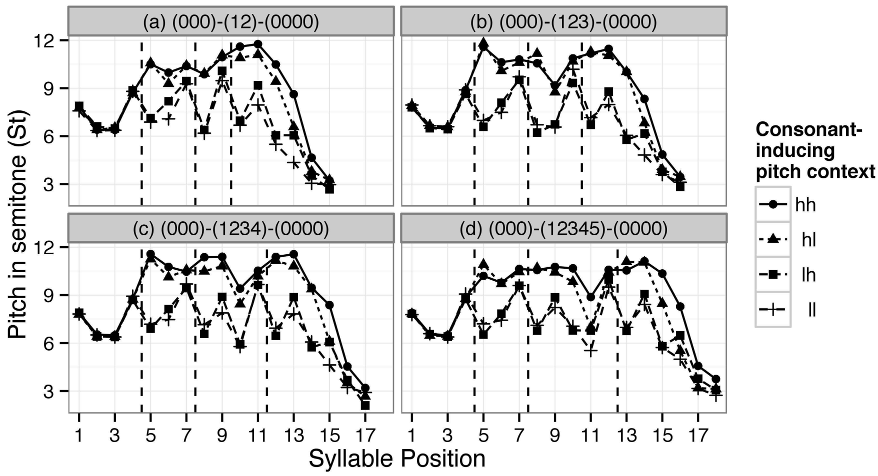

Figure 2. Mean pitch values of all speakers by the size of the second phone-number strings (where the number of digits varies from (a) 2 syllables to (d) 5 syllables). The x-axis shows syllable positions of the target sentences, and the dashed lines represent AP boundaries. The first AP is /ne.pʌn.ho.nɨn/ 'my number-TOPIC' in the carrier sentence and the other APs are the target phone-number strings. The phone numbers in the title of each plot (000-12…-0000) are for illustration.

Lastly, the pitch difference between *h*-initial APs (*hh* and *hl*) and *l*-initial APs (*lh* and *ll*) decreases in the phrase-final position to the extent that the mean pitch values overlap somewhat. This seems to be mainly due to the phrasal-final intonational H tone (TH-L**H**), and the effect of the phrase-final intonational H tone is observed in the second and third APs of the carrier sentence in all plots in Figure 2.

To determine whether the observations are statistically significant, five linear mixed-effects models (for each syllable position within an AP) were built with each pitch pattern within each speaker as a nested random effect. Instead of using all f0 values of the target sentences as the response variable in the analysis, only the third AP, where the number of syllables varied from two to five, was considered in the models. As for syllable positions, all AP-initial and AP-final syllables are coded as Initial and Final, respectively. All AP-penultimate syllables are defined as Penultimate for APs with more than three syllables. The rest of the syllables are defined relative to their positions to the beginning of an AP (2-syllable AP: Initial–

Final, 3-syllable AP: Initial–Penultimate–Final, 4-syllable AP: Initial–Second–Penultimate–Final, 5-syllable AP: Initial–Second–Third–Penultimate–Final). Table 3 summarizes the output of the models.

The models estimate that the pitch differences between *h*-initial and *l*-initial APs (e.g., ***hh*** vs. ***ll,*** ***hh*** vs. ***lh,*** ***hl*** vs. ***ll,*** and ***hl*** vs. ***lh***) are significant in all syllable positions ($p = 0.003$ for the comparison of *lh* vs. *hl* and $p = 0.001$ for the comparison of *ll* vs. *hl* in the AP-final position; $p < 0.001$ for all the other comparisons). This result indicates that the pitch of AP-initial onset consonants affects the pitch values of the rest syllables of the same AP, confirming the observation that the entire pitch range of an AP is affected by an AP-initial onset consonant as illustrated in Figure 2. The pitch differences between *h*-initial and *l*-initial APs in the AP-final syllable (from 0.88 St to 1.02 St) are smaller than those of the other syllable positions, but they are still found to be significant ($p < 0.001$ for both comparisons of *hh* vs. *lh* and *hh* vs. *ll*, $p = 0.003$ for *lh* vs. *hl,* and $p = 0.001$ for *ll* vs. *hl*).

Table 3. The outputs of the linear mixed-effects models of the phone-number strings: each table represents the results of group comparisons (of the pitch contexts) in a given syllable position. The first column of each table shows the estimated coefficient (i.e., the estimated pitch difference between two groups), and the second column is its standard error. The *z* values are the estimates divided by the standard errors, and the probabilities in the fourth column are given based on the calculated *z* values.

	Initial				Second			
	Est.	St. Error	*z* value	Pr(>|*z*|)	Est.	St.Error	*z* value	Pr(>|*z*|)
hl - hh	-0.03	0.42	-0.08	0.999	-0.46	0.31	-1.46	0.46
lh - hh	-4.1	0.42	-9.76	< 0.001 ***	-2.21	0.31	-7.06	< 0.001 ***
ll - hh	-3.8	0.42	-8.96	< 0.001 ***	-3.02	0.31	-9.67	< 0.001 ***
lh - hl	-4.07	0.42	-9.58	< 0.001 ***	-1.75	0.31	-5.6	< 0.001 ***
ll - hl	-3.77	0.42	-8.88	< 0.001 ***	-2.57	0.31	-8.2	< 0.001 ***
ll - lh	0.3	0.42	0.7	0.896	-0.82	0.31	-2.6	0.045 *

	Third				Penultimate			
	Est.	St. Error	*z* value	Pr(>|*z*|)	Est.	St.Error	*z* value	Pr(>|*z*|)
hl - hh	-0.75	0.55	-1.35	0.529	-1.06	0.39	-2.71	0.035 *
lh - hh	-3.78	0.55	-6.84	< 0.001 ***	-2.64	0.39	-6.73	< 0.001 ***
ll - hh	-3.52	0.55	-6.37	< 0.001 ***	-3.16	0.39	-8.05	< 0.001 ***
lh - hl	-3.03	0.55	-5.49	< 0.001 ***	-1.58	0.39	-4.03	< 0.001 ***
ll - hl	-2.78	0.55	-5.02	< 0.001 ***	-2.1	0.39	-5.35	< 0.001 ***
ll - lh	0.26	0.55	0.46	0.967	-0.52	0.39	-1.32	0.55

	Final			
	Est.	St. Error	*z* value	Pr(>|*z*|)
hl - hh	-0.15	0.23	-0.64	0.92
lh - hh	-0.97	0.23	-4.13	< 0.001 ***
ll - hh	-1.02	0.23	-4.37	< 0.001 ***
lh - hl	-0.82	0.23	-3.49	0.003 **
ll - hl	-0.88	0.23	-3.73	0.001 **
ll - lh	-0.06	0.23	-0.25	0.995

The models also estimate that the pitch differences between *hh* and *hl* are not significant in any syllable positions ($p < 0.999$ for the AP-initial position, $p = 0.46$ for the AP-second position, $p = 0.529$ for the AP-third position, $p = 0.92$ for the

AP-final position), but in the penultimate syllable (p = 0.035). This result suggests that *hh* and *hl* generally pattern together, and the onset consonant of the AP-second syllable plays a less important role than that of the AP-initial syllable. However, this does not mean that high-pitch inducing consonants in non-AP-initial positions have no effect on the Korean intonational melody. In the result of the penultimate syllable, the model estimates that the pitch value of *hh* is 1.06 St higher than that of *hl* and this difference is significant (p = 0.035), suggesting that the high-pitch inducing segment in the penultimate syllable has an effect on the intonational L tone in TH-**L**H, making the pitch value of the penultimate syllable of *hl* higher than that of *hh* (see Figure (2d)). However, the models reveals that the pitch value of *hl* in the penultimate syllable is still 2.1 St and 1.58 St higher than those of *ll* and *lh* (APs starting with a low-pitch inducing segment), respectively, and these pitch differences are significant ($p < 0.001$ for both comparisons). This result suggests that even the pitch value of the penultimate syllable (where the intonational L tone in TH-**L**H falls) is higher when the AP-initial consonant is a high-pitch inducing segment, i.e., aspirated or tense, than when it is a low-pitch inducing segment.

Similarly, the models estimate that the pitch difference between *ll* and *lh* are not significantly different in most syllable positions (p = 0.896 for the AP-initial syllable, p = 0.967 for the AP-third syllable, p = 0.55 for the AP-penultimate syllable, p = 0.995 for the AP-final syllable), except in the AP-second syllable (p = 0.045). This result also suggests that *ll* and *lh* generally pattern together in terms of their pitch values within an AP. Again, the significant difference between *ll* and *lh* in the AP-second syllable seems to be due to the consonant-induced pitch. The high-pitch inducing segment in the second syllable of *lh* seems to make the pitch value of its second syllable 0.82 St higher than that of *ll,* and this difference is significant in the model.

3.2 Natural words

Figure 3 shows the results of the natural words. The most important result is that the pitch range of an AP is mainly affected by the AP-initial onset consonant, as shown in the results of the phone-number strings. When an AP-initial onset is a high-pitch inducing segment (aspirated or tense), the pitch range of the AP is much higher than when it is a low-pitch inducing segment (lenis, sonorant, or vowel). There is a large pitch difference between the target words starting with aspirated or tense consonants (*hh* and *hl*) and those starting with other consonants or vowels (*lh* or *ll*). In addition, similar to the results of the phone-number strings, the effect of high-pitch inducing onsets in non-AP-initial syllables is much smaller than that of the AP-initial syllable. That is, *hh* and *hl* pattern together regardless of the size of the target words, and *ll* and *lh* are not much different from each other when the size of the target AP is 2 syllables, 3 syllables or 4 syllables. Yet, the target word of *lh* is noticeably higher than that of *ll* when the target AP is a 5-syllable word in Figure (3d). This seems to be because of the 5-syllable target word of *lh* /*in.t*ʰʌ.net.čʰin.ku/ 'online friend.' The speakers tended to produce /in.tʰʌ.net/ 'internet' and /čʰin.ku/ 'friend' as separate APs, raising /čʰin.ku/ 'friend' to the same pitch range of *hh* and *hl* due to the high-pitch inducing onset consonant /čʰ/ (aspirated affricate). An example of the pitch track of this target word produced with two separate APs is shown in Figure 4.

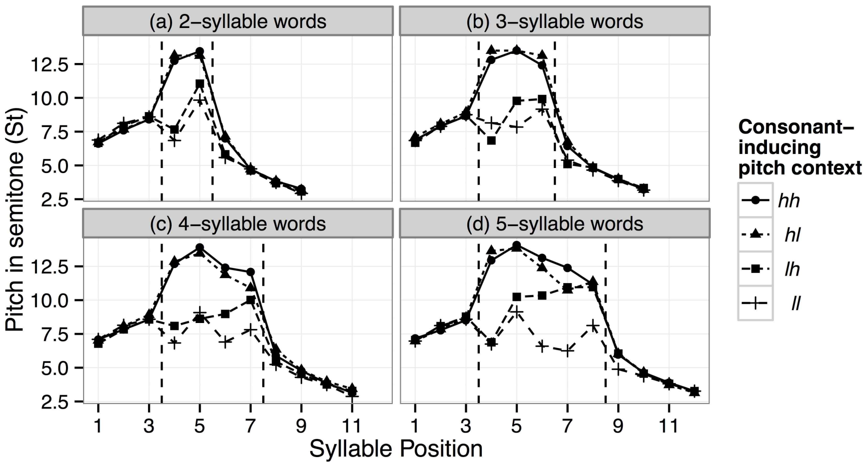

Figure 3. Mean pitch values of all speakers by the length of the target words. The x−axis shows syllable positions in the target sentences, and the y−axis shows the mean pitch values. The first three syllables are for /i.če.nɨn/ 'now−TOPIC', the last four syllables are for /mal.ha.se.jo/ 'say' in the carrier sentence /i.če.nɨn ____ mal.ha.se.jo/ 'Now say _____.', and the syllables in the middle (inside the dashed lines) are the target words.

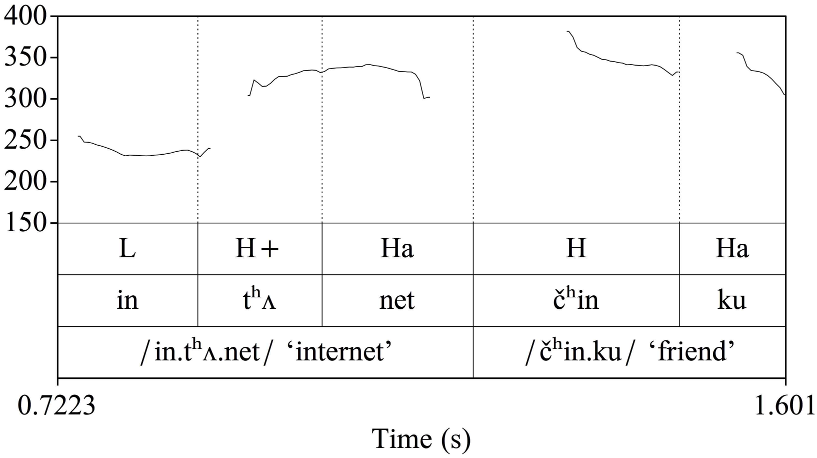

Figure 4. An example pitch track of the lh target word /in.t^{h}ʌ.net.č^{h}in.ku/ 'online friend' produced by a female speaker (JM). The pitch value is expressed in Herz (Hz). The first tier of the text grid shows the Korean intonational tones in the K−ToBI system (Jun 2006), and the second tier is for the syllable level. The third tier shows the word level with gloss in English.

For statistical analyses, five linear mixed-effects models (for each syllable position within an AP) were built with each pitch pattern within each speaker as a nested random effect. In these models, the f0 values of the target words were included as a dependent variable. Similar to the analyses of the phone-number strings, syllable positions within the target words were coded as the following: 2-syllable AP: Initial–Final, 3-syllable AP: Initial–Penultimate–Final, 4-syllable AP: Initial–Second–Penultimate–Final, 5-syllable AP: Initial–Second–Third–Penultimate–Final. Table 4 summarizes the outputs of the models.

The models estimate that the pitch values of the target words starting with a high-pitch inducing consonant (*hh* and *hl*) are significantly different from those of starting with a low-pitch inducing consonant (*lh* and *ll*) in all syllable positions ($p <$ 0.001 for all comparisons). In particular, the estimated pitch difference between *h*-initial and *l*-initial APs is the largest in the AP-initial syllable (from 5.35 St to 5. 89 St) and the second largest in the AP-second syllable (from 4.25 St to 4.6 St). The smallest pitch difference is found in the AP-final syllable position (from 1.65 St to 3.54 St), but the difference is still significant ($p <$ 0.001 for all four comparisons).

However, it is notable that the estimated pitch differences between *h*-initial and *l*-initial words in the AP-third and penultimate syllables vary considerably. The pitch differences between the *h*-initial target words (*hh* and *hl*) and that of *ll* are large (6.41 St and 5.79 St for Third, 5.32 St and 4.9 St for Penultimate), whereas the pitch differences between the *h*-initial target words (*hh* and *hl*) and that of *lh* are relatively small (2.66 St and 2.04 St for Third, 2.44 St and 2.02 St for Penultimate). This result seems to be due to the 5-syllable target word of *lh*, /in.tʰʌ.net.čʰin.ku/ 'online friend.' Because the participants tended to produce the target word with two separate APs, a 3-syllable AP (/in.tʰʌ.net/ 'online (LHH)') and a 2-syllable one (/čʰ in.ku/ 'friend (HH)'), the third syllable of this word in fact bears a final intonational H tone. Also, the penultimate and final syllables of this word are in a similar pitch range to those words starting with a high-pitch inducing consonant. This fact seems to be reflected in the models, estimating the pitch difference between the *h*-initial words and the target word of *lh* in the third and penultimate syllables smaller than the one between the *h*-initial words and that of *ll.*

The models also show that the pitch values of *hh* and *hl* are not significantly different from each other in all syllable positions (p = 0.543 for Initial, p = 0.999 for Second, p = 0.489 for Third, p = 0.613 for Penult, p = 0.912 for Final). This result confirms that *hh* and *hl* pattern together, regardless of AP sizes. Also, the pitch values of *ll* and *lh* are not significantly different in the AP-initial and second syllables (p = 0.714 for Initial and p = 0.853 for Second). However, the models indicate that the pitch differences between *ll* and *lh* in the other syllable positions are significant ($p < 0.0001$ for all three syllable positions). Again, this result seems to be due to the 5-syllable target word of *lh*, /in.tʰʌ.net.čʰin.ku/ 'online friend.' Because the participants tended to produce the target word with two separate APs, the last three syllables of that target word were produced with higher pitch values than those of *ll* and the models find the difference between *ll* and *lh* significant.

Table 4. The outputs of the linear mixed−effects models of the natural words: each table represents the results of group comparisons (of the pitch contexts) in the given syllable position. The first column of each table shows the estimated coefficient (i.e., the estimated pitch difference between two groups), and the second column is its standard error. The z values and the probabilities are given in the third and fourth columns.

	Initial				Second			
	Est.	St. Error	z value	Pr(>|z|)	Est.	St.Error	z value	Pr(>|z|)
hl - hh	0.3	0.23	1.33	0.543	-0.06	0.36	-0.16	0.999
lh - hh	-5.35	0.22	-23.82	< 0.001 ***	-4.31	0.36	-12.09	< 0.001 ***
ll - hh	-5.59	0.22	-24.89	< 0.001 ***	-4.6	0.36	-12.89	< 0.001 ***
lh - hl	-5.65	0.23	-25.1	< 0.001 ***	-4.25	0.36	-11.95	< 0.001 ***
ll - hl	-5.89	0.23	-26.18	< 0.001 ***	-4.54	0.36	-12.76	< 0.001 ***
ll - lh	-0.24	0.22	-1.06	0.714	-0.29	0.36	-0.8	0.853
	Third				Penultimate			
	Est.	St. Error	z value	Pr(>|z|)	Est.	St.Error	z value	Pr(>|z|)
hl - hh	-0.62	0.44	-1.42	0.489	-0.42	0.34	-1.22	0.613
lh - hh	-2.66	0.44	-6.09	< 0.001 ***	-2.44	0.34	-7.13	< 0.001 ***
ll - hh	-6.41	0.44	-14.68	< 0.001 ***	-5.32	0.34	-15.54	< 0.001 ***
lh - hl	-2.04	0.44	-4.67	< 0.001 ***	-2.02	0.34	-5.94	< 0.001 ***
ll - hl	-5.79	0.44	-13.27	< 0.001 ***	-4.9	0.34	-14.42	< 0.001 ***
ll - lh	-3.75	0.44	-8.6	< 0.001 ***	-2.88	0.34	-8.48	< 0.001 ***
	Final							
	Est.	St. Error	z value	Pr(>|z|)				
hl - hh	-0.16	0.25	-0.66	0.912				
lh - hh	-1.81	0.25	-7.33	< 0.001 ***				
ll - hh	-3.54	0.25	-14.33	< 0.001 ***				
lh - hl	-1.65	0.25	-6.67	< 0.001 ***				
ll - hl	-3.38	0.25	-13.66	< 0.001 ***				
ll - lh	-1.73	0.25	-6.99	< 0.001 ***				

4. Discussion and conclusion

The results of the phone-number strings and the natural words led to the same conclusion: when an AP starts with a high-pitch inducing type (aspirated or tense), all syllables in the AP are produced within a higher pitch range than when an AP starts with a low-pitch inducing segment (lenis, sonorant, or vowel). The results show that even the penultimate syllable, where the intonational L tone in TH-**L**H falls, is produced with a higher pitch in *h*-initial APs (*hh* and *hl*) than in *l*-initial ones (*lh* and *ll*). Also, we find that there is a structural pitch difference between *h*-initial and *l*-initial APs. In the results of the phone-number strings and the natural words, *hh* and *hl* are not significantly different from each other in most of the AP positions, but *lh* and *ll* are considerably different from *hh* and *hl* in all AP positions. This result suggests that the pitch distinction in the AP-initial position is no longer due to automatic phonetic effects from consonantal perturbation, as it affects not only the initial but also non-initial syllables of APs.

Also, the pitch difference between *ll* and *lh* in the 5-syllable natural words provides a clue that the AP-initial onset consonant is a primary factor affecting the entire pitch range of an AP. The *lh* target word tends to be produced with two separate APs (Figure 4), and the pitch range of the second AP is similar to that of the *h*-initial APs because of the high-pitch inducing onset segment /$\check{c}^h$/ of /$\check{c}^h$in.ku/ 'friend'. The high pitch values of the 5-syllable word of *lh* in the penultimate and the final syllables contribute to the significant pitch differences between *ll* and *lh* as demonstrated in Table 4. Again, this result suggests that the pitch distinction in the AP-initial position is the most important factor in deciding the pitch range of an AP.

In addition, the results of the present study reveal that high-pitch inducing consonants in non-AP-initial positions seem to increase the pitch values of non-initial syllables, although their effect is much smaller than those in AP-initial positions. For example, in the phone-number strings, the pitch value of *hh* in the penultimate syllable is found to be slightly higher than that of *hl* and this seems to be because the penultimate syllable of *hh* always has a high-pitch inducing onset. (This effect is most clearly seen in Figure (2d).) Also, the results of the phone-number strings show that the pitch value of *lh* in the AP-second syllable is higher than that of *ll* due to the segment-induced *h* pitch of *lh.*

One important question raised by these results is why we see such a large pitch difference between *h*-initial and *l*-initial APs (in all syllable positions). Considering the results of the present study and other previous studies on the Korean stops, we hypothesize that this is due to the tonogenetic sound change in the AP-initial syllable position (Silva 2006; Wright 2007; Oh 2011; Kang 2014; among others). Previous studies show that there is a trade-off between the VOT contrast and the f0 contrast in the AP-initial syllable in Seoul Korean. That is, while the VOT difference between the aspirated and the lenis categories are merging, the f0 difference between those categories is increasing among younger generations. The results of the current study seem to indicate that the pitch difference in the AP-initial position extends to later syllables of the same AP. Since the pitch of the AP-initial syllable with an aspirated or tense onset consonant is much higher than the one with a lenis or sonorant onset, the pitch of *h*-initial APs seems to stay relatively high through the end of the APs.

To conclude, the current study investigated the effect of the consonant-induced pitch on the Korean intonational melody, and we found that the pitch range of APs starting with a high-pitch inducing segment (aspirated or tense) is much higher than those starting with a low-pitch inducing segment (lenis, sonorant, or vowel) in all syllable positions within the APs. We also showed that the high-pitch inducing consonant in non-initial syllables has a local segmental effect on f0 of the following vowel, but this effect is smaller than the one in the AP-initial position. However, since the current study only investigated younger generations (speakers in their 20s), it is not yet clear when the consonant-induced pitch in the AP-initial syllable started to affect later syllables of the same AP. To provide a complete picture of the development of the AP-initial pitch distinction in Seoul Korean, a diachronic study that investigates speakers with a wide age range will be needed for future research.

References

Cho, Taehong, Sun-Ah Jun, and Peter Ladefoged. 2002. Acoustic and aerodynamic correlates of Korean stops and fricatives. *Journal of Phonetics* 30: 193-228.

Choi, Hansook. 2002. Acoustic cues for the Korean stop contrast: Dialectal variation. *ZAS papers in linguistics* 28: 1-12.

Halle, Morris and Kenneth Stevens. 1971. A note on laryngeal features. *MIT Research Laboratory of Electronics Quaterly Report* 101: 198-213.

Han, Meiko S. and Raymond S. Weitzman. 1970. Acoustic features of Korean /p, t, k/, /p', t', k'/ and /ph, th, kh/. *Phonetica* 22: 112-128.

Jun, Sun-Ah. 1993. The phonetics and phonology of Korean prosody. PhD Dissertation, Ohio State University.

Jun, Sun-Ah. 1996. The influence of the microprosody on the macroprosody: A case of phrase initial strengthening. *UCLA Working Papers in Phonetics* 92: 97-116.

Jun, Sun-Ah. 1998. The accentual phrases in the Korean prosodic hierarchy. *Phonology* 15:

189-226.

Jun, Sun-Ah. 2006. Intonational phonology of Seoul Korean revisited. In Timothy J. Vance and Kimberly Jones (eds.), *Japanese-Korean Linguistics* 14: 15-26. Stanford: CSLI.

Jun, Sun-Ah and Jihyeon Cha. 2015. High-toned [il] in Korean: Phonetics, intonational phonology, and sound change. *Journal of Phonetics* 51: 93-108.

Kang, Kyoung-Ho and Susan G. Guion. 2008. Clear speech production of Korean stops: Changing phonetic targets and enhancement strategies. *Journal of Acoustical Society of America* 124(6): 3909-3917.

Kang, Yoonjung. 2014. Voice Onset Time merger and development of tonal contrast in Seoul Korean stops: A corpus study. *Journal of Phonetics* 45: 76-90.

Kim, Mi-Ran. 1994. Acoustic characteristics of Korean stops and perception of English stop consonants. PhD Dissertation, University of Wisconsin-Madison.

Kim, Midam. 2004. Correlation between VOT and F0 in the perception of Korean stops and affricates. *Interspeech 2004–8th Interanational Conference on Spoken Language Processing*, 49-52. Jeju, Korea.

Kong, Eun Jong, Mary E. Beckman, and Jan Edwards. 2011. Why are Korean tense stops acquired so early: The role of acoustic properties. *Journal of Phonetics* 39: 196-211.

Lee, Hyunjung and Allard Jongman. 2012. Effects of tone on the three-way laryngeal distinction in Korean: An acoustic and aerodynamic comparison of the Seoul and South Kyungsang dialects. *Journal of the International Phonetic Association* 42: 145-169.

Lee, Yong-cheol. 2015. Prosodic focus within and across languages. PhD Dissertation, University of Pennsylvania.

Lisker, Leigh and Arthur S. Abramson. 1964. A cross-language study of voicing in initial stops: Acoustic measurements. *Word* 20: 384-422.

Matisoff, James A. 1973. Tonogenesis in Southeast Asia. In Larry M. Hyman (eds.), *Consonant types and tone,* 71-95. CA: Linguistics Program, University of Southern California.

Oh, Eunjin. 2011. Effects of speaker gender on voice onset time in Korean stops. *Journal of Phonetics* 39: 132-142.

Silva, David J. 2006. Acoustic evidence for the emergence of tonal contrast in Contemporary Korean. *Phonology* 23: 287-308.

Whalen, Douglas H., and Andrea G. Levitt. 1995. The universality of intrinsic F0 of vowels. *Journal of Phonetics* 23: 349-366.

Wright, Jonathan D. 2007. The phonetic contrast of Korean obstruents. PhD Dissertation, University of Pennsylvania.

Yoon, Tae-Jin and Yoonjung Kang. 2012. A forced-aligned-based study of declarative sentence-ending 'ta' in Korean. *Proceedings of the 6th International Conference on Speech Prosody 2012.* Shanghai: Tongji University Press.

Sunghye Cho
Department of Linguistics
University of Pennsylvania
619 Williams Hall, 255 S. 36th Street, Philadelphia, PA 19104, USA
E-mail: csunghye@sas.upenn.edu

Yong-cheol Lee
Department of English Language and Literature
Cheongju University
298 Daesung-Ro, Sangdang-Gu, Cheongju 28503, South Korea
E-mail: yongcheol@cju.ac.kr

Received: 2016. 03. 02.
Revised: 2016. 04. 04.
Accepted: 2016. 04. 04.