

Pedagogical and Acquisitional Implications of the Intonational Map Provided by Korean Textbook Example Conversations

Bonnie J. Fox

Abstract

Through the analyzation of a corpus of K-ToBI annotated speech taken from beginning level textbook conversation recordings, this paper aims to determine the global attributes of Slow, Clear Speech (SCS) on Korean prosody production, and to implicate these effects in the pedagogy of beginning-level Korean. In an analysis of the features that make Korean SCS distinct, four common themes emerged. First, there is final lengthening on Accentual Phrases (APs). Second, there are additional pauses and breaks between APs. Third, there is broad use of pitch reset and of focus in small syntactic frames. And fourth, boundary tones are typically flat and disaffected. Intonation plays a key role in the pursuit of L2 Korean intelligibility and is integral to strong acquisition of Korean. However, instructors rarely speak at normal speech rates (SR) with normal articulation, and typically use SCS with their beginning students. Students will recall frequently heard or salient intonational patterns, so instructors must take care to use intonational patterns intentionally. Thus, it is proposed that instructors of beginner students give explicit instruction and direct feedback on intonation and show natural speech examples often from various speakers, among other strategies to mitigate the effects of SCS on student intonational acquisition.

Key words: intonation, prosody, Korean, speech rate, clear speech

1 Introduction

This paper aims to accomplish two goals by creating and analyzing a corpus of K-ToBI annotated speech samples taken from the first textbook in the KlearText Korean language textbook series (Edition 2) (Cho, Lee, Schulz, Sohn, & Sohn, 2010). First, I aim to determine the global attributes of Slow, Clear Speech (SCS) on Korean prosodic production and second, I aim to implicate these effects in the pedagogical process of teaching Korean to beginning students. There has been ongoing work on Korean spoken at a normal articulation and speech rate, especially stemming from the development of the K-ToBI system (Jun, 2000; Jun, 2005; Jun, 2011), however, speakers of a language do not always speak at such optimal speech rates. Given such, work has since been done to look into the effects of increased speech rate and poor enunciation on Korean speech (Jun, 2003), however, the specific attributes of opposite, slow and clear speech are still unclear. Additionally, in the classroom and in for-textbook recorded speech, instructors rarely speak at normal speech rates (SR) with normal articulation (Kuder, 2017) and specifically in Korean, thanks to ambiguous syntactic constructions such as in-situ wh-questions, intonation plays a key role in the pursuit of L2 Korean intelligibility. There is also evidence that phonology begins to be internalized as soon as a person is first exposed to a language (Chang, 2012; 2013), meaning these first exposures to intonation can be critical to guiding students to intelligible intonational patterns.

Thus, the research questions of this paper are twofold: First, what are the attributes of the prosody of Korean SCS? Second, what are the implications of these attributes on L2 beginner speakers? It is predicted that the attributes of Korean SCS will center around 1) increased breaks (i.e. more APs, ips, and IPs) where they would not usually appear, and 2) on less affectual boundary tones (i.e. limited or non-existent use of longer tones such as HLH%, LHL%, LHLH%, and HLHL%). I hypothesize that we should see more mismatches between the intended prosodic category and the actual break produced, as well as a mismatch between the expected boundary tone complexity and the actual one used. The speech rate reduction is expected to account for the break mismatches and increased frequency, while the hyper-enunciation of clear speech is expected to account for the reduced tone complexity. Extrapolating from work in Exemplar Theory (Pierrehumbert, 2001, 2002, 2006; Wedel, 2006; Johnson, 2007), it can be predicted that students will recall intonational patterns they frequently hear or find salient among other weighting factors, indicating that instructors must take care to use intonational patterns intentionally in the classroom. Thus I propose that instructors of beginner students give explicit instruction and direct feedback on intonation (i.e. when students can pause, give alternate examples of how intonation changes the meaning of a sentence, etc.), show natural speech examples often from various speakers, among a few other strategies to mitigate the effects of SCS on Korean instructor speech.

2 Literature Review

2.1 The Effects of Increased Speech Rate (SR) & the Effects of Clear Speech (CS) on Prosody

In prior literature, the effects of increased Speech Rate (SR) and the effects of Clear Speech (CS) on prosody have been researched and documented. For the purposes of this paper, Speech Rate (heretofore SR) refers to the pace at which a speaker is speaking. According to Chen et. al. (2014), it is “a prosodic feature that influences many speech phenomena such as syllable duration, pause duration, prosodic phrasing, occurrence frequency of pause, word pronunciation, phone contraction, pitch contour shape, and so on” (p. 1158). Clear Speech (heretofore CS), according to Krause & Braida (2002), is “a speaking style that many talkers adopt in order to be understood more easily in difficult communication situations” (p. 2165) and it involves more careful articulation and often increased speech amplitude.

2.1.1 Effects of Increased Speech Rate on Prosody

With regard to the effects of increased SR on prosody, several studies have been conducted looking into both the acoustic effects of fast speech and the prosodic grouping changes in fast speech. In French, increased speech rate resulted in reduced pitch range and pitch displacements, as well as a reduction in the number of phrases (Fougeron & Jun, 1996), indicating that prosodic phonetic realization and organization are effected by speech rate. In Korean speech, an AP was found to include 5 or fewer syllables at normal rate, but can include up to 7 syllables at fast rate (Jun, 2003), thus indicating that the amount of viable syllables per phrase is determined by how fast the speech is. Controlling for SR in a Mandarin TTS system was found to have more accurate output than other non-controlled models (Chen et. al. 2014). This particular study also controlled the occurrence frequencies of different break types and their pause durations according to the given speaking rate, indicating that breaks and pauses are clear markers of speech rate integral to properly

accessing prosodic output. Peak amplitude event clustering was affected by speech rate in English where fast speech shifted nested clustering into shorter timescales and slower speech in longer timescales decreased the nested clustering (Ramirez-Aristizabal, Médé, & Kello, 2018). This suggests that speech will contract to fit faster pacing and will chunk out the slower the pace gets and also provides evidence for opposite effects from those found with fast speech to be seen in slow speech. Looking at these studies overall, it can be hypothesized that the slow SR may show opposite effects to increased SR, leading to greater numbers of groups, stronger junctures between phrases, and increased pitch range.

2.1.2 Effects of Clear Speech on Prosody

In terms of the effects of CS, there has been work done looking into the intelligibility of CS and its acoustic properties independent of SR. Chang, S.-e. (2017) looked into specifically the effects of CS on Korean glides and found that there was production enhancement at the onsets of glides with an expanded vowel space. While this is a segmental finding, the expansion of the glide vowel space can lead to a larger TBU or even to the addition of a TBU in places where the glides are typically heavily reduced in conversational speech which could affect the realization of different prosodic phrases.

Krause & Braidá (2002) indicate that CS is often studied in conjunction to slow SR since the slower SR is in part thanks to more frequent and longer pauses and lengthened phones, however, they aimed to isolate out SR from CS in their work (Krause & Braidá, 2002; 2004). Krause & Braidá (2002) looked into the intelligibility of CS in different rates (slow, normal, and fast) and amplitudes (soft and loud) as compared to conversational speech and found a consistent advantage in intelligibility for CS in all conditions over conversational speech. These results indicated that there are properties of CS which are independent of slow SR. Their follow-up paper, Krause & Braidá (2004), thus attempts to isolate these acoustic properties of CS as opposed to conversational speech. By comparing CS and conversational speech produced at the same SR, they were able to isolate out two global properties of CS with regards to prosody. They noted increased energy in 1000~3000 Hz range and increased modulation depth of low frequency modulations of intensity were indicators of CS. Their results indicate that SR and CS are independent factors which influence each other and do not inherently go together. Thus CS, though often produced with a reduced speech rate, can also be produced at normal speech rates with certain increased efforts (Krause & Braidá, 2002; 2004), thus when looking for prosodic characteristics in this corpus of Korean speech, it will be necessary to try and isolate out SR.

2.1.3 SR and CS-Adjacent Speech: Student Directed Speech & Child Directed Speech

Looking into two SR and CS-adjacent manners of speech, as Child Directed Speech (CDS) and Student Directed Speech (SDS), we can note that there are some other factors which may come into play within the results of the corpus analysis. First, looking at Child Directed Speech (CDS), Ko & Soderstrom (2013) found in an analysis of sentence long utterances, that the elongation effects of CDS stretch across the entire sentence with proportionally lengthened final syllables. The elongation effects were found to have the same temporal ratio to that of standard speech. These results indicate that other modes of native speaker speech modified to a more carefully produced and slower tempo would also likely maintain a similar ratio of elongation. Second, looking at Student Directed Speech (SDS), Kuder (2017) analyzed the speech of L2 Spanish TAs and found

a few global prosodic patterns. TAs consistently spoke louder, with a higher pitch register, and with a wider pitch range. They produced shorter prosodic phrases, a larger number of pauses, less prenuclear f0 rises, more f0 peaks, and showed less utterance-medial f0 suppression. These results indicate that speech which is directed at L2 learners, as is the speech produced in this corpus, could bear similar qualities. However, a key difference is that these corpus utterances are not live-produced to students, so certain strategies may be less prominent such as louder speech, which is more useful for grabbing student attention, something unnecessary in a recording context.

2.2 Korean Intonation and K-ToBI

2.2.1 The Prosodic Hierarchy of Korean

Korean intonation has been described in detail and subsequently researched for various prosodic, syntactic, and semantic properties since the creation of the K-ToBI (Korean Tones and Break Indices) annotation system (Jun, 2000) which is based in an Autosegmental-Metrical perspective on prosody. K-ToBI assumes a close relationship between intonational phonology and a hierarchical model of prosodic features (Jun, 2005). Korean prosody is composed of four main prosodic units: the *Intonation Phrase* (IP), the *intermediate phrase* (ip), the *Accentual Phrase* (AP), and the *Prosodic Word* (PW) (Jun, 2005; Jun 2007; Jun & Cha 2015).

There are several defining characteristics of each constituent of the prosodic hierarchy in Korean. Each of these characteristics are what drive the perception of naturalness in native speech and how these are realized is integral to understanding why changes to the realizations of these characteristics in slow, careful speech (SCS) is perceived as sounding unnatural. Starting with the most all-encompassing level, the Intonational Phrase (IP) is marked by a boundary tone (%) on the right or sometimes left edge of the phrase, final lengthening, and an optional pause afterwards (Jun, 2005). Boundary tones are realized on the final tone bearing unit (TBU), which is typically the final vowel of a phrase, but can span over the last few syllables of a phrase (Jun, 2014).

Moving to lower levels, the Accentual Phrase (AP) and the Intermediate Phrase (ip) are both units that are higher than the Prosodic Word (PW) and lower than the IP (Jun, 2007). A single word utterance will encompass all four levels of the hierarchy. The intermediate phrase (ip) is marked by syntactic or focus marking (-) and slight lengthening (Jun, 2011). The ip is usually a phrase consisting of a few words and indicates a pitch reset from down-stepping with a smaller amount of final lengthening than an IP. There are two kinds of ip according to Jun (2011): syntactic and prominence. The syntactic ip is marked on the right edge tone of the phrase (T-) and slight final lengthening and follows larger syntactic groupings such as noun-modifying constructions (Jun, 2011). The prominence ip denotes the beginning of an expanded pitch range for the upcoming focused or prominent word and there is no final lengthening on the preceding word's final syllable (Jun, 2011). The ip, which was added as a revision to the original K-ToBI system, to explain the correlation between the processing (the attachment preference) and production (prosodic phrasing) phenomena in Korean (Jun, 2007).

The AP is defined as smaller than an ip, but larger than a Phonological Word (PW). Korean is an agglutinative language and thus an AP typically consists of a lexical item and its case marker(s) and postpositions. Thus, the AP is slightly longer than a word. The AP is marked by a phrasal edge tone or melody, where the standard tune is THLH, where T is H (High) if the initial segment is aspirated or tense, and L (Low) if otherwise, and TBUs are undershot if there are not

four or more syllables. aPs typically do not show final lengthening. Finally, the Prosodic Word (PW) most typically aligns with a single lexical item, thus often an AP consists of only one PW.

2.2.2 The K-ToBI Labelling System

In terms of annotating actual speech, the K-ToBI system uses a set of tiers to indicate levels of prosody. The first tier is the word tier, where words should be labelled following spelling or romanization conventions (Jun, 2000; Jun, 2005; Jun, 2011). For this corpus, an additional sentence tier has been added above this. Following the words tier, the Phonological tier (U-tones) marks the phonological boundary tones for the AP, ip and IP (Jun, 2000; Jun, 2005; Jun, 2011). APs are typically marked with LH_a, ips are indicated with T-, and IPs are marked with T%. Below this tier is the Phonetic tier (S-tones) which marks surface tones for the APs, ips, and IPs (Jun, 2000; Jun, 2005; Jun, 2011). Underneath this is the Break Index tier, (BI) which marks the breaks between words or between words and the end of the utterance. The original K-ToBI denoted from 0-3 (Jun, 2000; Jun, 2005), but to include the addition of the ip, this index was expanded to include 0 through 4, where 2 indicates a typical AP break, 3 indicates a typical ip break, and 4 indicates a typical IP break (Jun, 2011). There is also an optional Miscellaneous tier (misc) which can be used to add comments or markings.

Of particular use to this project is the “m” and the “-” markers in the Break Index. When there is a mismatch between the tonal evidence and the actual break disjuncture, a labeler can mark “m” after the break index number to indicate that they are certain there is a mismatch between the prosodic structure and the perceived break. The original version of K-ToBI only included 1m, 2m, and 3m, but with the induction of the ip into K-ToBI, for this paper the label 4m has been added in. In this case, the index has been modified so that now 1m indicates a low degree of disjuncture like that between words within an AP but marked with the tonal structures of an AP (disjuncture up the hierarchy). 2m indicates a medium disjuncture as with between APs but marked either down like an AP or up like an ip. 3m describes a high disjuncture as found between ips but marked prosodically either down like an AP or up like an IP. Finally, 4m indicates the highest degree of disjuncture like those which should be between IPs, displaying tonal marking down the hierarchy like an ip.

As for the “-”, in the original K-ToBI (Jun, 2000), this marker was added after the Break Index number to indicate transcriber uncertainty about the degree of disjuncture so, for instance, a 2- would indicate uncertainty between a 1 and 2 break. However, in the revised K-ToBI (Jun, 2007; Jun, 2011), this symbol appears to have been repurposed to indicate ip boundaries such that 3- would indicate a break between ip boundaries. This has led to a mismatch between symbol uses as a 3- technically now indicates both an ip disjuncture and a 2 or a 3-level break. At least for now, 3- in this corpus has been used to indicate ip break and not uncertainty; however, this is an issue which is worth returning to in the future. Perhaps it may be prudent enough to simply mark ips with a 3 and leave 3- to indicate uncertainty. Further scrutiny is required to determine if that is the right direction to go.

3 Methods

3.1 Background Corpus Information

The conversations taken for this corpus are recordings of example conversations created for the Kleartext Korean Language Series Edition 2 for Beginning 1 (Cho, et. Al., 2010). These conversations were recorded by Korean linguistics graduate students and Korean language instructors in the Korean Section of the East Asian Languages and Literatures Department at the University of Hawai'i at Mānoa (M.-j. Park, personal communication, November 14, 2019). These conversations were designed to be speech samples for lower-level students in Korean language courses, thus, they are good samples of slow, clear speech, in addition to being samples of how Korean beginning instructors modulate those features. These modulating features include placing emphasis on key contrasts between segments which are difficult to perceive as a non-native speaker or giving extra emphasis to key vocabulary or grammar being learned in each specific chapter.

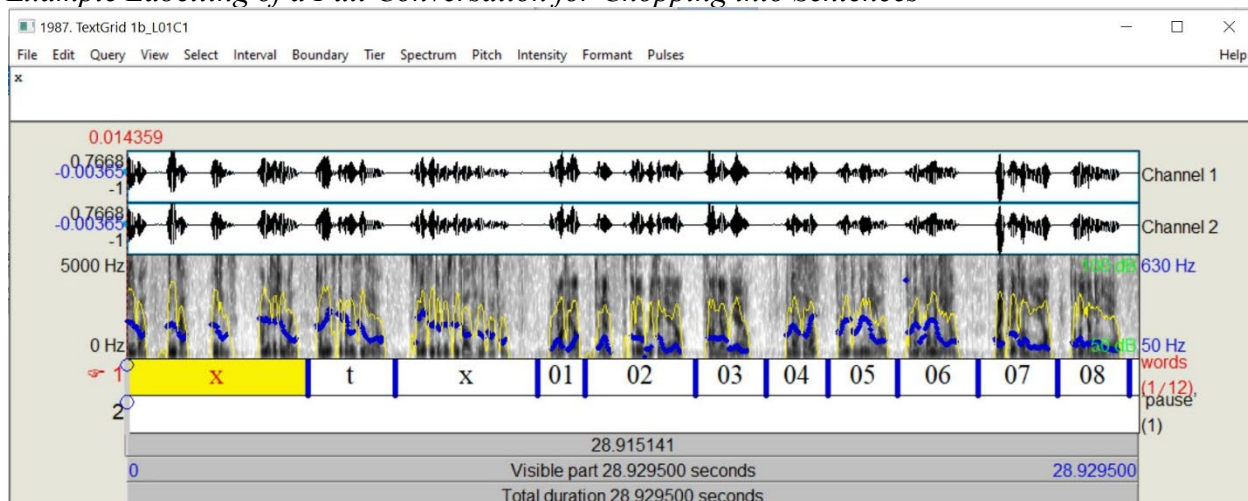
3.2 The Creation of the Corpus

The following section will go into the specifics of creating this corpus in detail. Scripts and text files used are available upon request. First, all conversations were downloaded from the Kleartext website. Conversations for this textbook were recorded in .mp3 format, so before beginning any annotations or file chopping, the sound files need to be converted into .wav format to prevent sound quality degradation. The long sound files were converted from .mp3 to .wav format using a Praat script.

After the long sound files were converted, the conversations were labelled with a TextGrid creation script. Utterances were divided by marking intervals at pauses between them. There are three main kinds of utterances in the conversations: 1) conversation context spoken in English, 2) title information in Korean, and 3) sentences of the actual conversations. To denote English utterances, 'x' was used. To denote title information, 't' was used. To denote sentence information, '01-99' were used, where '01' indicated the first uttered sentence '02' indicated the second, and so on and so forth as noted in Figure 1.

Figure 1

Example Labelling of a Full Conversation for Chopping into Sentences



Annotated TextGrids were then used to chop the conversations into smaller files via another Praat script. These chopped files were saved with the sentence number or type of sentence

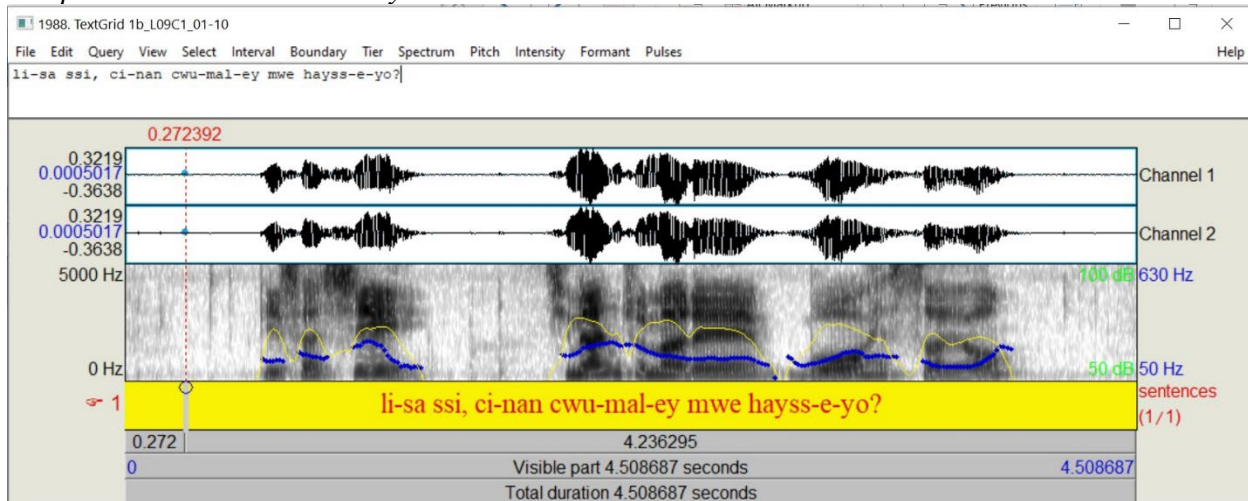
label (x, t, or 01-99) and the time in seconds in which the utterance began in the long sound appended to the end of the original file name.

Before continuing in the annotation process, it was necessary to then create a transcription of the conversation contents. All conversations can be found in Korean using the Hangul script in the respective textbooks. Thus, these conversations were copied from the textbooks into a text file and then converted into the Yale System of romanization, to allow for the use of a K-ToBI default tones annotator script modified from Yoon (2006) who used the original K-ToBI annotation system for their defaults. This script requires the use of romanization to run. After converting conversations into romanized script, each word needed to be broken down into syllable blocks with dashes indicating where syllables you usually be marked in Korean Hangul script. These syllable block indicators allow for the K-ToBI default tones annotator script to add in the appropriate defaults for the existing number of default TBUs. To allow for this script to work, punctuation for sentence endings must be included and spaces need to be indicated between each word.

Once the transcriptions were adequately romanized, then each chopped sound file sentence was annotated to a TextGrid using another Praat script, as can be seen in Figure 2. Sentences were added one by one to check for differences in both the textbook script compared to the actual recording, and also to check for differences between how the sentences were chopped as opposed to how the conversations chunked sentences in the textbooks. Instances of replacing or dropping words or particles were found, along with instances of mostly utterance leading yes's, no's, conjunctions, or discourse markers where periods in the text corresponded to prosodic AP or ip breaks rather than IP breaks, or the reverse, where commas in the text corresponded to IP sentence breaks rather than AP or ip breaks.

Figure 2

Sample Sentence Label with Syllable-delineated Yale Romanization



Next, using the K-ToBI default tones annotator Praat Script, each of the sentences was labelled by the author following the revised K-ToBI conventions. This script works by first creating an interval tier and chunking each word delineated by a space into their own intervals, asking the annotator to realign the boundaries for the words, and then inserting the default K-ToBI tones (phonology tones first, then phonetic tones) based on syllable numbers, syntactic punctuation markers, and the AP-leading consonant (which determines if an AP begins L or H), and then

inserting the standard break indexes based on whether the tunes are labelled as APs or IPs. Once the defaults were added in and the words adjusted for intra-sentential pauses, the default labels were adjusted to the actual tones in the utterances.

After K-ToBI annotating each of the utterances, another script was used to open the annotated TextGrid and corresponding .wav files to write out the data from each annotation into a spreadsheet and double check the annotated work for errors. Data extracted included the speaker gender, the sentence type (declarative, command/request, prepositive, interrogative, other), the word count, the AP count, ip count, the IP count, the boundary tone(s), the full tune, the full break index, a rough rating for naturalness of the intonation (1-5, with 5 being very natural), and for any other comments on the sentence.

4 Corpus Results

4.1 Descriptive Statistics

In the complete corpus for Beginning I, there are a total of 170 sentences and 16 conversations. The length of both the conversations and the sentences themselves generally increases as the series goes on. Looking at the sentence type make-up of the corpus, sentences were labelled as either a declarative, command/request, proposition, interrogative, or other. The other category mostly ended up comprising greetings such as hello, goodbye, or name acknowledgements and discourse markers comprising the entire sentence. The overall majority of sentences were declaratives at 96 total. The second most common type of sentence were interrogatives at 58 total. This was followed by commands/requests 7, other at 7, and propositions 2. This lack of commands, requests, and propositions is most likely due to the limited syntax available to use in the examples. There was a total of 15 single word sentences.

Table 1

Number of Words per Sentence

Sentence Length	Number of Sentences at this Length
1 word	13
2 word	53
3 word	48
4 word	36
5 word	29
6 word	12
7 word	6
<i>Average Sentence Length</i>	<i>3.6 words</i>

In terms of the number of words, there were 13 single-word sentences, 53 two-word, 48 three-word, 36 four-word, 29 five-word, 12 six-word, and 6 seven-word sentences in Beginning 1 for an average of 3.6 words per sentence as seen in Table 1. A word was counted as a lexical item plus its attached particles, but the copula, which typically attaches to the final noun, was counted

as its own word. Looking at the speaker make-up, 97 sentences from Beginning 1 were spoken by female speakers and 73 by male speakers as seen in Table 2.

Table 2

Number of Sentences by Gender

Speaker Gender	Number of Sentences
Women	97
Men	73

Naturalness ratings have been given based on K-ToBI labeler intuition. In terms of intonational naturalness rating 70% of all sentences produced in the Beginning 1 sample set are perceived as unnatural (given a 1 or 2) with the remaining 30% being considered majority neutral to acceptably natural.

Table 3

Number of Sentences per Unnaturalness Rating

Intonation Naturalness Rating	Number of Sentences	Unnaturalness Groups
1 (Very Unnatural)	52	Completely Unnatural 70%
2 (Unnatural)	67	
3 (Neutral)	45	Neutral to Very Natural 30%
4 (Natural)	6	
5 (Very Natural)	0	

Adding in a speaker gender factor, male-speaker-uttered sentences were found slightly less natural than female-speaker-uttered sentences. Combining the two largest rated categories for both, 80.8% of sentences uttered by male speakers were perceived as a 1 or 2 level naturalness, while 71% of sentences by female speakers were rated as a 2 or 3 level naturalness, shown in Table 4. These gender by naturalness statistics indicate that in these samples male SCS speech may be perceived as more unnatural overall than female speech. This is perhaps thanks to the general characteristics of greater pitch range and higher overall pitch, combined with greater pitch movement which is already perceived as more natural in female Korean speech as opposed to male Korean speech which tends to be flatter and have less pitch excursions (Jun & Fletcher, 2014). More research is necessary to confirm this trend.

Table 4

Unnaturalness Ratings by Speaker Gender

Intonation Naturalness Rating	Women (Number of Sentences Spoken by)	Men (Number of Sentences Spoken by)
1 (Very Unnatural)	23 (23.7%)	29 (39.7%)
2 (Unnatural)	37 (38.1%)	30 (41.1%)
3 (Neutral)	32 (32.9%)	17 (17.8%)
4 (Natural)	5 (5.1%)	1 (1.4%)
5 (Very Natural)	0 (0%)	0 (0%)

In terms of the boundary tones, as seen in Table 5, there were 67 rising tones and 159 falling tones, coming to a total of 226 boundary tones as can be seen in Table 4. Many of these extra 56 boundary tones (compared to total sentences) came from sentence leading yes, no, or conjunctions. This is somewhat natural for yes or no to be pronounced as their own leading IP. However, conjunctions are much more unnatural to pronounce with this degree of disjuncture. Many of the 3m, 3-, 4m, and 4 break index leading sentences were actually led by a yes, a no, a conjunction, or another discourse marker such as ‘ah’.

Table 5
Count of Each Boundary Tone

Boundary Tone	Boundary Tone Count	Rising/Falling Count
H%	13	Rising Tones 67
LH%	53	
HLH%	1	
L%	84	Falling Tones 159
HL%	35	
LHL%	38	
HLHL%	2	
Total	226	

Looking at rising boundary tones, there were 13 H%, 53 LH%, and 1 HLH%, where LH% consisted of 79.1% of all rising boundary tones. For falling tones, there were 84 L%, 35 HL%, 38 LHL%, and 2 HLHL%, where 52.8% of all falling tones were L%. LH% is a very common tune in Korean, especially for default questions, however the overt prominence of L% is surprising. L% is not nearly as common in conversational speech and is usually more proportionately balanced by other falling tones, as can be compared to findings in a corpus of conversational speech by Fox (2018). This L% prevalence is perhaps one defining aspect of slow, careful speech as seen in this corpus. This is especially evident when compared to the degree to which L% ending sentences were marked at a naturalness rating of 1 ‘very unnatural’. 32 of all of the items given a 1 (52 total) were marked with L% boundary tone, or 61.5% of all 1’s. HL% was second most common, at 14 items rated at a 1, LH% following with 10 items, LHL% with 4, H% with 2, and finally HLHL% with 1 item rated at a naturalness of 1.

4.1.1 Common Attributes of Slow, Careful Speech (SCS)

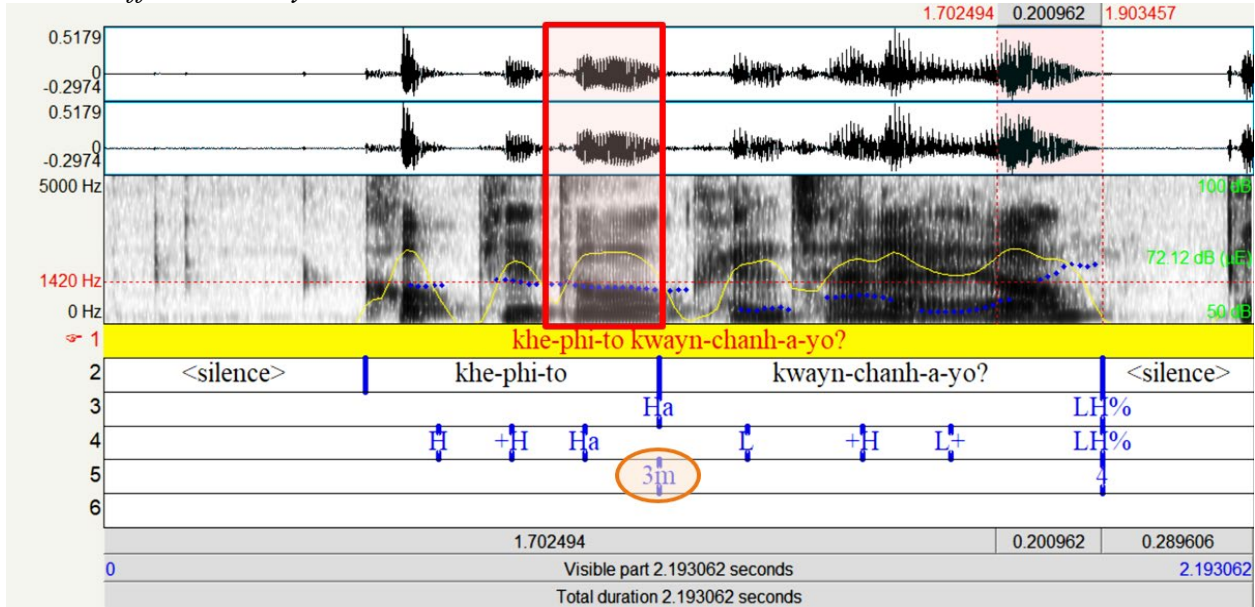
In an analysis of the features that make Korean SCS distinct, four common themes emerged. First, there is unusual lengthening on APs. Second, breaks are more disjointed than normal speech, adding unnatural pauses/breaks between words and APs. Third, there is much more pitch reset and use of focus in small syntactic frames. And fourth, boundary tones are typically flatter, and more disaffected.

Looking at AP lengthening and break disjuncture, in the following example, Example 1 ‘Is the coffee also okay?’, there is extra length on the final *-to* ‘also’ leading to a mismatch between the prosodic evidence and the break index. There were many instances where the AP was lengthened leading to break index disjuncture, thus this is something that should be measured

numerically later to see how long the lengthening becomes on APs in SCS. Additionally, Example 1 exhibits disjointed breaks, with added pausing after the *-to* ‘also’.

Example 1

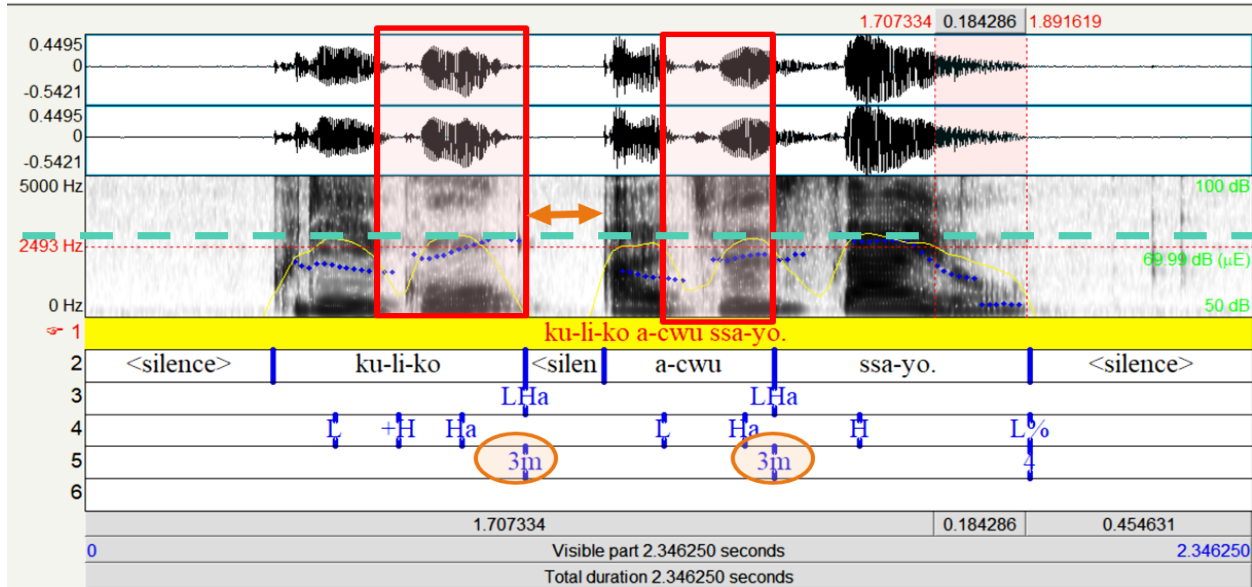
“*Is the coffee also okay?*”



The following example, Example 2 “And it’s very cheap.” exhibits AP lengthening, unusually break disjuncture, and pitch reset. There is AP lengthening on the final syllables of *kuliko* ‘and’ and *ajwu* ‘very’ in addition to added pauses after the two APs, especially prominent after *kuliko* ‘and.’ This leads to distinct prosodic blocking between words with pausing similar to what is expected of an ip after each AP. Further proof that this blocking is more ip-like than AP-like is provided in the use of pitch reset after the APs. It is unusual to find this degree of pitch reset, pausing, and lengthening on APs, especially given the short length of the utterance as a whole.

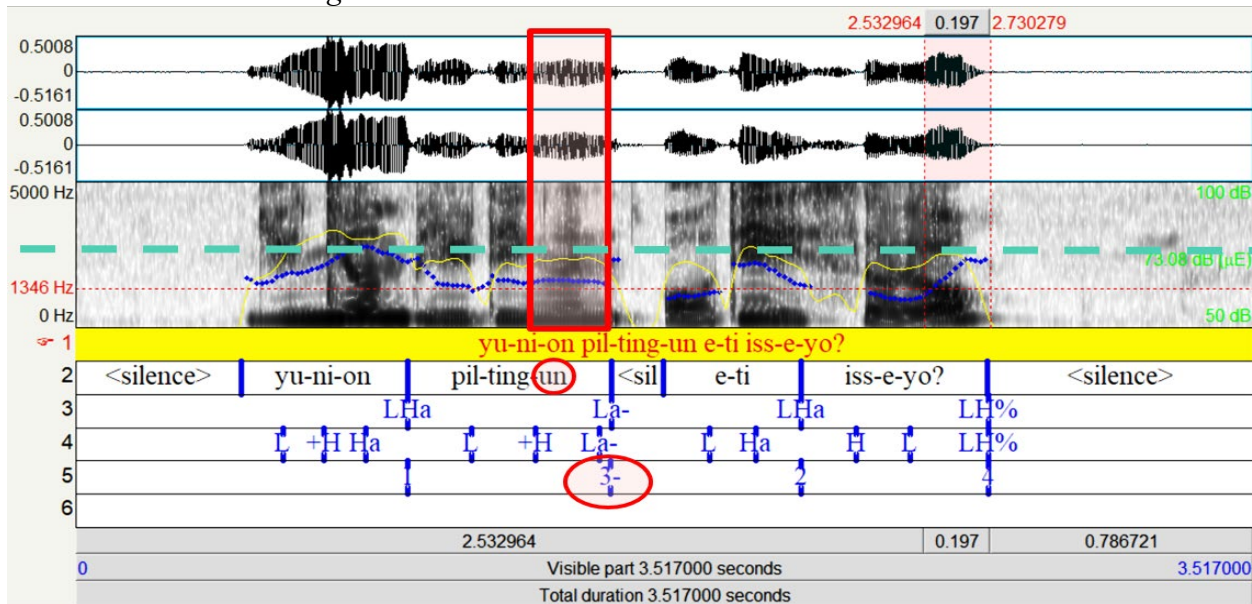
Example 2

“*And it’s very cheap.*”



Another example of ip use in a small syntactic frame that would be unlikely in conversational speech can be seen in Example 3 “Where is Union Building?”. In this example, the ip is signaling focus onto the syntactic significance of the topicalizing particle *-un* ‘topic marker’ which usually would only receive an AP level marking. This particular example is used in the textbook to draw learner attention to the function of *-un* ‘topic marker’ as a contrast marker. This marking then resets the pitch following the topic, further marking this AP like an ip.

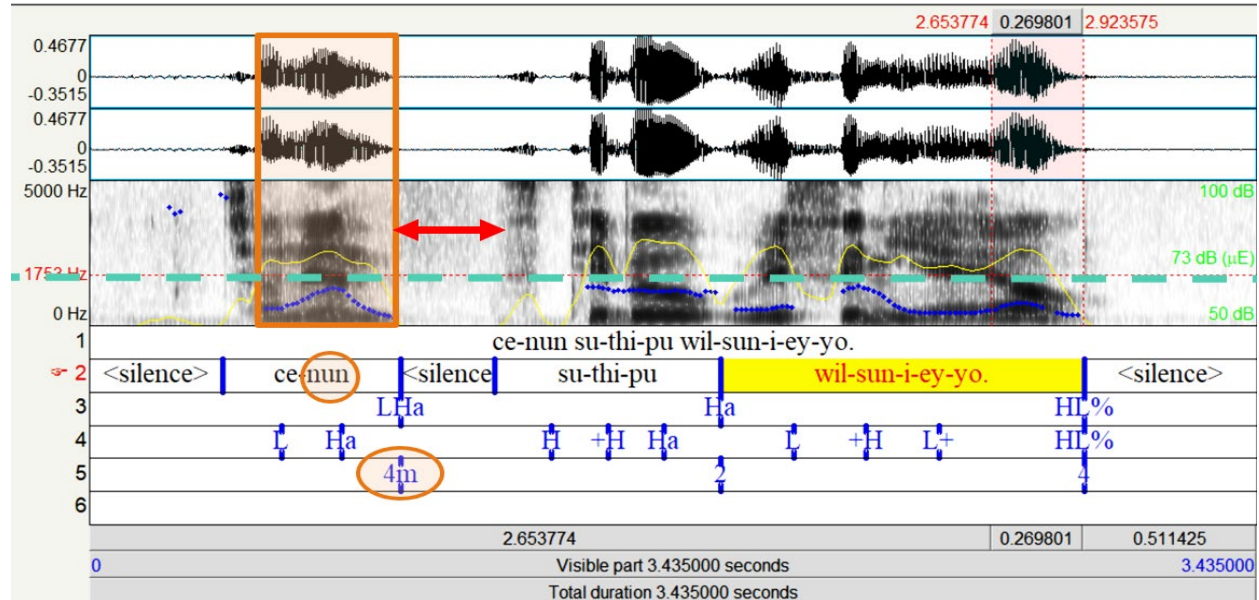
Example 3
 “Where is Union Building?”



The following Example 4 “I am Steve Wilson.” Exemplifies unusual break disjuncture and focus. Here there is even stronger breaking after the *-un* ‘topic marker’ AP phrase, where it is used as an ip with markings of an IP break when the small syntactic domain typically would call for AP

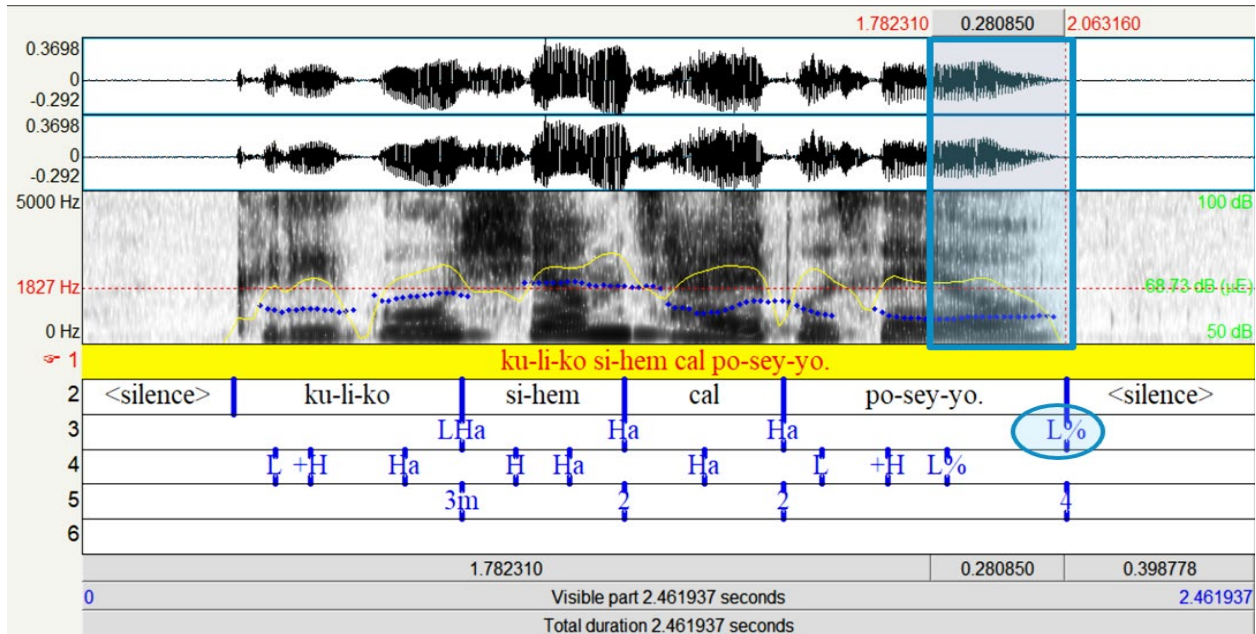
features. The ip here is signaling the focus of the topicalizing particle *-un* with overtly strong marking. Additionally, there is a long pause after the topic particle, which then leads to pitch reset on the following AP *suthipu* ‘Steve.’

Example 4
 “I am Steve Wilson.”



Finally, in terms of the use of flatter, less affectual boundary tones, such as the over-use of L%, Example 5 “And do well on your exam.” depicts the use of a long, flat L% boundary tone where one would expect to see a LH%, HL%, or even LHL% based on the tone’s length and context. This leads to a very effectually bland realization of the utterance and a mismatch between the expected boundary tone complexity and the produced tone. The use of the L% could be considered more natural if the boundary tone were much more truncated, but this tone here is proportionately long compared to the affect it lacks. This theme was found throughout the dataset, where L% was overrepresented compared to more complex tones such as HLH%, LHL%, LHLH%, HLHL%.

Example 5
 “And do well on your exam.”



5 Discussion and Pedagogical Implications

The intonational map being shown to especially beginning-level Korean students is guided largely by these intonational attributes of SCS examined in the previous section. This is especially the kind of speech that instructors are using when trying to slow down their speech rates for students to understand clearly. Instructor intonational accuracy and variety can often be sacrificed for segmental, lexical, or grammatical clarity especially at the beginning levels, but with the knowledge of how SCS effects prosody, instructors can more effectively produce more intonationally-varied speech examples. Thus, to help students acquire more comprehensive intonational patterns at earlier stages in the L2 acquisition process, intonational variety and accuracy must be included into classroom routines.

Intonational accuracy and variety at even beginner levels is critical to proper intonational development as there is growing evidence that beginner speakers will internalize frequent and salient intonation patterns regardless of intelligibility (Chang 2012; 2013). Additionally, according to Exemplar Theory (Johnson, 2007; Pierrehumbert, 2001, 2002, 2006; Wedel, 2006), instances of language input or production are stored with varying weights based on relative saliency or frequency etc. This indicates that students will recall intonational patterns they frequently hear or find salient for some reason, such as uniqueness or strong emotion attached to the memory (such as with a recast or immediate correction). Instructors should take care to ensure that any prosodic cues used to draw attention to other newer linguistic features are not overly repeated to prevent internal conceptualization of the pattern's prosody as the SCS style exclusively.

Given the importance of incorporating mindful intonational use into the classroom, I recommend the following actions be integrated into regular language course lectures and activities. First and simplest, instructors should show natural speech examples often from various speakers. This can be done by simply providing students with a variety of natural video or audio clips in class, or creating assignments that integrate a variety of speakers to complete at home. Active listening assignments would nicely incorporate intonation into regular content instruction. Second,

I recommend that instructors give explicit instruction and direct feedback on intonation wherever possible. I recommend giving instructions on how students can pause naturally, how use of a particular intonation pattern will make them misunderstood, and giving alternate examples of how intonation changes the meaning of a sentence. I recommend also taking time to instruct students simply in how to group phrases together, and giving examples of high affect speech (i.e. samples in a context where there is clear emotion incorporated into the intonation) regularly during class time, among other explicit explanations. Additionally, intonational cues can be taught to help students distinguish the differences between the three-way plosive system in Korean, since APs which begin with a lax consonant begin with a low tone, while those which begin with a tense or aspirate begin with a high tone. This would provide beginning students with another tool to help them navigate this common difficult phonemic contrast.

To facilitate intonation acquisition, I recommend utilizing call and repeat activities and recasting student speech into more natural intonation to help build a beginning level student's awareness of intonation and build their Korean intonational map. The focus on instructing intonation should be intelligibility, and so recasting and repetition speaking activities etc. should focus on cases where incorrect intonation can cause intelligibility errors. For example, it is helpful to explicitly explain the LHa an HHa AP tunes as aspirated & tense vs lax, or to explain the role of intonation in the three way ambiguity phenomena found in the use of wh-questions words (e.g. *mwe hesseyo* 'what did' being able to be interpreted as "what did you do?", "did you do something?", and "I did something." all of which's interpretations hinge on differentiated intonational realizations).

6 Conclusions

In conclusion, the four global attributes of the prosody of Korean Slow, Clear Speech include final lengthening on APs, presence of disjointed breaks and pauses between APs, common use of ip-like pitch reset and use of focus in small syntactic frames, and overall, less use of affect through shorter, less-complex boundary tones. The hypotheses made at the beginning of this inquiry have been found to be generally correct. There were increased breaks where they would not usually appear, there were mismatches between the intended prosodic category and the break produced, there were less affectual boundary tones with incredibly limited or non-existent use of longer tones such as HLH%, LHL%, LHLH%, and HLHL%, and there was a mismatch between the expected boundary tone complexity and the actual one used. Somewhat unexpectedly, gender appears to play a role in the overall perceived naturalness, with male speech being slightly more likely to be considered unnatural sounding. This is likely due to the extra effort it requires to produce a higher pitch and greater pitch range for male speakers as compared to female speakers. This merits further investigation beyond labeler interpretation, however.

The intonational map being provided to beginning Korean students is guided largely by these intonational attributes of SCS. This inquiry into Korean SCS ultimately can help inform instructors on how to effectively include intonational variety and accuracy into their classroom routines and lesson plans to help students acquire more comprehensible intonational patterns at earlier stages in the L2 acquisition process. Often in the beginning levels of language classes, instructor intonational accuracy is sacrificed for segmental or lexical clarity, but if we can look into how slow, hyper-enunciated speech effects prosody, we can more effectively produce more intonationally-accurate speech samples without sacrificing the articulation of phonemes or new

grammar patterns and vocabulary. Instructor intonational accuracy and variety can be maintained through the use of conscious intonation use by the teacher, through native speech samples, and through explicit intonation instruction and feedback. Understanding how SCS affects instructor speech can help instructors mitigate their effects and help students acquire more comprehensive intonational patterns at earlier stages in the L2 acquisition process. This paper proposed preliminary analyses and instructional methods, however further work is necessary to see which methods of intonational instruction and in which contexts materializes the best results, in addition to further research into the more fine-grained features of Korean SCS.

References

- Chang, C. (2012). Rapid and multifaceted effects of second-language learning on first-language speech production. *Journal of Phonetics*, 40(2), 520-533.
- Chang, C. (2013). A novelty effect in phonetic drift of the native language. *Journal of Phonetics*, 41(6), 520-533.
- Chang, S.-e. (2017). Enhancement effects of clear speech and word-initial position in Korean glides. *The Journal of the Acoustical Society of America*. 141, 4188. doi: 10.1121/1.4984061.
- Chen, S.-H., Hsieh, C.-H., Chiang, C.-Y., Hsiao, H.-C., Wang, Y.-R., Liao, Y.-F., & Yu, H.-M. (2014). Modeling of speaking rate influences on Mandarin speech prosody and its application to speaking rate-controlled TTS. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*. 22, 7.
- Cho, Y.-m, Lee, H.-s., Schulz, C., Sohn, H.-m. & Sohn S.-o. (2010). *Integrated Korean: Beginning I*. 2nd ed. Honolulu: University of Hawai'i Press.
- Fougeron, C. & Jun, S.-a. (1998). Rate effects on French intonation: prosodic organization and phonetic realization. *Journal of Phonetics*. 26, 45-69.
- Fox, B. J. (2018). Korean wh-question boundary tones: A systematic corpus-based tone categorization. General Sessions (Phonology). 2018 Silence Workshop/Linguistics Society of Korean Conference. Kyung Hee University, Seoul, South Korea.
- Johnson, K. (2007). Decisions and mechanisms in exemplar-based phonology. In Solé, M.-J., Beddor, P. S., and Ohala, J. J., editors, *Experimental Approaches to Phonology*, 25–40.
- Jun, S.-a. (2003). The effect of phrase length and speech rate on prosodic phrasing. in the Proceedings of the 15th International Congress of Phonetic Sciences, Barcelona, Spain, ed. by M. J. Solé, D. Recasens, and J. Romero; Barcelona 2003, ISBN 1-876346-48-5; ICPhS Archive, <http://www.internationalphoneticassociation.org/icphs/icphs2003>. 483–486.
- Jun, S.-a. (2000). K-ToBI (Korean ToBI) labelling conventions - Version 3. *The Korean Journal of Speech Sciences*, 7, 143-169.
- Jun, S.-a. (Ed.). (2005). *Prosodic typology: The phonology of intonation and phrasing*. Oxford: Oxford University Press.
- Jun, S.-a. (2007). The intermediate phrase in Korean intonation: Evidence from sentence processing. In *Tones and Tunes: Studies in Word and Sentence Prosody*, eds. by C. Gussenhoven and T. Riad. Berlin: Mouton de Gruyter. 143-167.
- Jun, S.-a. (2011). Prosodic markings of complex NP focus, syntax, and the pre-/post-focus string. In the Proceedings of the 28th West Coast Conference on Formal Linguistics (WCCFL) (eds. by Mary Byram Washburn, Katherine McKinney-Bock, Erika Varis, Ann Sawyer, and Barbara Tomaszewicz). www.lingref.com, document #2454. 214-230.
- Jun, S.-a. (Ed.). (2014). *Prosodic typology II: The phonology of intonation and phrasing*. Oxford: Oxford University Press.
- Jun, S.-a. & Cha, J.-h. (2015). High-toned [il] in Korean: Phonetics, intonational phonology, and sound change. *Journal of Phonetics*. 51, 93-108.
- Ko, E.-s. & Soderstrom, M. (2013). Additive effects of lengthening on the utterance-final word in child-directed speech. *Journal of Speech, Language, and Hearing Research*, 56, 364–371.
- Krause, J. C. & Braid, L. D. (2002). Investigating alternative forms of clear speech: The effects of speaking rate and speaking mode on intelligibility. *The Journal of the Acoustical Society of America*, 112, 2165. doi: 10.1121/1.1509432.

- Krause, J. C. & Braidia, L. D. (2004). Acoustic properties of naturally produced clear speech at normal speaking rates. *The Journal of the Acoustical Society of America*, 115, 362. doi: 10.1121/1.1635842.
- Kuder, E. (2017). En voz alta: Prosodic Features of “Teacher Talk” in the Spanish Second Language Classroom (ProQuest Dissertations Publishing). Retrieved from <http://search.proquest.com/docview/1925346102/>.
- Park, M.-j. (2019, November 14). Personal interview.
- Pierrehumbert, J. (2001). Exemplar dynamics: word frequency, lenition and contrast. In Bybee, J. and Hopper, P. (eds) *Frequency and the emergence of linguistic structure*, (pp. 137–158). John Benjamins, Amsterdam.
- Pierrehumbert, J. (2002). Word-specific phonetics. In Gussenhoven, C. and Warner, N., editors, *Laboratory Phonology VII*. Mouton de Gruyter, Berlin.
- Pierrehumbert, J. B. (2006). The next toolkit. *Journal of Phonetics*, 34 (4), 516–530.
- Ramirez-Aristizabal, A. G., Médé, B., Kello, C. T. (2018). Complexity matching in speech: Effects of speaking rate and naturalness. *Chaos, Solitons & Fractals*, 111, 175-179. <https://doi.org/10.1016/j.chaos.2018.04.021>.
- Yoon, K.-c. (2006). Hankwuke thopi phyokilul wihan phulas sukhulipthu thwul [Unified Praat Script Tools for Facilitating Korean ToBI Annotation]. *Inmwun nonchong [Humanities Collection]*, 56, 1-63.
- Wedel, A. (2006). Exemplar models, evolution and language change. *The Linguistic Review*, 23 (3), 247–274.