1. Summary

This talk will present current issues and future prospects of research in Korean prosody, focusing on how intonational phonology interacts with segmental phonology and sound change. First, I will introduce the model of intonational phonology of Seoul Korean and how that affects the current debate on tonogenesis in Korean. Then, I will report ongoing exceptions to the rule of intonational phonology of Seoul Korean and provide future prospects of studies in Korean prosody.

2. Intonational Phonology of Seoul Korean

The intonational phonology of Seoul Korean was proposed in Jun (1993, 1998) within the Autosegmental-metrical (AM) framework (e.g., Pierrehumbert 1980, Beckman & Pierrehumbert 1986, Ladd 1996/2008), where an intonation contour is categorized as a linear sequence of two tonal targets, High (H) and Low (L), and tones have two functions, marking either a prosodic boundary or prominence. In this framework, the prosodic structure of an utterance is parsed with a hierarchical structure of prosodic units defined by intonation. In the model of Korean intonational phonology, the highest prosodic unit in Seoul Korean is the Intonational Phrase (IP), which is larger than the Accentual Phrase (AP), which is slightly larger than the Word (average 1.2 words; average length 3.5 syllables). This hierarchy is similar to the model of Japanese intonational phonology proposed by Pierrehumbert & Beckman (1988), but a unique feature of Korean intonational phonology lies in the fact that the tonal pattern of an AP is determined by its segmental property. The ‘tones’ in intonational phonology do not form a lexical category but rather a post-lexical category, thus not affecting the form or the meaning of a word. Therefore, intonational tones typically do not interact with the segmental properties of a word, though they can be sensitive to morpheme boundaries within a word. In Seoul Korean, however, the tonal pattern of an AP is determined by the laryngeal property of the AP-initial segment. When the segment begins with either an aspirated (ㅍ, ㅌ, ㅋ, ㅊ, ㅅ, ㅎ) or tense consonant (ㅃ, ㄸ, ㄲ, ㅉ, ㅆ), having a [tense] of [stiff] laryngeal feature (Kim 1965, Halle & Stevens 1971), the AP-initial syllable is realized with a High tone, and otherwise (i.e. when the segment begins with a sonorant sound or a lenis consonant) the AP-initial syllable is realized with a Low tone. Regardless of the AP-initial tone, however, the AP-final tone is typically a High tone. In other words, the basic tonal pattern of an AP, when there are more than three syllables, is either HHLH or LHLH.

This model was revised in Jun (2006, 2007, 2011, Jun & Cha 2015) by adding an Intermediate Phrase (ip) between the IP and the AP. A tree diagram illustrating the revised model of the intonationally defined prosodic structure of Seoul Korean is given in Figure 1. The highest unit, the IP, is marked by a boundary tone (e.g., L%, H%, LH%, HL%, LHL%, HLH%, LHLH%) and can have more than one ip, which is marked by an optional boundary tone, T- (= H- or L-), and can have more than one AP. The ip has two functions: either marking the end of a heavy syntactic constituent or marking prominence (focus) (Jun & Jiang 2019). For the former function, the right edge of the ip is marked by a boundary tone. To mark prominence, however, the left edge of the ip is marked by raising the ip-initial H tone, which is realized on the prominent word, thus creating pitch reset beginning on the prominent word.

When the basic tonal pattern of an AP is fully realized, the first two tones (TH) mark the left edge of the AP, and the last two tones (LHa) mark the right edge of the AP. But when an AP has fewer than four syllables, which is very common, one or both of the two AP-medial tones, H (labeled ‘+H’
in K-ToBI (Jun 2000) and L (labeled ‘L+’ in K-ToBI), can be undershot, showing a rising tone pattern (AP-initial L and AP-final H) (Kim 2004, Yoo & Jun 2016).

![Diagram](image)

Figure 1. Intonationally defined prosodic structure of Seoul Korean (Jun 2006, 2011, Jun & Cha 2015)

- IP = Intonational Phrase, ip = Intermediate Phrase, AP = Accentual Phrase
- W = phonological word, S = syllable, T = AP-initial H (when the AP-initial segment is aspirated or tense) or L (otherwise), (H) = optional H on AP-second syllable,
- (L) = optional L on AP-penultimate syllable, Ha = AP-final H boundary tone,
- (T-) = optional ip-final boundary tone of H- or L-, % = IP-final boundary tone (L%, H%, LH%, HL%, LHL%, HLH%, LHLH%, HLHL%, LHLHL%)

3. The debate on Tonogenesis

Tonogenesis is a linguistic process describing the emergence of a tonal contrast at the lexical level by phonologizing redundant pitch patterns conditioned by the loss of a segmental contrast. The three-way laryngeal contrast of Korean stops (lenis, aspirated, and fortis) has been traditionally known to be distinguished by the VOT (Voice Onset Time) duration as a primary cue along with other secondary cues such as f0 and phonation of the following vowel and stop closure duration. In general, VOT is shortest after tense stops, intermediate after lenis stops, and longest after aspirated stops, while f0 is higher after aspirated and fortis stops than after lenis stops (Lisker & Abramson 1964, Kim 1965, Han & Weitzman 1970, Cho et al. 2002). However, studies since the early 2000s (Silva 2002, 2006a, 2006b; Silva et al. 2004) have found that the VOT difference between lenis and aspirated stops has become very small among younger speakers of Seoul Korean (especially those born in late 1960s onward) when the stop is produced phrase (AP) initially (AP-medial lenis stops are typically voiced, so the VOT distinction between lenis and aspirated stops is maintained) (Jun 1993, 1998). See Figure 2. To my knowledge, Silva and his colleagues' work is the first study linking the possible emergence of a tonal system to a diachronic shift in VOT cues, i.e. tonogenesis. But he noticed that word-initial consonantal effects on f0 are not restricted to the initial syllable but rather extend into the higher prosodic domain (i.e. the AP tonal pattern in Jun 1993, 1998), and unlike the shift in VOT, there were no discernable diachronic effects on f0. In other words, in his apparent-time study data, the tonal pattern has been stable over time across age groups.
Subjects' year of birth

Figure 2. VOT difference between aspirated stops and lenis stops as a function of subjects' year of birth. The best-fit curve of a quadratic function (R² = 0.686) is shown; based on 36 speakers' data (21 females). (Fig.2 from Silva (2006b, p.293)). A vertical blue line marks the birth year 1970.

Among these studies, Bang et al. (2018) is the most comprehensive study investigating what they called the 'quasi-tonogenetic' sound change of Seoul Korean. They used the phrase quasi-tonogenesis because “the change does not to date exhibit all features of tonogenesis, where lexical tonal contrast develops from consonant-induced f0 distinction” (p.121). Based on a dataset from a large apparent-time corpus of Seoul Korean (1250 tokens from 11 words, across 118 speakers born between the 1930s and 1984, in utterance-initial position from 11 stories found in The Speech Corpus of Reading Style Standard Korean (The National Institute of the Korean Language 2005)), they claim that the sound change is due to transphonologization from a distinctive segmental feature to a tonal feature, and it is driven by production bias and adaptive reinterpretation of the speech signal. Based on the word-frequency effects on VOT and f0 data, they propose that “quasi-tonogenesis in Seoul Korean may be driven by contrast reduction – namely, production bias affecting VOT— and that production bias in VOT may be one source of tonogenetic sound change more generally”. That is, the weakening VOT distinction may have caused f0 enhancement. Figure 3 shows the f0 (left) and VOT (right) data of stop categories for female and male speakers as a function of the speaker’s birth year. They further claim, based on the by-speaker variability data, that “VOT contrast reduction has ended earlier than f0 contrast enhancement.” But it is important to note that unlike a VOT merger which threatens a phonemic distinction, f0 enhancement found in their study is phonetic and the distinction between L vs. H category (AP-initially) still remains the same across all age groups.

Figure 3. Empirical plot for f0 (left) and for VOT (right) of three laryngeal categories for female and male speakers as a function of speaker year of birth. (Fig.4 from Bang et al. 2018, p.131).
Contra to the studies mentioned above, Choi et al. (2020) propose a prosodic account for the phenomenon. Since previous studies have only examined the segment-tone interaction in phrase-initial positions, Choi and colleagues examined the stop distinctions in various prosodic positions (phrase-initial vs. medial) and prominence conditions (focused vs unfocused). They found that the way speakers use VOT versus f0 varies as a function of prosodic position and prominence. Specifically, VOT is still used as a primary cue in phrase-medial and unfocused conditions. Rather than f0 being transphonologized to the phonemic feature system, therefore, the sound change is best characterized as a prosodically-conditioned change in the use of a segmental feature (VOT) using the already available post-lexical tones in the intonational phonology of Korean. That is, since the lenis and the aspirated stops are always accompanied by categorically different L and H tones in the phrase-initial position due to the intonational phonology of Seoul Korean, the tonal differences may have become perceptually salient in association with these stops, and the principle of effort minimization may have caused the VOT feature be redundant in this particular prosodic position.

This prosodic account is further supported when we consider the timing of the VOT merger relative to the development of the intonational phonology of Korean. While we now know that the VOT neutralization is found among people born in the late 1960s and beyond, we first discovered this phenomenon in the early 2000s (when those speakers were younger than 40 years old). Before that time, multiple studies reported the VOT values as the main difference between the lenis and aspirated stops. However, the connection between the laryngeal feature and the AP-initial High vs. Low tone in Chonnam Korean was first observed in 1988 (Jun 1989) and similarly in Seoul Korean in 1992 (Jun 1993) for speakers born in 1960s and earlier (in their mid 20s~early 30s). This does not mean the L vs. H AP-initial intonation pattern in Korean “started” in late 1980s. I am very sure it existed at least in the speech of people born in the 1900s because my grand-relatives spoke with the same intonation, as did the generation of my parents, which is the source of the intonation model of the current generation. The point is that the tone-segment mapping in AP-initial position has been present for generations but the VOT merger started sometime during 1980s~1990s when those born in the late 1960s and early 1970s were adolescents and young adults. As shown in Jun (1996), the f0 difference after lenis or aspirated stops in Seoul Korean was much larger than what a typical low-level phonetic perturbation shows; the impact of the segment-tone mapping was not limited to the AP-initial syllable, but rather it affected the f0 level of subsequent syllables of the AP (as mentioned in Silva 2006). That is why it was suggested that the f0 difference after lenis vs. aspirated (and fortis) stops in Korean is phonologized at the post-lexical level as part of the representation of Korean intonational phonology. In other words, it is not likely that a VOT merger caused the phonologization of f0, but it is the f0 distinction at the AP-initial intonation pattern that caused the VOT merger. Since most words form one AP in the default prosody of Korean (Jun & Fougeron 2000, Schafer & Jun 2002, Kim 2004), most word-initial lenis vs. aspirated stops would be produced with a L vs. H toned vowel onset. So, native speakers would use the f0 difference at vowel onset as a primary cue in perceiving lenis vs. aspirated stops in word initial position if they hear the word in isolation, which is the common stimulus design in a perception test.

Therefore, we ought to consider the other direction of tonogenesis: how has this tone-segment mapping (High after aspirated or tense consonants vs. Low after other segment) been a part of the intonation system of all non-tonal dialects of Korean? This tone-segment mapping at AP-initial position is not limited to Seoul Korean; it is also found in Chonnam (Jun 1989, 1993, Oh 2020), Chungcheong (Oh & Jun 2019), and Jeju dialects (Lee 2014), while it is not found in tonal-dialects such as North and South Kyungsang dialects (Jun et al. 2006, Lee et al. 2013) and Yanbian Korean dialect (Jun & Jiang 2018). It has been proposed that the tonal features of Middle Korean are simplified to the lexical pitch accent system in tonal dialects and are lost in non-tonal dialects (Martin 1951, Ramsey 1978, Lee, S-O 2000, Jun & Kubozono in press). Instead of assuming a total loss of the tonal features in non-tonal dialects, however, it is possible to assume that they are further simplified to the AP-initial L vs. H in the intonational system. This means that the current L vs. H
distinction at the post-lexical level in Seoul and other non-tonal dialects could be indicative of a final stage in the process of tonoexodus, instead of the beginning stage of tonogenesis from a non-tonal language. Of course, these are hypotheses and only time will tell which direction Seoul Korean is following.

4. Ongoing exceptions to Seoul Korean intonational phonology

Since the tonal pattern of the AP in Seoul Korean is post-lexical, the mapping between the AP-initial segment and the tone type is not as categorical as the mapping in the lexicon. In natural speech, we can easily find exceptions to the mapping. For example, an AP may begin with a H when the segment is a sonorant consonant or begin with a L when the segment is an aspirated/tense consonant. Kim (2004) found about 5~7% exceptions to this tone-segment mapping (4.9% of AP-initial H were with L-tone inducing segments and 6.7% of AP-initial L were with H-tone inducing segments), based on the radio corpus of interviews and spontaneous speech and the read speech corpus. But the mismatch cases are not consistent or systematic with a certain lexical item or within or across speakers. Rather, they are in free variation, influenced by various factors such as adjacent words or segments or pragmatic motivations.

However, starting from the mid 1990s, the syllable [il] ‘digit ‘1’ or ‘the first’ (e.g., il.cho ‘1 second’, il.tung ‘the first place’) started to show a consistent mismatch with the tone, i.e., High tone, when the syllable is located at the beginning of an AP. Based on the data collected in 2011 (Jun & Cha 2015), the phenomenon has spread to other meanings of [il], such as ‘day’ (e.g., il.ki ‘diary’, il.pon ‘Japan’) and ‘work’ (e.g., il.son ‘worker’, il.saks ‘wages’), though the degree of High-[il] production was lower for the ‘work’ meaning than the ‘day’ meaning, which was in turn lower than the meaning of ‘#1/the first’.

What caused the onset of this exceptional rule? It was proposed that the most likely motivation of this mismatch is to avoid confusion between [il] ‘one’ and [i] ‘two’. The monosyllabic word [i] ‘two’ used to be a long vowel until Seoul Korean lost the length contrast. These two syllables/words were distinguished by the vowel length as well as the intonational tones. Since a Tone Bearing Unit (TBU) of the intonational contour is a mora (‘m’) if there is a vowel length contrast in Korean, a vowel length difference would create a different tone-mora alignment, as shown in (1a). For example, a word [il.tʃo] ‘1 second’, produced in isolation, would have L on [il] and HL on [tʃo], but [i.i.tʃo] ‘two seconds’ would have L on the first mora [i], H on the second mora [i] and L on the third mora [tʃo]. That is, the last mora [cho] would carry a falling tone (HL) when the first syllable is short, but a Low tone when the first syllable is long. However, after the loss of the vowel length, these two syllables/words must have been pronounced very similarly, as shown in (1b), only differing in the presence of the coda [i], probably causing confusion to listeners. This could have motivated a change from a L tone to a H tone on [il] since a high pitch is typically used to emphasize a word, and [il] meaning ‘the first’ is more likely to be emphasized than [i] ‘the second’ in Korean society. If the loss of vowel length is the cause of this H-[il] phenomenon, it is expected that this will not happen among speakers of Chonnam or Chungcheong dialects where the vowel length distinction is still preserved. As of 2010, young speakers of Chonnam dialect did not show any High-[il] phenomenon (Jun & Cha 2011).

(1) a. Seoul dialect (before V-length loss) b. Seoul dialect (after V-length loss)

```
   [m m]  [m m m]  [m m]  [m m]
   il tʃo i i tʃo il tʃo i tʃo
   L HL   L HL   L HL   L HL
  ‘one second’  ‘two second’  ‘one second’  ‘two second’
```

Though no study is available yet, it seems that a newer exceptional case of the intonational tone rule has been happening with the syllable [sa] ‘number 4’. [s], being an aspirated fricative, belongs to a High tone inducing segment (e.g., sa.ki ‘scam’), but these days it is often produced with a Low tone when it means ‘digit 4’ (e.g., [sa] in {sapon sonsu} ‘the athlete whose back number is #4’). This
exceptional mapping seems to be spreading to other words beginning with [sa] such as sa.ko ‘accident’. Again, it is possible that [sa] meaning ‘#4’ is produced with a L tone to enhance the perceptual distinction from [sam] ‘#3’. As in [i] vs. [j], [sam] has a short vowel but [sa] ‘#4’ used to have a long vowel, and now without the length contrast, [sam] and [sa] are much more similar to each other when uttered in the middle of a phrase. One could wonder if the exceptional mapping could spread further and the AP-initial tone-segment mapping is not predictable any more in Seoul Korean, causing Seoul Korean to be a tonal language.

Whatever the motivation was, what is interesting is that the age group who started the High-[il] exception (i.e., the speakers born in the early 1970s and later; See Figure 4) is similar to the age group who merged the VOT duration between aspirated and lenis stops in Seoul Korean (see Figure 2 and the right panel in Figure 3). And even more interesting is that, according to Kang et al. (2015), the loss of vowel length in Seoul Korean is completed by those born in the 1960s. Figure 5 shows the mean vowel duration of short and long vowels (ms) across speakers’ birth years (Figure 3 from Kang et al. 2015). So, it seems that these three types of sound change are related, all started by those born in the 1960s. Based on the data shown in Figures 2-5, the loss of vowel length, which was completed around the time when those born in the 1960s were linguistically active, might have triggered High-on-[il] change because, as mentioned earlier, the lack of length contrast would have reduced the perceptual distinction between [il] ‘one’ and [i] ‘two’ by losing the difference in the AP tonal pattern when [il] and [i] are the first syllable.

Figure 4. Percentage of High-toned [il] usage for each speaker, arranged by their birth year; based on 80 speakers (40 females). (Fig.8 from Jun & Cha 2015, p.104).

Figure 5. Mean vowel duration (ms) of long and short vowels aggregated over speakers’ decade of birth (from Figure 3 in Kang, Yoon, & Han 2015; measurement of 3,368 vowel tokens from The Reading-Style Speech Corpus of Standard Korean).
The loss of vowel length might have also contributed to the VOT merger. It has been shown that the VOT merger is found in Seoul and Jeju dialects (Holliday and Kong 2011), where a vowel length contrast is lost, but not in other dialects where vowel length is preserved (e.g., Chonnam (Choi 2002), Daegu (Holliday & Kong 2011, Kenstowicz & Park 2006), and Suncheon (Km M. 2017)). As shown in Jun (1998), the domain of Vowel Shortening is the AP. That is, though vowel length is lexical in Korean, a long vowel only occurs word-initially. However, when a word occurs AP-medially, a word-initial long vowel is shortened. So, on the surface, only AP-initial syllables can have a length contrast. So, by losing the length contrast at the AP-initial position, Seoul and Jeju speakers might have lost a sensitivity to VOT length. A similar view is given in Choi (2002).

So, if the VOT merger and the H-on-[il] phenomenon are both triggered by the loss of vowel length, one might wonder why a vowel length contrast is lost in Seoul and Jeju but not in other non-tonal dialects such as Chonnam and Chungcheong. One possibility could be due to the tonal pattern of an AP in an Intonational Phrase (IP) medial position. In Seoul and Jeju (Lee 2014), the most common tonal pattern of an AP is a rising tone, [L H], with a L on the first mora and a H on the final mora (when we only consider APs beginning with a L-tone inducing segment, for the sake of simplicity), but in Chonnam (Jun 1989, 1998; Oh 2020) and Chungcheong (Oh & Jun 2019) dialects, the most common tonal pattern of an AP is a rise-fall, [LHL], with a L on the first mora, a H on the second mora, and a L on the final mora. So, as shown in (1), the length of the AP-initial syllable would affect the tonal pattern of an AP in Chonnam and Chungcheong dialects, but not in Seoul and Jeju dialects because the tonal pattern of an AP is always a [L H] rising tone regardless of the vowel length difference in the first syllable. That is, a vowel length difference must have not been salient in the Seoul and Jeju dialects compared to that in the Chonnam and Chungcheong dialects.

Regardless of the cause of the loss of the vowel length contrast, it is not clear what caused multiple sound changes to happen among speakers born in the 1960s. We can only guess that when those speakers were old enough to participate in sound change, typically during their adolescence, which would be around mid-late 1980s, there might have been some social events or movements happening in Korean society that caused the sound change.

5. Future prospects and Conclusion

I have proposed that the Intonational phonology of Seoul and other dialects of Korean might be involved in multiple sound changes happening over the past few decades. This suggests how important it is to consider the prosody and intonation patterns of a language when investigating segmental phonology and sound change. It has been shown in the literature that the prosodic phrasing and prominence affect how segments are realized and perceived (e.g., Cho 2016, Holger et al. 2016, Kim & Cho 2013), how words are segmented (e.g., Kim & Cho 2009), and how sentences are processed (e.g., Kjelgaard & Speer 1999, Snedeker & Trueswell 2003, Jun & Bishop 2015). Prosody interfaces with multiple subareas of grammar. As such, studying the prosodic structure and intonational phonology of a language will help us to understand and investigate various questions in other subareas in linguistics. Now with the addition of the Intermediate Phrase (ip) in the intonational phonology of Seoul Korean, we need to examine how this prosodic unit is realized, how it affects segmental phonology, how the two functions of the ip are different or similar, and how the syntax-marking function of the ip could improve the syntax-prosody interface. Future studies should also examine the prosody of spontaneous data and build a speech corpus labeled with K-ToBI (version 4) or other intonational annotations such as an IPrA (International Prosodic Alphabet; Hualde & Prieto 2016) or PoLaR (Points, Levels, And Ranges Annotation; Ahn et al. 2019) which are more faithful to the surface prosodic events.

Finally, by studying Korean prosody, we can also examine the intonation patterns and prosody in children’s speech and the acquisition of second language prosody by comparing the intonation system of Korean to that of a second language. Studying Korean prosody can also contribute to the development of teaching materials for Korean as a foreign language and improving the quality of Korean speech synthesis and recognition through the collaboration with engineers and speech scientists.


