Central Lancashire Online Knowledge (CLoK)

Title	Prosodic marking of narrow focus in Seoul Korean
Type	Article
URL	https://clok.uclan.ac.uk/id/eprint/15765/
DOI	10.17863/CAM.6795
Date	2017
Citation	Jeon, Hae-Sung and Nolan, Francis (2017) Prosodic marking of narrow focus in Seoul Korean. Laboratory Phonology, 8 (2). pp. 1-30. ISSN 1868-6346
Creators	Jeon, Hae-Sung and Nolan, Francis

It is advisable to refer to the publisher's version if you intend to cite from the work.
10.17863/CAM.6795

For information about Research at UCLan please go to http://www.uclan.ac.uk/research/

Title: Prosodic marking of narrow focus in Seoul Korean

Full authors' names: Hae-Sung Jeon[1] and Francis Nolan[2]

Affiliations:

[1]School of Language and Global Studies, University of Central Lancashire, Preston, UK

[2]Department of Theoretical and Applied Linguistics, University of Cambridge, Cambridge, UK

Corresponding author's email address: HJeon1@uclan.ac.uk

Abstract

This paper explores prosodic marking of narrow (corrective) focus in Seoul Korean. Korean lacks lexical stress and it has a phonologized association between the Accentual Phrase (AP) initial segment and intonation. In the experiment, four speakers read sentences including a two-item list which were designed to elicit either an L or H AP-initial tone. The durational variations, the pitch events at prosodic boundaries, and F_0 span in 32 sentences read neutrally and 64 sentences read with one of the items under focus were analysed. The results show that the focused constituent consistently initiates a new prosodic phrase. In comparison to the neutrally spoken or defocused counterpart, the focused constituent was more likely to be realised as an Intonational Phrase (IP) in some contexts. Bitonal IP boundary tones were more likely to occur under focus than monotonal tones. In addition, in focused constituents, durational expansion particularly at the phrase-edges, expansion in F_0 span, and raising of the phrase-initial pitch were observed. On the other hand, defocused constituents were not phonetically reduced compared to the neutral counterparts. The results imply that the phonetic cues spreading over the focused constituent complement the exaggerated prosodic boundaries.

Keywords: Korean, prosody, intonation, focus

1. Introduction

The information structure in a given discourse influences the phonetic shape of an utterance. For example, across languages, emphasized elements tend to be realized with an increase in articulatory effort such as increased duration, amplitude and pitch excursion size (e.g., Gussenhoven, 2002; Gussenhoven, 2004). Our current understanding of the relationship between information structure and the speech signal is significantly influenced by West Germanic languages. In West Germanic languages, the location of pitch accents

contributes to marking information structure, and gradient variation in the realisation of a pitch accent, mainly in its pitch excursion and duration, convey the degree of emphasis (e.g., Baumann, Grice and Staindamm, 2006). Pitch accents are associated with lexically stressed syllables.

However, the assumption that there is a direct link between focal structure and pitch accent distribution in all languages is empirically inadequate (see Ladd, 2008, Ch. 7). Even in the West Germanic languages, the presence of a pitch accent by no means always signals focus, and in other languages such as French or Japanese without lexical stress, the concept of a focal pitch accent is not clearly applicable. In these languages, phrasing and increase in the acoustic salience in focused constituents have been reported as prosodic markers of narrow focus (see Section 1.2). Previous studies on Korean, which is another language without lexical stress (see Jeon, 2015 for a recent survey), reveal that Korean shares some similarities with French and Japanese in that focal structure directly affects prosodic structuring. However, the prosodic marking of information structure in Korean has not previously been thoroughly described.

This paper explores the prosodic markers of one kind of narrow focus, corrective focus, which is related to rejecting and correcting what has been already said in a conversation (see Gussenhoven, 2007) in Seoul Korean (henceforth Korean). Narrow focus here refers to a word-sized unit being highlighted, as opposed to broad focus on a larger unit (see, for example, Ladd, 2008, p. 215). Although there are morphological or syntactic means for focus marking (see Féry, 2013 and references therein), narrow focus is marked prosodically in many languages (for example, see Baumann, Grice and Steindamm, 2006 for German, Féry, 2001 for French, Venditti, Maekawa and Beckman, 2008 for Japanese, Wang and Xu, 2011 for Mandarin Chinese).

The broad aims of the study are as follows. First, we aim to investigate the manifestation of an equivalent to 'focal accent' in Korean, a language which lacks word-level stress. The Western-Germanic type of pitch accent associated with lexical stress and focal prominence is theoretically undefinable in Korean (e.g., Jun and Fougeron, 2000). Unlike in West Germanic languages where focus triggers intonational events in prosodic phrases (see Section 1.2), in other languages such as French and Japanese with unclear word-level prominence, speakers' adjustment of the phonetic shape is more apparent near the prosodic phrase edges (see Sections 1.2 and 1.3). The study of a language such as Korean would therefore contribute to understanding the relationship between the (lack of) word-level prominence and the higher-level prosodic organisation related to the focal structure.

Second, we aim to examine focus-related intonational variation in Korean.[1] The F_0 contour shapes over prosodic phrases in Korean show a wide range of variation (see Section 1.1). This variation poses complications in speech data analyses and the majority of previous studies limited their scope to F_0 contour shapes common at the level of the Accentual Phrase (AP) such as LHLH or HH (Jun and Lee, 1998; Jun and Kim, 2007; Lee and Xu, 2010; Cho, Lee and Kim, 2011; Yang, Cho, Kim and Chen, 2015). As a consequence of this experimental control of contours, it is still not clear whether focal structure is one of the factors determining the F_0 contour shapes (e.g., the location of F_0 turning points in a phrase) and phonetic shape of the right edge of the phrase, and whether the segmentally induced AP-initial tones (reviewed in Section 1.1) interact with the focus-related prosodic variations. Furthermore, there has been little systematic comparison between neutral, focused and defocused constituents in speech.

Third, we will quantify speakers' durational adjustment related to focal structure. In West Germanic languages, lengthening of the word carrying the pitch accent – with the greatest magnitude of such lengthening associated with the stressed syllable (e.g., Cambier-Langeveld and Turk, 1999; Turk and White, 1999; Dimitrova and Turk, 2012) – would serve as an important cue to the focus. In Korean, on the other hand, it is difficult to delineate the extent of such accentual lengthening due to the lack of a definable lexical stress and pitch accent.

1.1. The Prosodic Hierarchy in Korean

We adopt Jun's analysis of the prosodic structure (Jun, 1996a, 1998, 2000, 2005, 2006, 2012) which is currently the most widely used model for prosodic analysis in Korean. Jun's model is based on the Prosodic Hierarchy (Selkirk, 1984, 1986; Nespor and Vogel, 1986), and represents intonation with L and H targets following the Autosegmental-Metrical theory (see Ladd, 2008). In addition, Jun (1996a, 1998, 2000, 2005, 2006, 2012) classifies Korean as a language without any head prominence related to lexical stress, pitch accent or tone, but a language with edge prominence. Four levels above the lowest unit, the syllable, are defined in the Prosodic Hierarchy: the Phonological Word (PW), the Accentual Phrase (AP), the Intermediate Phrase (ip), and the Intonational Phrase (IP). Although the ip was added in a later revision (Jun 2006), demarcating ips in speech data is not straightforward and therefore the ip is excluded in the following discussion.[2]

In Jun's model (1996a, 1998, 2000, 2005), the PW does not have any prosodic specification and the AP is the basic unit for prosodic analysis. The AP is a word-sized unit defined with particular reference to the pitch contour (see Schafer and Jun, 2002; Jun and Kim, 2004).[3] An AP tends to have 3–4 syllables and 1.14–1.2 content words on average (Kim, 2004; Kim, 2009). Jun (1996a, 1998, 2006) proposes that the AP has the underlying tonal pattern

THLH, where the realisation of the initial tone (T) tends to depend on the laryngeal configuration of the phrase initial segment. When the initial segment is a fortis (/p*, t*, k*, ts*/) or aspirated (/p^h, t^h, k^h, ts^h/) consonant, /s/ or /h/, the initial tone tends to be H, but otherwise it is L. Segments triggering AP-initial H and segments triggering AP-initial L (i.e., lenis consonants, nasals, semivowels, and vowels) will hereafter be referred to respectively as strong segments and weak segments. The association between the type of segment and the tonal target (L vs. H) in AP-initial position is considered phonologised in Korean (Jun, 1996a, 1998), although the AP-initial tones are not always predictable (e.g., Jun, 2000; Kim, 2004). Jun (1998) states that all four tones are realized when there are four or more syllables in an AP, but if there are fewer than four syllables, some of the tones are not realised. There are 14 pitch contours of the AP reported in Jun (2000) (i.e., LH, LHH, LLH, HLH, HH, HL, LHL, HHL, HLL, LL, HHLH, LHLH, LHLL, and HHLL) and the last tone, either L or H, is associated with the AP-final syllable. The AP and IP are differentiated by the size of the perceivable disjuncture at the phrase boundary. For instance, the IP is often marked by boundary tones (e.g., L%, H%, LH%, LHL%, etc.) and significant final lengthening. Unlike the AP, boundary tones can consist of two or more tones which can be realised on the IP-final syllable.

The causes of intonational variation in Korean have not been investigated much so far. What is known is that there is a tendency referred to as the see-saw effect in Jun (1996b) which avoids the same type of tones occurring adjacent to each other, that speaking style affects the shape of the pitch contour (Kim, 2004; Kim, 2009; Kim, et al. 2007), and that there is a strong preference for the AP-final H (Kim 2004 and references therein).

1.2. Prosodic marking of narrow focus across languages

Across languages, the constituent under focus tends to involve hyper-articulation of a kind consistent with the Effort Code (Gussenhoven, 2002; Gussenhoven, 2004) and the focus can affect the phonetic shape of the following constituent in the utterance (as in post-focus compression; see below). There are multiple ways of signalling narrow focus prosodically including, for example, an increase in the segmental duration or F_0 span (e.g., Xu, 1999, Chen, 2006 for Standard Chinese; Eady and Cooper, 1986, Xu and Xu, 2005 for American English; Baumann, Grice and Steindamm, 2006 for German; Peters, Hanssen and Gussenhoven, 2014 for varieties in Dutch, Frisian or German); varying the alignment of F_0 turning points (e.g., Xu and Xu, 2005 for American English; Peters, Hanssen and Gussenhoven, 2014 for varieties in Dutch, Frisian or German); creating prosodic breaks by phrasing (e.g., Féry, 2001 for French); using boundary tones (e.g., Venditti, Maekawa and Beckman, 2008 for Japanese); varying pitch accent types (e.g., Baumann, Grice and Steindamm, 2006 for German); and compressing the pitch range, duration, and/or intensity of the constituent preceding or following the focused constituent (e.g., Eady et al., 1986, Xu and Xu, 2005 for American English; Chen, Wang and Xu, 2009 for Taiwanese and Mandarin; Lee and Xu, 2012 for Japanese).

In West Germanic languages such as English, German, and Dutch, the marking of narrow focus is often discussed in relation to the placement and acoustic salience of a pitch accent, which are respectively interpreted to be phonological (i.e., categorical) and phonetic (i.e., showing gradient variation within a category). The common view is that the focused word receives a pitch accent while the defocused counterpart is de-accented (e.g., Beckman and Pierrehumbert, 1986) especially if following the focused item. Alternatively, the type of pitch accent employed may differ between focused and defocused constituents (e.g., Baumann, Grice and Steindamm, 2006 for German). The type of pitch accent may also be determined by the distinction between narrow focus and broad focus, though this

distinction is not always clearly marked (e.g., Baumann, Grace and Steindamm, 2006, Kügler, 2008, Féry and Kügler, 2008 for German). The gradience is related to duration (e.g., Baumann, Grace and Steindamm, 2006 for German) and scaling or alignment in F_0 (e.g., earlier peak in narrow focus in American English, Xu and Xu, 2005; Standard German, Baumann, Grice and Steindamm, 2006).

In languages without clear lexical stress such as French and Japanese, prosodic phrasing is often referred to as a phonological marker to focus. In French, the focused constituent is generally realised in a separate phrase with its own tonal structure and with optional dephrasing of the following constituents, and sometimes with short breaks before and/or after the phrase boundaries (e.g., Jun and Fougeron, 2000)[4]. Phonetic markers of narrow (or contrastive) focus include the raising of the phrase-initial pitch and a higher F_0 peak in the prosodic phrase (Féry, 2001). In Japanese, which has lexical pitch accent, dephrasing or prosodic subordination seems to occur in relation to focal structure (e.g., Beckman and Pierrehumbert, 1986; Gussehnhoven, 2004). For instance, Venditti, Maekawa and Beckman (2008) discuss a variety of intonational means to mark focus in Japanese, including pitch range expansion, F_0 reset at the left edge of the focused constituent which may co-occur with the insertion of the IP boundary at the beginning of the focused constituent, dephrasing (i.e., post-focal prosodic subordination), and boundary pitch movement (H%, HL%). Prosodic subordination (dephrasing) is considered the most crucial way of focus marking in Japanese, and the rest are optional (see Venditti, Maekawa and Beckman, 2008 and references therein).

Narrow focus affects the phonetic shape of the out-of-focus constituent in some languages. Xu, Chen and Wang (2012) classify languages into those with or without post-focus compression (PFC) in F_0 span, intensity and/or duration, whilst the pre-focus constituent

generally does not show systematic variation. PFC occurs independently of whether lexical tone, pitch accent, or lexical stress is present in a language. Languages such as English (Eady and Cooper, 1986; Xu and Xu, 2005), Japanese (Lee and Xu, 2012), and Beijing Mandarin (Chen, Wang and Xu, 2009) show PFC.

In summary, the focused constituent tends to be associated across languages with an increase in articulatory effort. In general, studies of West Germanic languages suggest a strong link between lexical stress and the prosodic manifestation of the focus, whilst the manifestation of focal structure in languages without lexical stress seems to be best described in terms of prosodic phrasing and phonetic events near prosodic boundaries. The presence of post-focus compression is known to be independent of the presence of lexical tone or stress in a language.

1.3. Previous studies on prosodic marking of focus in Korean

Previous studies show that narrow focus in Korean is signalled by phrasing together with other phonetic markers.[5] A focused constituent initiates a new prosodic phrase and tends to be longer than neutral or post-focus counterparts, and the lengthening tends to be greater in magnitude at the edges of the focused constituent; furthermore the lengthening at the left edge seems to be more consistent than that at the right edge (e.g., Chung and Kenstowicz, 1997; Jun and Lee, 1998; Jun and Kim, 2007). Jun and Lee (1998) show that the phrase-initial lengthening under focus is mainly due to extra articulatory strengthening of the phrase-initial segment, although Cho, Lee and Kim (2011) found significant lengthening of the second syllable of focused words. The focused constituent also tends to be spoken more loudly than neutral or post-focus words (Lee and Xu, 2010; Cho, Lee and Kim, 2011) and shows an increased F_0 excursion, whilst the F_0 peak of the post-focus words tends to be compressed (Chung and Kenstowicz, 1997; Jun and Kim, 2007; Lee and Xu, 2010). Jun and

Lee (1998) show that the F_0 peak in the post-focus constituent may not necessarily be lower than that of the neutral counterpart, although it tends to be lower than that of the focused constituent within the utterance.

An important phonological marker seems to be that the focused constituent always begins a new prosodic phrase. In previous studies, the phrase headed by a focused constituent has the status minimally of an Accentual Phrase (AP) (Jun and Lee, 1998) or of an Intonational Phrase (IP) (Jun and Kim, 2007). When the new phrase is headed by a focused constituent, the post-focal constituent, which would have been produced as an AP with acoustic disjuncture at both edges in the neutrally spoken utterance, may be realised without the perceivable acoustic disjuncture at its left edge. That is, the prosodic boundary following the focused constituent may be deleted, but such dephrasing is optional in Korean (Jun and Kim, 2007; Lee and Xu, 2010).[6]

The findings on the out-of-focus constituents are inconsistent. Although Lee and Xu (2010) report PFC in Korean, Jun and Lee (1998) and Yang, Cho, Kim and Chen (2015) offer only inconsistent evidence of the F_0 span compression or shortening of the syllable duration in pre- or post-focus constituents. Jun and Lee (1998) reports mixed results; only three speakers out of five reduced the F_0 peak height compared to neutral speech, whilst the defocused constituents tended to be shorter than the neutral counterpart. In Yang et al. (2015), F_0 span was wider in focused words than in defocused words when the defocused words were pre-focus but not post-focus, although the opposite trend would be expected from the findings in other languages in which PFC has been demonstrated. Yang et al. (2015) also reported an interaction between the size of the focused constituent and the prosodic marking of focus; the focus effect was shown when the target word had four syllables but not with disyllabic words.

1.4. Aims and hypotheses

We explore the prosodic variation associated with information structure. In particular, we investigate prosodic marking of narrow focus not only at the level of AP but potentially at a larger domain, and, complementarily, the prosodic properties of defocused constituents. The prosodic characteristics of Korean lead us to expect that focus marking in Korean may exploit cues similar to those reported in French and Japanese (see Section 1.2). Our expectation follows from shared features of the three languages: they lack lexical stress, and the smallest prosodic unit – the AP in Korean for instance – is demarcated by the F_0 contour (see, for example Welby, 2006 for French, Venditti, 2005 for Japanese, and Jun, 2005 for Korean; note, however, that the AP in Japanese may include a lexical pitch accent). In both French and Korean the lack of lexical stress means there is no culminative prominence marking focus; rather, it is phrasal structure which fulfils the role, with support from overall adjustment of the F_0 span of APs. In addition, the frequent use of a complex pitch movement at the right edge of the IP, which contributes to pragmatic interpretation of the utterance in Japanese (see Venditti, 2005; Venditti, Maekawa, and Beckman, 2008), seems to resemble that in Korean (e.g., Jun, 2005; Park, 2012)

In the experiment, we focus on examining the following specific hypotheses formulated based on the previous findings in Korean and other languages. First, there would be lengthening of the focused constituent and the linguistic units therein, and there would be shortening of the defocused constituent. We expect lengthening spreading over the focused constituent but a larger magnitude of lengthening near the constituent edges. Second, narrow focus would affect the formation of prosodic phrases in that the focused constituent begins a new prosodic phrase. Third, as in Japanese (e.g., Venditti, Maekawa, and Beckman, 2008), the focused constituent may be associated with a higher-level prosodic phrase than

the AP and boundary tones with a complex pitch movement at the right edge. Fourth, the F_0 contour shapes in the AP would be affected by focal structure and speakers' F_0 adjustment in relation to the focal structure may interact with the segmentally-induced F_0 variation. Fifth, F_0 span would be affected by focal structure; the F_0 span would be widest in the focused constituent, whereas the F_0 span would be compressed in the defocused constituent in comparison to the focused or neutrally spoken counterpart. In addition, we expect to observe an association between the AP-initial segment type and tone in speech data as widely reported in previous studies (reviewed in Section 1.1).

2. Experiment

2.1. Experimental materials

There were 32 targets of 5- or 7-syllable sequences of two number units (Table 1). It was assumed that each number unit is a Phonological Word (PW) corresponding to an Accentual Phrase (AP) in speech. The term PW is used to refer to the semantically coherent number unit in the reading materials henceforth (while in speech, PW is produced without perceivable disjuncture at its edges, as discussed in Section 2. 3). They were classified into 4 phrasing types (2 + 3, 3 + 2, 3 + 4, and 4 + 3, where + indicates the PW boundary), and there were two types of initial segment ('weak' and 'strong', see Section 1.1) in the second PW. The location of the strong segments was systematically controlled to have APs beginning with a low tone (L) or a high tone (H) in the dataset. Each PW was designed to include between two and four syllables, as 2-, 3-, 4-syllable APs tend to occur frequently. It was intended to have experimental materials that were meaningful and familiar to speakers, while the possible confounding effect of the morpho-semantic structure was controlled. All constituents of the targets were monosyllabic numbers referring to, for example, 1 (/il/), 10 (/sip/), 100 (/pɛk/), 1,000 (/tsʰʌn/) or 10,000 (/man/), in Korean, which could be combined to create different numbers resembling words, depending on the location of a

prosodic boundary. Targets were designed not to include diphthongs, but otherwise different types of segments and syllable structures were used.[7]

As many syllables as possible were kept identical at a given position across the targets to minimize the variance unrelated to the factors of interest. The first two syllables /i.man/ and the last syllable /o/ were identical for all 5-syllable targets. For the 7-syllable targets, the first three syllables, /i.pɛk.man/, and the last two syllables, /i.pɛk/, were identical for all targets, with two exceptions (marked with † in Table 1), where /man/ was used as the penultimate syllable. Since the potential target APs had a similar number of syllables, fillers were constructed to have different phrasing structure from them (e.g., 3 + 3, 2 + 2, 1 + 2 + 1, or 2 + 2 + 2) to distract speakers from producing speech with strict rhythmic regularity. Various numbers which did not appear in the targets were used for fillers, together with the numbers used in the targets. Some of the experimental materials were originally designed for an experiment in which speakers' prosodic strategies for disambiguating two alternative phrasings (e.g., 2 + 3 vs. 3 + 2) are examined, reported in Jeon (2011, Chap. 4).

There were three FOCUS conditions: Neutral (read neutrally), PW1-focus (the first PW under narrow corrective focus) and PW2-focus (the second PW under narrow corrective focus). This design makes it possible to compare the same PW spoken in three different ways: neutral, focused (PW1 in PW1-focus, PW2 in PW2-focus) and defocused (PW2 in PW1-focus, PW1 in PW2-focus). For the recording of the Neutral utterances, a list of sentences ('neutral list') were created with the targets and fillers between two carrier phrases (meaning 'the numbers for this time are **[target]** ')[8]:

carrier 1,	**target**	carrier 2
/ipʌn#sutɯilɯn#	**target** (PW1#PW2)#	twɛkɛs*ɯmnita/

'this time' # 'numbers' + TOP #   target (PW1#PW2) # 'be' + ENDER

In the written sentences, commas were placed after the first carrier phrase

(/ipʌn#sutɯlɯn/) and after each PW in the target or filler.

 A 'focus list' of sentences was prepared for the recordings of PW1-focus and PW2-focus. On

the 'focus list', either PW1 or PW2 was underlined and was preceded by a phrase with a 'not

A, B but A, C' construction in parentheses, for example, '([it is] not twenty thousand and

one, a thousand and one), but the numbers this time are twenty thousand and one, ten

thousand and five.' The phonemic transcription and gloss are:

                    carrier 1   **target**   carrier 2

(imanil <u>tsʰʌnili</u> anila,)   /ipʌn sutɯlɯn **imanil, mano**   twɛkɛs*ɯmnita/

(20,001, 1,001 + NM + 'not',) this + time numbers + TOP 20,001, 10,005 'become' +

ENDER

This construction may appear to be a double correction in English (e.g., 'TEN thousand

FIVE' in the example above) but it was intended to have completely different numbers

resembling lexical units to be contrasted in the materials in Korean (/tsʰʌnil/, 1,001 vs.

/mano/, 10,005). That is, the English-type single correction (e.g., 'not twenty thousand and

one (20,001), a thousand and one (1,001), but twenty thousand and one (20,001), a

thousand and FIVE (1,005)) in Korean would not yield the AP structure as desired because

it could lead speakers to produce three APs as target, e.g., (imanil#<u>tsʰʌnili</u>#anila,)#

imanil#<u>tsʰʌn</u>#<u>o</u>, 'not 20,001, 1,001, but 20,001, 1,005' with monosyllabic APs.

The 'neutral list' consisted of 32 sentences with the targets and 64 sentences with the fillers between the carrier phrases. In the 'focus list', there were 64 sentences with the targets and 128 sentences with the fillers. The order of the sentences on each list was randomized for each subject, and two filler sentences in a random order were inserted between two sentences with a target, so that there would always be two filler sentences separating sentences with a target.

The inherent nature of the number sequence in the form of a two-item list might lead to readings deviating from those of ordinary sentences which do not include numbers. However, the use of the number sequence did not elicit any noticeably different prosodic properties from what is reported in literature for ordinary sentences when read neutrally (see Jeon, 2011, Ch. 4). Further, the number sequences are likely to be associated with a phrase-final rising intonation which can be interpreted as signalling phrase-finality, continuation, and the organisation of the successive items (Park, 2012) across all materials, and therefore tight experimental control was achieved.

Table 1. *Experimental materials.* /il/ (1), /i/ (2), /sam/ (3), /sa/ (4), /o/ (5), /juk/ (6), /tsʰil/ (7), /pʰal/ (8), /ku/ (9), /sip/ (10), /pɛk/ (100), / tsʰʌn/ (1,000), and /man/ (10,000).

phrasing	PW2-initial	number	phonemic transcription
2 + 3	w	20,000#10,005	iman#ilmano
		20,000#105	iman#ilpɛko
		20,000#6,005	iman#juktsʰʌno
		20,000#1,005	iman#iltsʰʌno
	s	20,000#30,005	iman#sammano
		20,000#70,005	iman#tsʰilmano
		20,000#3,005	iman#samtsʰʌno

		20,000#7,005	iman#tsʰiltsʰʌno
3 + 2	w	20,001#10,005	imanil#mano
		20,001#105	imanil#pɛko
		20,003#10,005	imansam#mano
		20,007#10,005	imantsʰil#mano
	s	20,006#1,005	imanjuk#tsʰʌno
		20,001#1,005	imanil#tsʰʌno
		20,003#1,005	imansam#tsʰʌno
		20,007#1,005	imantsʰil#tsʰʌno
3 + 4	w	2,000,000#10,200	ipɛkman#ilmanipɛk
		2,000,000#1,002,000	ipɛkman#pɛkmanipɛk
		2,000,000#1,200	ipɛkman#iltsʰʌnipɛk
		2,000,000#1,00,100	ipɛkman#ilsipmanpɛk†
	s	2,000,000#30,200	ipɛkman#sammanipɛk
		2,000,000#70,200	ipɛkman#tsʰilmanipɛk
		2,000,000#3,200	ipɛkman#samtsʰʌnipɛk
		2,000,000#7,200	ipɛkman#tsʰiltsʰʌnipɛk
4 + 3	w	2,000,001#10,200	ipɛkmanil#manipɛk
		2,000,100#10,200	ipɛkmanpɛk#manipɛk
		2,000,003#10,200	ipɛkmansam#manipɛk
		2,000,007#10,200	ipɛkmantsʰil#manipɛk
	s	2,000,001#1,200	ipɛkmanil#tsʰʌnipɛk
		2,000,001#100,100	ipɛkmanil#sipmanpɛk†
		2,000,003#1,200	ipɛkmansam#tsʰʌnipɛk
		2,000,007#1,200	ipɛkmantsʰil#tsʰʌnipɛk

2.2. Experimental procedure

Four native speakers of Seoul Korean (2 females, YH, HKL and two males, KJ, CHJ) aged

between 20 and 22 participated in the experiment. Participants were given a small

payment. All recording was done in a sound-attenuated booth in the Hanyang University in Seoul, using a Tascam HD-P2 recorder and a Shure KSM 44 microphone. The sampling rate was 44.1 kHz.

Participants were given the 'neutral list' first. They had time to become familiar with the materials, and were allowed to practise if they wanted. They read through the 'neutral list' five times for recording. The 'focus list' reading was recorded after each speaker completed the 'neutral list' reading. With the 'focus list', speakers were asked to imagine a situation in which their interlocutor misunderstood the sentence, and they were told they could silently read the phrase in the parentheses if they wanted to (see, e.g. Jun and Lee, 1998 for a similar technique). They were asked to practise until they could read the materials naturally. The 'focus list' was read three times in total, and speakers took a short break after each list reading.

2.3. Data annotation and measurements

For Neutral, 640 utterances including the target (4 targets $\times$ 2 PW2-initial segment types $\times$ 4 phrasings $\times$ 5 repetitions $\times$ 4 speakers) were recorded. Fifteen utterances with hesitation or ambiguous phrasing were discarded, finally leaving 625 utterances for the analysis. Since we aimed to investigate natural intonational variation in speech data, no utterances were discarded by reason of being produced with F_0 contour shapes which are infrequently observed in Korean. In each of PW1-focus and PW2-focus, 378 and 382 utterances out of 384 utterances including the target (4 targets $\times$ 2 PW-initial segment types $\times$ 4 phrasings $\times$ 3 repetitions $\times$ 4 speakers) in each category were analysed.

In order to examine durational variations related to the focus, segment, syllable, and PW boundaries in each utterance were annotated using Praat (Boersma and Weenink, 2010),

following standard criteria suggested in Peterson and Lehiste (1960) and Turk, Nakai and Sugahara (2006). Glottal stops or creaky parts which often appeared at the vowel onset or offset were marked separately but included as part of a vowel. Glottal stops were marked only when there was a clear silent interval, and irregular but continuous pulses were marked as creak. When there was a sequence of consonants with a single closure, e.g., kts^h in /juk.ts^{h}ʌn/ ('6,000'), the closure was halved, and each half was treated as a closure of one consonant. The syllable boundary was defined as suggested in Korean orthography which reflects the morphological structure without regard to possible resyllabification, for instance, CVC.V to CV.CV (e.g., /man.il/ to /ma.nil/, '10,001'), for the ease and consistency in the analysis. Although it was assumed that one PW in the reading materials would form an AP in speech, the targets included sequences of various syllabic compositions, e.g., /manil/ ('10,001') and /mansam/ ('10 003'), and one orthographic syllable, e.g., /man/ ('10,000'), a morphonological unit as an independent noun, could form an AP. Therefore, the orthographic syllable division was used in all data. The duration of each syllable and PW was extracted using a Praat script.

In addition, the first author annotated the F_0 contour shape in the PW together with the type of prosodic boundary at the right edge of each prosodic phrase, following the criteria in Jun (2000). Annotation of the F_0 contour shape was necessary in order to investigate possible variations related to focal structure and also for analysing F_0 span in PW. The PWs realised with a static pitch contour shape (e.g., HH) would need to be separated from those in a PW with a clear F_0 peak (e.g., LHLH).

The prosodic boundary strength before PW1, after PW1 and after PW2 was labelled as PW (with no perceivable prosodic disjuncture), AP (with perceivable prosodic disjuncture and the AP boundary tone) or IP (with significant phrase-final lengthening and an IP boundary

tone). The points of F_0 maximum and minimum in each number unit (i.e., PW) were detected semi-automatically in Praat, their values were extracted using a Praat script, and they were used in calibrating F_0 span within each PW. The F_0 values were measured in semitones (ST) relative to 100 Hz. Outliers and octave jumps were manually corrected. Sample of 160 utterances in total were cross-checked by four native Korean speakers who are trained in prosodic annotation. The between-annotator agreement rate was high at 82% for F_0 contour shapes and at 93% for the prosodic boundary strength. For the F_0 contour shapes, out of the 18% of tokens where there was disagreement, 10% was caused by minor disagreement on the precise turning point of the F_0 (e.g., LH vs. LLH in multisyllabic PWs) which would not affect the result of the study. The rest of the disagreement was on the identification of the initial or final tone between L and H (e.g., LHLH vs. HHLH, LHL vs. LHLH). The cases where there was disagreement were re-examined by the first author, and a decision between the alternatives was made on the basis of the F_0 contour and the perceived pitch. For the prosodic boundary strength, annotators disagreed when the presence of the phrase-final lengthening in the target PW was not clear. The annotation on the tokens where there was disagreement was corrected so that prosodic phrases with clearly perceived final lengthening and a boundary tone would be referred to as the IP.

3. Results

For the statistical analysis, mixed-effect models were fitted to the data with R (R Development Team, 2015) and with the package *lmer4* (Bates, et al., 2014). P-values were corrected for multiple comparisons using the *mcp* function in the *multComp* package in the post-hoc tests (Hothorn et al., 2008). In the comparison of A vs. B, the positive parameter value indicates the relationship A > B and the negative parameter value indicates A < B. The dependent variables were PW duration, syllable duration, the prosodic boundary type (AP vs. IP), the type of IP boundary tone (monotonal vs. bitonal) when present, and F_0 span

in each PW. In the modelling process, the random factors, SPEAKER and ITEM, were always included with random intercepts and the fixed factors are provided in the relevant section. The initial full model was constructed with all relevant fixed factors and the best-fitting model was identified using the log-likelihood χ^2 tests. Only the results of the final models or the contrast tests directly relevant to hypotheses are reported.

3.1. Duration

3.1.1. PW duration

It was hypothesised that the PW duration would show the order focused > neutral > defocused. In order to explore this hypothesis, the initial full model was constructed with the dependent variable PW duration (ms) and fixed factors FOCUS (Neutral, PW1, PW2), PHRASING (2 + 3, 3 + 2, 3 + 4, 4 + 3), LOCATION (PW1, PW2), and PW-INITIAL SEGMENT (strong vs. weak). Since the interactions involving LOCATION (PW1, PW2) were statistically significant (LOCATION × FOCUS, $\chi^2 (2) = 368.2, p < 0.001$; LOCATION × PHRASING, $\chi^2 (3) = 1627.3, p < 0.001$), data were split by LOCATION in order to examine durational variations of each PW. PW-INITIAL SEGMENT was not included in the final models, since its effect was not significant for PW2 ($\chi^2 (1) = 0.48$, ns; there was only one level for PW1).

The effect of FOCUS was significant for PW1 ($\chi^2 (2) = 506.52, p < 0.001$) and PW2 ($\chi^2 (2) = 630.55, p < 0.001$). PHRASING also significantly affected PW duration ($\chi^2 (3) = 102.44, p < 0.001$ for PW1 and $\chi^2 (3) = 84.47, p < 0.001$ for PW2), whilst the FOCUS × PHRASING interaction was not significant ($\chi^2 (6) = 12.45$, ns for PW1 and $\chi^2 (6) = 10.92$, ns for PW2). Table 2 shows that for both PW1 and PW2, the focused PWs were significantly longer than Neutral (see PW1-f vs. Neutral for PW1 and PW2-f vs. Neutral for PW2 in Table

2), whilst the duration of the defocused PW did not show a significant difference from Neutral (PW2-f vs. Neutral for PW1 and PW1-f vs. Neutral for PW2 in Table 2).

Table 2. *Summary of the Tukey contrast test for FOCUS with a dependent variable PW duration (ms). A positive estimate indicates that the first level in the comparison has a higher group mean. n = 1385 for PW1; n = 1385 for PW2. Significance codes *p < 0.05, ** p < 0.01, *** p < 0.001.*

	PW1			PW2		
	Estimate	SE	z	Estimate	SE	Z
PW1-f vs. Neutral	106.47	4.62	23.05***	-5.06	5.15	-0.98
PW2-f vs. Neutral	-0.33	4.60	-0.07	133.26	5.13	25.96***
PW2-f vs. PW1-f	-106.80	5.14	-20.77***	138.32	5.73	24.12***

3.1.2. Syllable duration

The hypothesised relationship for syllable duration was focused > neutral > defocused, and it was expected that lengthening triggered by focus would be more pronounced at the focused constituent edges. Figure 1 demonstrates that all syllables in the focused PW tended to be longer than those in its Neutral or defocused counterpart. As expected, the magnitude of lengthening under focus was more pronounced at phrase edges than in phrase-medial syllables. Compared to Neutral, PW-initial syllables were lengthened under focus by 27 % on average (mean = 46.53 ms, range 26.09-62.12 ms), PW-medial syllables by 14.3 % on average (mean = 23.39 ms, range 15.01-35.53ms), and PW-final syllables by 25 % on average (mean = 46.88 ms, range 32.70-62.67 ms, also see larger absolute values of the estimates in Table 4 for the PW-initial syllable). However, the duration of syllables in the defocused PW tends to overlap with that in Neutral in Figure 1.

Linear mixed-effect models were fitted to the syllable duration (ms) as a dependent variable. The initial full models were constructed for 5-syllable targets and 7-syllable targets respectively with fixed factors, FOCUS, PHRASING, PW-INITIAL SEGMENT, and SYLLABLE

POSITION (1-5 in 5-syllable targets, 1-7 in 7-syllable targets). In the modelling process, the three-way interaction effect SYLLABLE POSITION × FOCUS × PHRASING was statistically significant (5-syllable targets, $\chi^2 (8) = 113.63, p < 0.001$; 7-syllable targets, $\chi^2 (12) = 132.5, p < 0.001$). Further models were constructed for the syllable duration (ms) in each SYLLABLE POSITION to explore the FOCUS × PHRASING interaction. Some syllables in the same position in the target occupy different position relative to the PW boundaries depending on the phrasing (e.g., the second syllable would be PW1-final in 2 + 3 phrasing but PW1-medial in 3 + 2 phrasing) and, therefore, it was necessary to examine the potential interaction between FOCUS and PHRASING in each SYLLABLE POSITION. Since there was a significant FOCUS × PHRASING effect in the majority of cases (see Table 3 for the effects of fixed factors), the pairwise comparisons between the focus conditions were conducted within each phrasing.
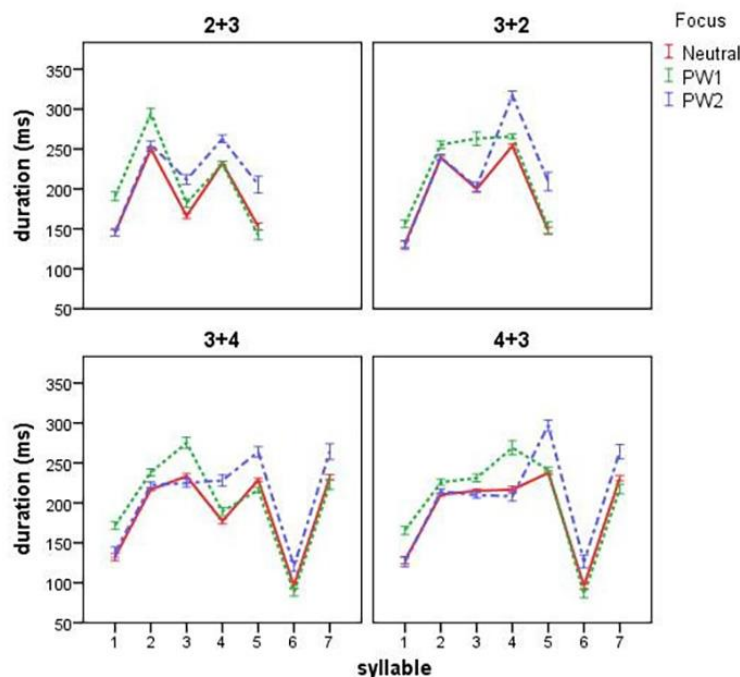

Figure 1. *Mean duration of each syllable (ms) for each PHRASING averaged over all speakers.* Error bars indicate one standard error, *n = 8328. Neutral: solid line, PW1-focus: dotted line, PW2-focus: dotdash line.*

The pairwise comparisons in each phrasing showing statistically significant differences are the following (see Table 4): for 2 + 3, syllables 1 (PW1-initial) and 2 (PW1-final) were significantly longer in PW1-focus than PW2-focus or Neutral, whereas there was no statistically significant difference between Neutral and PW2-focus. Syllable 3 (PW2-initial) was significantly longer in the order PW2-focus > PW1-focus > Neutral. Syllables 4 and 5 were significantly longer in PW2-focus than Neutral and PW1-focus, whereas there was no significant difference between Neutral and PW1-focus.

For 3 + 2, syllables 1 (PW1-initial), 2 (PW1-medial) and 3 (PW1-final) were longer in PW1-focus than in Neutral or PW2-focus, whereas there was no significant difference between Neutral and PW2-focus. Syllables 4 (PW2-initial) and 5 (PW2-final) were longer in PW2-focus than in Neutral or PW1-focus. The duration of syllables 4 and 5 did not show a significant difference between Neutral and PW1-focus.

For 3 + 4, syllables 1 (PW1-initial), 2 (PW1-medial), and 3 (PW1-final) were significantly longer in PW1-focus than in Neutral or PW2-focus, whereas there was no significant difference between Neutral and PW2-focus. Syllables 4 (PW2-initial), 6 (PW2-medial), and 7 (PW2-final) were significantly longer in PW2-focus than in Neutral or PW1-focus, and no statistically significant difference was found between Neutral and PW1-focus. Syllable 5 (PW2-medial) showed the order PW2-focus > Neutral > PW1-focus.

For 4 + 3, syllables 1 (PW1-initial), 2, (PW1-medial), 3 (PW1-medial), and 4 (PW1-final) were significantly longer in PW1-focus than in Neutral or PW2-focus, whereas the difference between Neutral and PW2-focus did not reach significance for all comparisons. Syllables 5 (PW2-initial) was also significantly lengthened in PW2-focus than in Neutral or PW1-focus

and the difference between PW2-focus and neutral did not reach significance. Syllables 6 (PW2-medial), and 7 (PW2-final) showed the order PW2-focus > Neutral > PW1-focus.

Table 3. *Summary of the main effects FOCUS and PHRASING and the FOCUS × PHRASING interaction in the mixed-effects model with syllable duration (ms) as a dependent variable for each syllable in 5-syllable targets and in 7-syllable targets. Significance codes *p < 0.05, ** p < 0.01, *** p < 0.001.*

Syll no.	5-syllable target		7-syllable target	
	χ^2	df	χ^2	df
FOCUS				
1	132.89***	2	157.37***	2
2	120.11***		112.95***	
3	118.02***		148.8***	
4	226.15***		65.524***	
5	219.60***		243.45***	
6			166.71***	
7			153.68***	
PHRASING				
1	21.59***	1	6.81**	1
2	5.34*		17.43***	
3	2.75		10.85***	
4	18.30***		2.60	
5	0.16		6.46*	
6			0	
7			0.67	
FOCUS × PHRASING				
1	10.40*	2	1.14	2
2	28.52***		3.63	
3	145.26***		30.62***	
4	31.51***		140.87***	
5	3.99		16.81***	
6			2.04	
7			1.03	

Table 4. *The output of a Tukey contrast test for syllable duration. A positive estimate indicates that the first level in the comparison has a higher group mean. 5-syllable targets, 2 + 3, n = 346; 3 + 2, n = 341; 7-syllable targets, 3 + 4, n = 350; 4 + 3, n = 349. Significance codes *p < 0.05, ** p < 0.01, *** p < 0.001.*

	5-syllable targets				7-syllable targets			
		Estimate	SE	z		Estimate	SE	z
1	2 + 3.PW1-f vs. 2 + 3.N	46.03	4.49	10.25***	3 + 4.PW1-f vs. 3 + 4.N	38.92	4.28	9.10***
	2 + 3.PW2-f vs. 2 + 3.N	1.46	4.51	0.32	3 + 4.PW2-f vs. 3 + 4.N	5.10	4.28	1.19
	2 + 3.PW2-f vs. 2 + 3.PW1-f	-44.57	5.00	-8.91***	3 + 4.PW2-f vs. 3 + 4.PW1-f	-33.82	4.77	-7.09***
	3 + 2.PW1-f vs. 3 + 2.N	25.94	4.56	5.69***	4 + 3.PW1-f vs. 4 + 3.N	37.14	4.29	8.67***
	3 + 2.PW2-f vs. 3 + 2.N	-1.47	4.50	-0.33	4 + 3.PW2-f vs. 4 + 3.N	-1.34	4.29	-0.31
	3 + 2.PW2-f vs. 3 + 2.PW1-f	-27.41	5.05	-5.43***	4 + 3.PW2-f vs. 4 + 3.PW1-f	-38.48	4.80	-8.02***
2	2 + 3.PW1-f vs. 2 + 3.N	43.16	3.72	11.61***	3 + 4.PW1-f vs. 3 + 4.N	21.08	2.31	9.13***
	2 + 3.PW2-f vs. 2 + 3.N	4.56	3.73	1.22	3 + 4.PW2-f vs. 3 + 4.N	5.37	2.31	2.33
	2 + 3.PW2-f vs. 2 + 3.PW1-f	-38.60	4.14	-9.32***	3 + 4.PW2-f vs. 3 + 4.PW1-f	-15.71	2.58	-6.10***
	3 + 2.PW1-f vs. 3 + 2.N	15.12	3.78	4.00***	4 + 3.PW1-f vs. 4 + 3.N	14.99	2.31	6.48***
	3 + 2.PW2-f vs. 3 + 2.N	-0.81	3.73	-0.22	4 + 3.PW2-f vs. 4 + 3.N	4.27	2.31	1.85
	3 + 2.PW2-f vs. 3 + 2.PW1-f	-15.94	4.18	-3.82**	4 + 3.PW2-f vs. 4 + 3.PW1-f	-10.71	2.59	-4.14***
3	2 + 3.PW1-f vs. 2 + 3.N	16.56	4.47	3.70**	3 + 4.PW1-f vs. 3 + 4.N	41.88	3.71	11.28***
	2 + 3.PW2-f vs. 2 + 3.N	45.60	4.49	10.16***	3 + 4.PW2-f vs. 3 + 4.N	-8.27	3.71	-2.23
	2 + 3.PW2-f vs. 2 + 3.PW1-f	29.04	4.98	5.83***	3 + 4.PW2-f vs. 3 + 4.PW1-f	-50.15	4.14	-12.11***
	3 + 2.PW1-f vs. 3 + 2.N	62.94	4.55	13.85***	4 + 3.PW1-f vs. 4 + 3.N	15.75	3.72	4.23***
	3 + 2.PW2-f vs. 3 + 2.N	1.86	4.48	0.42	4 + 3.PW2-f vs. 4 + 3.N	-5.56	3.72	-1.49
	3 + 2.PW2-f vs. 3 + 2.PW1-f	-61.08	5.02	-12.16***	4 + 3.PW2-f vs. 4 + 3.PW1-f	-21.31	4.16	-5.12***
4	2 + 3.PW1-f vs. 2 + 3.N	-1.435	4.022	-0.357	3 + 4.PW1-f vs. 3 + 4.N	10.59	5.09	2.08

	2+3.PW2-f vs. 2+3.N	29.70	4.04	7.36***	3+4.PW2-f vs. 3+4.N	50.71	5.09	9.97***
	2+3.PW2-f vs. 2+3.PW1-f	31.14	4.48	6.95***	3+4.PW2-f vs. 3+4.PW1-f	40.13	5.68	7.07***
	3+2.PW1-f vs. 3+2.N	11.38	4.09	2.79	4+3.PW1-f vs. 4+3.N	52.37	5.10	10.27***
	3+2.PW2-f vs. 3+2.N	62.10	4.03	15.40***	4+3.PW2-f vs. 4+3.N	-7.56	5.10	-1.48
	3+2.PW2-f vs. 3+2.PW1-f	50.72	4.52	11.23***	4+3.PW2-f vs. 4+3.PW1-f	-59.93	5.71	-10.51***
5	2+3.PW1-f vs. 2+3.N	-12.19	5.36	-2.28	3+4.PW1-f vs. 3+4.N	-12.31	4.33	-2.84*
	2+3.PW2-f vs. 2+3.N	50.05	5.38	9.31***	3+4.PW2-f vs. 3+4.N	35.39	4.33	8.18***
	2+3.PW2-f vs. 2+3.PW1-f	62.25	5.97	10.42***	3+4.PW2-f vs. 3+4.PW1-f	47.69	4.83	9.88***
	3+2.PW1-f vs. 3+2.N	2.50	5.45	0.46	4+3.PW1-f vs. 4+3.N	3.80	4.34	0.88
	3+2.PW2-f vs. 3+2.N	59.50	5.37	11.07***	4+3.PW2-f vs. 4+3.N	59.49	4.34	13.72***
	3+2.PW2-f vs. 3+2.PW1-f	57.00	6.02	9.47***	4+3.PW2-f vs. 4+3.PW1-f	55.70	4.85	11.48***
6					3+4.PW1-f vs. 3+4.N	-7.92	3.54	-2.24
					3+4.PW2-f vs. 3+4.N	24.71	3.54	6.99***
					3+4.PW2-f vs. 3+4.PW1-f	32.62	3.94	8.27***
					4+3.PW1-f vs. 4+3.N	-10.34	3.54	-2.92*
					4+3.PW2-f vs. 4+3.N	30.03	3.54	8.47***
					4+3.PW2-f vs. 4+3.PW1-f	40.37	3.96	10.18***
7					3+4.PW1-f vs. 3+4.N	-7.84	4.44	-1.77
					3+4.PW2-f vs. 3+4.N	32.91	4.44	7.42***
					3+4.PW2-f vs. 3+4.PW1-f	40.75	4.95	8.23***
					4+3.PW1-f vs. 4+3.N	-13.80	4.45	-3.10*
					4+3.PW2-f vs. 4+3.N	32.74	4.45	7.36***

4 + 3.PW2-f			
vs. 4 + 3.PW1-f	46.54	4.98	9.35***

3.2. Focus and the prosodic boundary type

It was hypothesised that the focused constituent begins a new prosodic phrase. As hypothesized, the PWs under narrow focus always initiated a new prosodic phrase, either the AP or the IP. Figure 2 shows that there was no PW-sized disjuncture (meaning there was always a perceivable prosodic disjuncture) before the focused constituent, i.e., Pre-PW1 for PW1 under focus and Post-PW1 for PW2 under focus. The hypothesis that focused PWs would be associated with a larger prosodic boundary than neutrally spoken or defocused PWs was supported in some contexts. Figure 2 shows that the most frequent prosodic boundary type was IP between the carrier phrase and the PW1 (Pre-PW1) and also after the PW2 (Post-PW2) in all FOCUS conditions. For Post-PW1, although AP was more common for Neutral or PW2-focus, when PW1 was under focus, the most frequent type boundary type was IP.

The probability of an IP boundary occurrence as opposed to an AP boundary was modelled by mixed effects logistic regression (e.g., Pinheiro & Bates, 2000; Baayen, Davidson, & Bates, 2008). The initial full model included fixed factors, FOCUS, PHRASING, LOCATION, and PW-INITIAL SEGMENT. All two-way interaction effects were statistically significant (FOCUS × PHRASING, $\chi^2 (6) = 12.76, p < 0.05$; PHRASING × LOCATION, $\chi^2 (6) = 88.05, p < 0.001$; FOCUS × LOCATION, $\chi^2 (4) = 219.99$), although the FOCUS × PHRASING × LOCATION interaction was not ($\chi^2 (12) = 17.44,$ ns). Due to the interdependence between the three fixed factors, data were split by LOCATION (Pre-PW1, Post-PW1, and Post-PW2) in order to explore how the PW in different positions in the utterance were affected by FOCUS. There were significant FOCUS and PHRASING effects for Pre-PW1 (FOCUS, $\chi^2 (2) = 31.49, p < 0.001$; PHRASING, $\chi^2 (2) = 15.71, p < 0.01$), Post-

PW1 (FOCUS, $\chi^2 (2) = 315.17, p < 0.001$; PHRASING $\chi^2 (3) = 37.66, p < 0.001$) and Post-PW2 (FOCUS, $\chi^2 (2) = 25.48, p < 0.001$; PHRASING $\chi^2 (3) = 10.51, p < 0.05$). For Post-PW2, the FOCUS × PHRASING interaction was also significant ($\chi^2 (6) = 21.33, p < 0.01$).
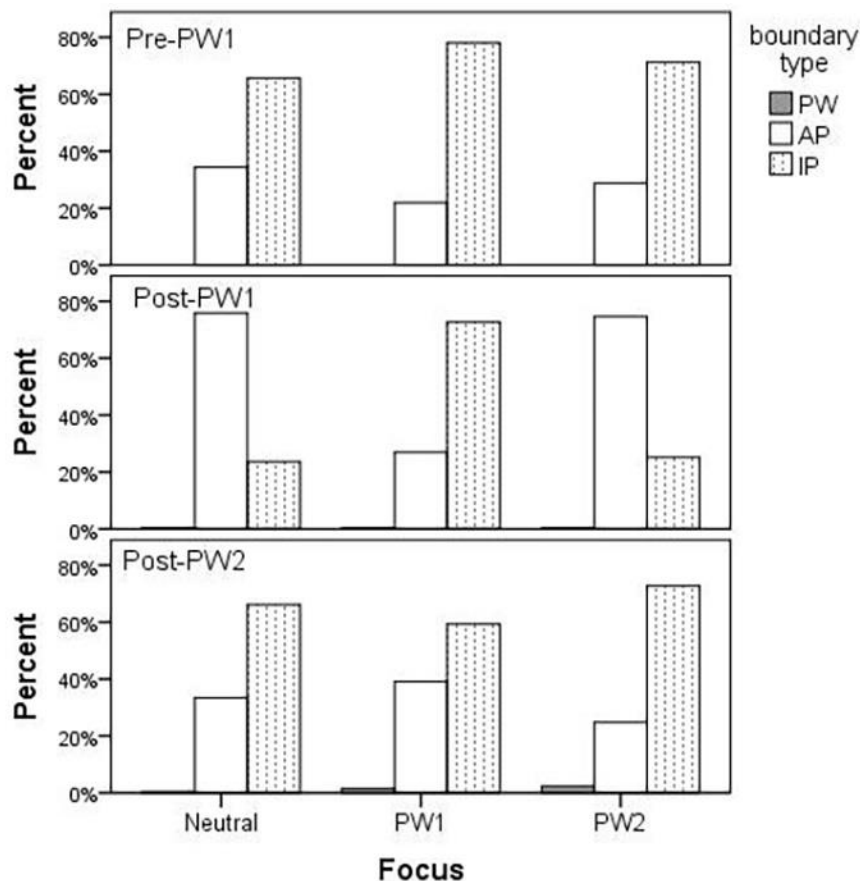

Figure 2. *The frequency (percentage) of each boundary type between the carrier phrase and PW1 (Pre-PW1), after PW1 (Post-PW1) and after PW2 (Post-PW2). Neutral, n = 626; PW1-focus, n = 378; PW2-focus, n = 382.*

The result of the contrast test (Table 5) demonstrates the effect of FOCUS for Pre-PW1 and Post-PW1. For both Pre-PW1 and Post-PW1, more IP boundaries were likely to occur in PW1-focus than in Neutral or PW2-focus. PHRASING affected the likelihood of the IP boundary presence in some cases. For Pre-PW1, the target in 2 + 3 phrasing was more likely to be preceded by the IP boundary than other phrasing types. For Post-PW1, the IP

boundary was more likely observed when PW1 had more syllables than PW2. For Post-PW2, a statistically significant difference in the contrast test was observed in only a few comparisons; for 3 + 2, the IP boundary was more likely to be present in PW2-focus than in Neutral or PW1-focus. For 3 + 4, more IP boundaries were present for Neutral than for PW1-focus.

Table 5. *The output of a Tukey contrast test in a generalised linear mixed effects model of the IP boundary presence, main effects of FOCUS and PHRASING for Pre-PW1, Post-PW1, and Post-PW2. In the Tukey test result, a positive estimate indicates that the first level in the comparison has a higher group mean. Significance codes *p < 0.05, ** p < 0.01, *** p < 0.001.*

Pre-PW1	Estimate	SE	z	Post-PW1	Estimate	SE	z
PW1-f vs. N	1.12	0.21	5.41***	PW1-f vs. N	2.56	0.17	14.89***
PW2-f vs. N	0.47	0.19	2.44	PW2-f vs. N	0.10	0.16	0.63
PW2-f vs. PW1-f	-0.65	0.23	-2.87*	PW2-f vs. PW1-f	-2.46	0.19	-13.11***
3 + 2 vs. 2 + 3	-0.76	0.24	-3.20**	3 + 2 vs. 2 + 3	0.72	0.23	3.11*
3 + 4 vs. 2 + 3	-0.94	0.23	-4.03***	3 + 4 vs. 2 + 3	-0.23	0.24	-0.98
4 + 3 vs. 2 + 3	-0.67	0.24	-2.84*	4 + 3 vs. 2 + 3	1.48	0.23	6.42***
3 + 4 vs. 3 + 2	-0.19	0.22	-0.84	3 + 4 vs. 3 + 2	-0.95	0.24	-4.05***
4 + 3 vs. 3 + 2	0.09	0.23	0.39	4 + 3 vs. 3 + 2	0.76	0.22	3.44**
4 + 3 vs. 3 + 4	0.27	0.22	1.24	4 + 3 vs. 3 + 4	1.72	0.24	7.29***
Post-PW2							
2 + 3.PW1-f vs. 2 + 3.N	-0.14	0.37	-0.38	3 + 4.PW1-f vs. 3 + 4.N	-1.38	0.35	-3.89**
2 + 3.PW2-f vs. 2 + 3.N	0.63	0.39	1.59	3 + 4.PW2-f vs. 3 + 4.N	-0.41	0.37	-1.09

2 + 3.PW2-f vs. 2 + 3.PW1-f	0.77	0.43	1.78	3 + 4.PW2-f vs. 3 + 4.PW1-f	0.97	0.39	2.50
3 + 2.PW1-f vs. 3 + 2.N	0.03	0.33	0.08	4 + 3.PW1-f vs. 4 + 3.N	-0.46	0.34	-1.33
3 + 2.PW2-f vs. 3 + 2.N	1.83	0.41	4.41**	4 + 3.PW2-f vs. 4 + 3.N	0.27	0.36	0.75
3 + 2.PW2-f vs. 3 + 2.PW1-f	1.80	0.44	4.02**	4 + 3.PW2-f vs. 4 + 3.PW1-f	0.73	0.39	1.85

3.3. F_0

3.3.1. F_0 contour shapes over the PW

It was hypothesised that the F_0 contour shapes in the PW would be affected by focal structure and speakers' F_0 adjustment would interact with the segmentally-induced F_0 variation. First, the analysis of the annotated F_0 contour shapes revealed that when the PW was realised as an AP, the most common F_0 contour shapes were LH, HH, HLH, LHLH and HLHL. These five types accounted for 79.25% of the total APs. Further details are not provided since the variations in the F_0 contour shapes within the AP were not related to FOCUS in contrast to the hypothesis, although some cases with static pitch contours (LL and HH) are presented below. The assumption that one PW in the reading materials would be realised as one AP was met in general; there were only 23 cases of the PW boundary percept (i.e., no marked prosodic disjuncture), and 9 cases out of 23 were observed in Post-PW2 when PW2 was under focus indicating post-focus dephrasing. The AP-final H frequently appeared as expected (84.33% of total APs) and the AP-initial tone tended to be determined by the type (weak/strong) of its initial segment. However, there were exceptions; Table 6 shows that not all PW2s beginning with a strong segment were realised

with a phrase-initial H and there were cases of association of H with a PW-initial weak segment. In Neutral, 2 % of PW1s (which all begin with a weak segment) and 23% of PW2s with a weak initial segment began with H. This is likely to be an artefact related to the composition of the experimental materials. The PW2 of 2 + 3 and 3 + 4 often begins with /il/ ('one'); some Korean speakers raise AP-initial pitch for /il/ ('one') in order to distinguish the two numbers /il/ ('one') and /i/ ('two') which could sound similar to each other (Jun and Cha, 2011, 2015).

Table 6. *Frequency of PW-initial H (percentage) in PW1 and PW2 for each the type of the PW2-initial segment for Neutral, PW1-focus and PW2-focus.*

	Weak		Strong	
PW1	Neutral (n = 626)	2%		
	PW1-focus (n = 378)	8%		
	PW2-focus (n = 382)	5%		
PW2	Neutral (n = 314)	23%	Neutral (n = 312)	92%
	PW1-focus (n = 189)	20%	PW1-focus (n = 189)	88%
	PW2-focus (n = 190)	28%	PW2-focus (n = 192)	97%

Second, a final rise (H% and LH%) occurred frequently when the PW was aligned to the right edge of the IP. The frequency statistics in Table 7 suggests that in Neutral, H% was the most common IP boundary tone and the frequency of LH% increased when the PW was focused. The probability of the bitonal IP boundary tone (LH%, HL%) being present was modelled using mixed effects logistic regression with the fixed factors FOCUS, LOCATION and PHRASING. Since the FOCUS × PHRASING × LOCATION effect was statistically significant ($\chi^2 (6) = 21.28, p < 0.01$), the data were split by LOCATION (PW1, PW2). The effect of FOCUS was significant for both PW1 ($\chi^2 (2) = 30.67, p < 0.001$) and PW2 ($\chi^2 (2) = 222.31, p < 0.001$). However, the PHRASING effect ($\chi^2 (3) = 25.69, p < 0.001$) and the FOCUS × PHRASING interaction ($\chi^2 (6) = 37.13, p < 0.001$) were significant only for

PW2. A Tukey contrast test revealed that for PW1, the bitonal IP boundary tones were more likely to occur for PW1-focus than in PW2-focus, whilst there was no significant difference between Neutral and either of PW1-focus or PW2-focus (Table 8). For PW2, although no statistically significant differences between the FOCUS levels were revealed for 4 + 3, the bitonal IP boundary tones were more likely to occur when PW2 was under focus in other phrasings in comparison to Neutral or PW1-focus.

Table 7. *Frequency of IP boundary tones (percentage) in Neutral, PW1-focus and PW2-focus.*

	PW1			PW2		
	Neutral (n = 147)	PW1-focus (n = 274)	PW2-focus (n = 96)	Neutral (n = 410)	PW1-focus (n = 223)	PW2-focus (n = 277)
H%	96	56	67	94	78	47
L%		18	17	1	1	2
LH%	2	23	2	5	21	52
HL%	2	4	15			

Further analyses of the APs produced with the static pitch contour such as LL or HH were carried out. There were no APs with LL in the data, but the APs produced with HH in Neutral could be identified (23 utterances from YH, 1 utterance from HKL, 24 utterances from KJ, and 20 utterances from CHJ). The data showed that some speakers (YH and HKL, e.g., Figure 3) consistently produced LH% on the final syllable of the focused PW in all cases whereas the two male speakers preferred raising overall pitch for the HH (KJ and CHJ, e.g., Figure 4). The statistical analysis was not carried out since there were not sufficient data points.

Table 8. *The output of the post-hoc Tukey contrast test for the IP boundary tone type (monotonal vs. bitonal). A positive estimate indicates that the first level in the comparison has a higher group mean. Significance codes *p < 0.05, ** p < 0.01, *** p < 0.001.*

| PW1	Estimate	SE	z	PW2	Estimate	SE	z
PW1-f vs. N	5.95	17.18	0.35	2+3.PW1-f vs. 2+3.N	-0.35	0.45	-0.77
PW2-f vs. N	7.70	17.19	0.45	2+3.PW2-f vs. 2+3.N	2.69	0.40	6.75***
PW2-f vs. PW1-f	1.75	0.24	7.23***	2+3.PW2-f vs. 2+3.PW1-f	3.04	0.47	6.43***
				3+2.PW1-f vs. 3+2.N	3.19	0.68	4.70***
				3+2.PW2-f vs. 3+2.N	4.55	0.67	6.82***
				3+2.PW2-f vs. 3+2.PW1-f	1.35	0.42	3.22*
				3+4.PW1-f vs. 3+4.N	3.18	1.08	2.96
				3+4.PW2-f vs. 3+4.N	4.25	1.05	4.06**
				3+4.PW2-f vs. 3+4.PW1-f	1.07	0.49	2.18
				4+3.PW1-f vs. 4+3.N	18.48	209.03	0.09
				4+3.PW2-f vs. 4+3.N	20.04	209.03	0.10
				4+3.PW2-f vs. 4+3.PW1-f	1.56	0.49	3.18
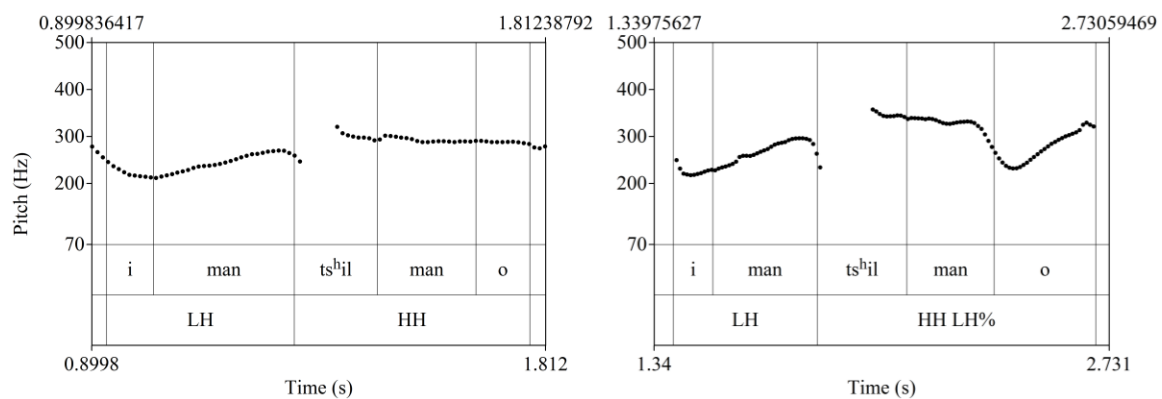
Figure 3. *Neutral (left) and PW2-focus (right), produced by YH (female).*
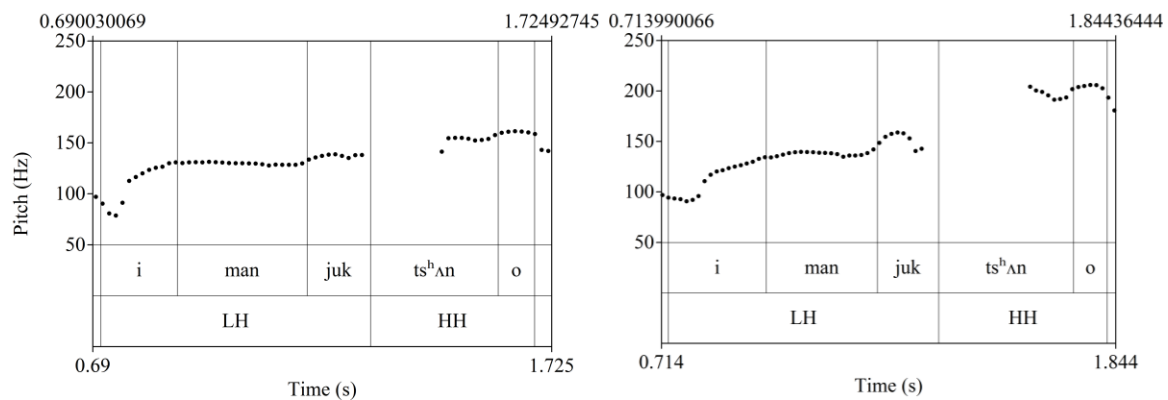

Figure 4. *Neutral (left) and PW2-focus (right), produced by KJ (male).*

3.3.2. F_0 span expansion under narrow focus

In order to examine the hypothesis that the size of F_0 span in the PW would show the order focused > neutral > defocused, data were split into two subsets prior to the statistical modelling: PWs with relatively static pitch (i.e., LL and HH) and the rest which involve a dynamic F_0 movement in the PW (e.g., LH, HL, LHL, etc.). Models are not reported for the static PW, since none of the factors showed a statistically significant effect. For dynamic F_0 contours, the initial full mixed effects model was constructed with F_0 span (ST) as the dependent variable and the fixed factors, FOCUS, LOCATION, and PW-INITIAL SEGMENT. PHRASING, which was not the factor of interest here, was excluded. Due to the significant interaction effect of FOCUS × LOCATION × PW-INITIAL SEGMENT ($\chi^2 (2) = 6.14, p < 0.05$), the data were split into subsets by LOCATION (PW1, PW2) for further modelling with FOCUS and PW-INITIAL SEGMENT as fixed factors.

For PW1, the only statistically significant effect was from FOCUS ($\chi^2 (2) = 396.95, p < 0.001$). The F_0 span showed the order PW1-focus > PW2-focus > Neutral (Table 9). For PW2, since the effect of FOCUS × PW-INITIAL SEGMENT was significant ($\chi^2 (2) = 13.44, p < 0.01$), the data were further split into PWs with a weak onset consonant and those with a strong onset consonant. The FOCUS effect was statistically significant for PW2s beginning with a strong consonant ($\chi^2 (2) = 160.49, p < 0.001$) and also for PW2s with a weak consonant ($\chi^2 (2) = 157.05, p < 0.001$). For PW2s beginning with a strong segment, the contrast test showed that the focused PW2 tended to have larger F_0 span under focus in comparison to PW1-focus or Neutral, whilst F_0 span in defocused PW2s (PW1-focus) was not significantly reduced compared to Neutral. For PW2s beginning with a weak segment, PW2 had the largest F_0 span under focus and shows the order PW2-focus > PW1-focus > Neutral (see Table 9).

Table 9. *The output of a Tukey contrast test in a linear mixed effects model for PW1s, for PW2s beginning with a strong consonant and for PW2s beginning with a weak consonant. All PWs had a dynamic F0 contour shape. A positive estimate indicates that the first level in the comparison has a higher group mean. Significance codes *p < 0.05, ** p < 0.01, *** p < 0.001.*

	Estimate	SE	Z
PW1			
PW1-f vs. N	0.21	0.01	21.59***
PW2-f vs. N	0.08	0.01	8.58***
PW2-f vs. PW1-f	-0.13	0.01	-11.67***
PW2, strong segment			
PW1-f vs. N	0	0.02	-0.25
PW2-f vs. N	0.20	0.02	12.45**
PW2-f vs. PW1-f	0.20	0.02	11.44**
PW2, weak segment			
PW1-f vs. N	0.063	0.01	4.41***
PW2-f vs. N	0.19	0.01	13.36***
PW2-f vs. PW1-F	0.13	0.02	7.89***

4. Discussion

4.1. Prosodic marking of narrow focus in Seoul Korean

Overall, the results show that Korean speakers actively adjust the prosodic organisation of the utterance in response to its focal structure. Although an exact analogue of the culminative pitch accent in West Germanic languages is not apparent in Korean due to the lack of lexical stress, similar prominence-lending acoustic properties are employed and the focus-marking phonetic events are concentrated at phrase boundaries. The speakers' focus marking strategies are in line with general cross-linguistic trends showing an increase in the articulatory effort, e.g., lengthening, more pitch movements, pitch raising, and wider pitch excursions, in the element under narrow focus (see Gussenhoven, 2004).

In the experiment, speakers were under pressure to convey emphasis in the reading materials *verbatim* and they were not allowed to use any of the alternative non-phonetic strategies such as changing the word order. The reading materials were sentences of two PWs which were numbers between carrier phrases, and speakers produced one or other, or neither, of the PWs under corrective focus. The type of PW-initial segment in the target included both weak segments and strong segments.

The hypothesis that the focused constituent and syllables therein would be lengthened in comparison to the neutrally spoken or defocused constituent was supported, since the durational marking of the focused constituent was robust in the data. Duration of the target PWs was affected by FOCUS; the PW-initial and PW-final syllables were lengthened more (on average 27 % and 25%, respectively, compared to Neutral) than phrase-medial syllables (14.3%), although the lengthening could spread over the focused constituent. Unlike the West Germanic languages in which a pitch-accented lexically-stressed syllable undergoes the greatest amount of lengthening under focus (e.g., Cambier-Langeveld and Turk, 1999), Korean speakers lengthen syllables at the phrase edges to the greatest magnitude. Although both edges of the prosodic phrase under focus are subject to lengthening in Korean, its motivation seems to be different between the left and the right edges. This study did not examine segmental duration, but previous studies reveal that the lengthening of the phrase-initial syllable is partly due to articulatory strengthening (Jun and Lee, 1998; Cho and Keating, 2001; Cho, Lee & Kim, 2011). On the other hand, the lengthening at the right edge is probably caused by speakers' attempt to produce the IP boundary tones, as discussed further below.

As hypothesised, the focal structure affects the formation of prosodic phrases in that the focused constituent always initiated a new AP or IP. The interpretation that the boundary

strength varies between the AP and the IP may appear to be inconsistent but it is justifiable. Following the K-ToBI criteria (Jun, 2000), at the right edge of the focused constituent, there may be a simple rise or fall with no perceived final lengthening (AP) or an IP-type boundary tone with significant final lengthening (IP); or there may not be a prosodic disjuncture. What is of importance is that the focused constituent is preceded by prosodic disjuncture.

The focused constituent was more likely to form an IP than an AP and it was associated with more pitch movements near the phrase boundary in comparison to its neutrally spoken or defocused counterpart. In the present experiment, in which speakers' intonational choice was not constrained, they seemed to enjoy the freedom to manipulate the pitch movement at the right edge of the prosodic phrase by employing monotonal (H%, L%) or bitonal (HL%, LH%) IP boundary tones. For example, the constituent which carried LHLH in the AP in Neutral was realized with a bitonal boundary IP tone LH% (LH LH%) under narrow focus (e.g., Figure 5). The bitonal LH% was used by two speakers when the Neutral PW2 had a high-static pitch (HH) (e.g., Figure 3), whilst the other two speakers seem to raise the overall level of HH (e.g. Figure 4). When an IP was formed in the utterance, the most frequent boundary tone was a rise: 91% of the IP boundary tones were H% or LH% on average, and specifically for PW2 99% of the IP boundary tones were either H% or LH% regardless of the focus condition. This would be partly due to the nature of the reading materials which included a two-item list, but speakers' choice of IP boundary tone was affected by FOCUS and PHRASING. In general, the IP boundary was commonly observed between the first carrier phrase and PW1 (Pre-PW1) and then between PW2 and the final carrier phrase (Post-PW2), in, for example, '([it is] not twenty thousand and one, a thousand and one), but the numbers this time are twenty thousand and one, ten thousand and five.' The phonemic transcription and gloss are:

_{IP}[carrier 1]	_{IP}[**target (PW1, PW2)**]	_{IP}[carrier 2]
ipʌn sutɯlɯn	**imanil, mano**	twɛkɛs*ɯmnita
'this time' 'numbers' + TOP	*20,001, 10,005*	*'become' + ENDER*

The right edge of the PW2 corresponded to the end of the target phrase and speakers seemed to have produced the IP boundaries to separate the target phrase from the surrounding carrier phrases.
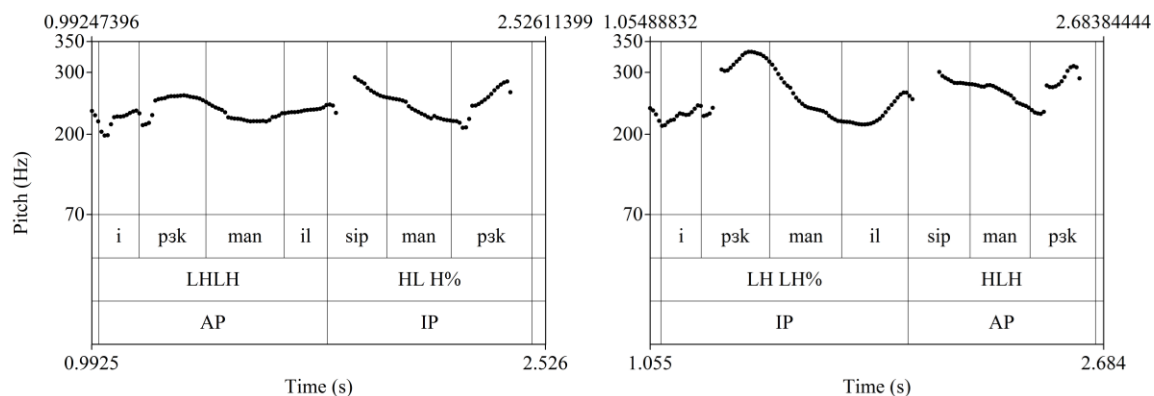

Figure 5. *Speaker HKL (female), Neutral (left) and PW1-focus (right).*

PW1 and PW2 were affected by the FOCUS in a different way. PW1 was frequently realised as an AP when neutrally spoken or defocused (i.e., when PW2 was focused), but focused PW1 was more likely to be realised as an IP. In addition, the focus on PW1 seemed to affect the magnitude of its preceding prosodic boundary together with PHRASING; the first carrier phrase was more likely to form the IP when PW1 was focused and also when the target was 2 + 3. The FOCUS effect was reliable for Pre-PW1 (right edge of the first carrier phrase) and Post-PW1 (right edge of PW1). In particular, PW1 under focus was more likely to have a bitonal IP boundary tone (LH% or HL%) than a monotonal boundary tone (H%, L%) regardless of PHRASING. The same tendency was observed at the right edge of PW2. However, the FOCUS effect was less clear, and the statistically significant FOCUS ×

PHRASING interaction suggests that the prosodic boundary strength after PW2 was more likely to be affected by other factors such as the number of syllables in the prosodic phrase (see Section 3.2).

The effect of narrow focus on prosodic structuring in Korean shares similarities to that in French (e.g., Féry, 2001) and Japanese (e.g., Beckman and Pirrehumbert, 1986; Venditti, Maekawa and Beckman, 2008). One difference is that the focused constituent is aligned to the right in the prosodic phrase in French (Féry, 2013) but left in Japanese and Korean. In addition, Japanese speakers use a salient rise (H) or a rise-fall (LHL) pitch movement at the final syllable of the focused word (Venditti, Maekawa and Beckman, 2008, sect. 3.4.2). These seem similar to Korean speakers' use of the IP boundary tones (e.g., H%, L%, LH%, HL%) observed in the present study.

The occurrences of IP boundary tones in relation to focal structure may not be surprising in that in Korean, complex IP boundary tones are commonly used to deliver pragmatic meaning (Jun, 2000; Park, 2012). Although K-ToBI suggests nine boundary tones (i.e., L%, H%, LH%, HL%, LHL%, HLH%, LHLH%, HLHL% and LHLHL%, see Jun, 2000), even more complex pitch movement such as HLHLHLHLHL% is observed in spontaneous speech (Park, 2012). Yet there is little research investigating the complex pitch movement which seems to be accompanied by substantial phrase-final lengthening, and the use of the IP boundary tones in relation to focal structure has not been previously reported in Korean. This is probably because the experimental materials in previous studies were generally in the form of pairs of a simple WH-question and an answer, and also unexpected F_0 contours were discarded in order to investigate phonetic variations within one type of F_0 contour shape.

The IP boundary tone is a potentially important cue to information structure for listeners. Speakers seem to highlight the focused constituent by increasing its prosodic boundary strength and separating it from adjacent constituents. The adjustment of duration or pitch (e.g., the use of complex pitch movements) at the right edge seems to be exploited when speakers make a semantic or pragmatic change from a neutral reading. These strategies would complement the consistent marker of narrow focus, the prosodic disjuncture occurring at the left edge of the focused constituent, since the disjunctural cue may not be a sufficient marker to information structure. Although the strength of the left edge is enhanced under narrow focus by F_0 raising in some cases and lengthening, these cues would be ambiguous with respect to the domain of the focus. When listeners perceive a large magnitude of prosodic disjuncture or complex pitch movements, they would be able to confirm that the preceding part of the utterance carried informational prominence. However, it should be noted that the rise on the phrase-final syllable (H% or LH%) does not exclusively signal narrow focus or emphasis in the corrective sense. For example, LH% is associated with questions, continuation rises, explanatory endings, annoyance, irritation or disbelief (Jun, 2000) and also, surprise, incredulity or confirmation (Park, 2012). It is possible that the use of narrow corrective focus in conversation is linked to a negative speaker attitude such as irritation or disbelief.

One of the aims of the present study was to investigate whether focal structure affects the F_0 contour shapes in the utterance. The result is that narrow focus affects the presence or the type of boundary tones in the IP as discussed above, but not within the AP. In Korean, the distribution of the F_0 turning points is less predictable than in West Germanic languages. In West Germanic languages, the turning point is likely to be associated with the lexically stressed syllable (see Chen, 2012 and references therein) and its distribution within an utterance is considered phonologically motivated (Ladd, 2008). In this case, the

conventionalised phonetic cues include the height and scaling of the F_0 turning points which systematically vary depending on the focal structure (e.g., Hanssen, Peters and Gussenhoven, 2008). On the other hand, in Korean, the F_0 turning point locations are strongly affected by phrasing as often signalled by boundary tones, and also partly determined by the number of syllables in the AP and the type of phrase-initial segment. The distribution of pitch turning points apart from the IP boundary tones at the right edge does not seem to be what native Korean speakers directly manipulate in relation to the focal structure.

In the design of the experimental materials, the type of AP-initial segment, which is known to affect the AP-initial tone (L vs. H), was taken into account, since it was expected to interact with focus-related variations in F_0. However, the statistical analysis results regarding temporal adjustment and also F_0 show that the AP-initial segment type does not significantly affect the way speakers adjust the phonetic parameters under narrow focus (one exception is the treatment of F_0 span in defocused constituents, discussed in Section 4.2).

When the target PWs were realised as APs, the expectation about their phonetic shape was generally met in that the majority of the APs (79.25%) showed previously reported frequent F_0 contour shapes such as LH, HH, HLH, LHLH and HLHL and the AP-final H was commonly observed in the data (84.33% of total APs). The phonologised association between the type of segment and the AP-initial tone (see Section 1.1) was observed in general in the dataset, although there were some exceptions (Table 6). The main cause of this exception seems to be that /il/ 'one', which tends to trigger AP-initial H in young people's speech, often appeared in the reading materials (Jun and Cha, 2011, 2015). Although /il/ 'one' begins with a vowel, which is commonly associated with an AP-initial L tone as an initial segment,

Korean speakers often produce the word with a H tone. Jun and Cha (2011, 2015) suggest that the H tone associated with /il/ 'one' is probably originated from a need to distinguish it from /i/ 'two', since the two numbers can be confused in communication.

An expansion in the F_0 span under narrow focus was observed in the data as hypothesised (also e.g., Jun and Lee, 1998; Lee and Xu, 2010), similarly to the increase in the F_0 excursion size of the pitch accent in West Germanic languages. In addition, Cho, Lee and Kim (2011) showed that speakers enhance the consonantally-induced F_0 perturbations at high-information sites in prosodically important positions in Korean; F_0 was higher in the vowels of focused words with HH than those of defocused words. Although the data in the present study were not sufficient to determine corrective focus as a factor conditioning the F_0 variation in phrase-initial syllable, there were some cases where the focus was cued by the boosted F_0 when the neutral target had HH (Figure 4).

We conclude that the focused constituent is always preceded by a prosodic boundary and also signalled by lengthening in Korean, since these effects are consistently observed across studies (e.g., Jun and Lee, 1998; Jun and Kim, 2007; Lee and Xu, 2010; Cho, Lee and Kim, 2011). As hypothesised, the focused constituent could be aligned to higher-level prosodic domain, the IP, than the AP, which is produced in neutral or defocused reading. As demonstrated in Figure 2, although each PW is likely to be realised as an AP in neutral reading, the target PW1 under focus or PW2 in any experimental condition was frequently perceived as an IP. The IPs under focus were more likely to be associated with a bitonal boundary tone (e.g., LH%) than a monotonal one in comparison to neutral or defocused readings. However, there seem be restrictions on the environment in which the IP is formed; for example, IP formation seemed to be affected by the location of the focused constituent within a sentence and also the number of syllables in the target. The results also

suggest that post-focal dephrasing, i.e., deletion of the prosodic boundary as reported in some speakers' speech in Jun and Lee (1998), is not likely to occur when the utterance includes an itemised list and also confirm that dephrasing is only optional (also see Jun and Lee, 1998 and Lee and Xu, 2010 for similar findings). In the present analysis, there were only a few cases which were marked as having merely PW-sized disjuncture (i.e., no perceivable prosodic disjucture) between PW2 and the following carrier phrase.

These prosodic structural cues to narrow focus can be accompanied by the expansion in F_0 span and/or the boundary pitch movement at the right edge. The focus-related increase in acoustic parameters interacts with other factors such as the size of the prosodic unit and the morphosyntactic construction of the utterance. The prosodic markers of focus tend to be concentrated near the phrase edges. Probably due to the lack of lexical prominence, the syllables at the phrase edges seem to serve as the anchoring site of the adjustments in Korean.

It is possible that there are non-prosodic constraints on the phonetic manifestations of focal structure. One potential area for further research is the relationship between morphosyntactic and prosodic marking of narrow focus, particularly the effect of the position of the focused constituent and the omission of the de-focused constituent, which is a common strategy for signalling focal structure in Korean. Although native Korean speakers' tendency to place the focused constituent utterance-initially is noted (Sohn, 1999), the phonetic properties of the utterance-initial focused constituent are not well understood.

4.2. Defocusing in Seoul Korean

The experimental evidence does not fully support the hypothesis that systematic reduction in duration and F_0 would be observed in defocused constituents in comparison to neutrally spoken counterparts. Syllable duration was statistically significantly shorter in defocused PWs only for syllable 5 in 3 + 4 and for syllables 6 and 7 in 4 + 3 than in Neutral. There was no evidence showing F_0 span reduction in defocused constituents; when the PW-initial segment was a weak consonant (for all PW1s and part of PW2s), the F_0 span measurement showed the order focused > defocused > neutral. With the PW-initial strong segment in PW2, F_0 span did not show a statistically significant difference between defocused and neutrally spoken PWs, although it was wider in the focused PWs.

In the present study, the defocused constituent was either pre-focus (PW1 in PW2-focus) or post-focus (PW2 in PW1-focus). Our finding seems to be in contrast to Lee and Xu (2010) who found post-focus reduction in prosodic parameters. In fact, there are possible reasons for the inconsistency across the studies: first, different F_0 properties were measured in different studies. For example, Lee and Xu (2010) report the mean and maximum F_0 values in their target materials; the focused constituent had higher values than neutral, whilst these values were lower in the post-focus constituent. These measurements compare the F_0 level rather than the span, which is the difference between the maximum and the minimum F_0 values, measured in the present study. Second, the experimental tasks differed across studies. In Lee and Xu (2010), speakers answered prompt WH-questions (A: 'what did you say Minswu eats?', B: 'Minswu eats potstickers'). In contrast, in this study and Jun and Lee (1998), listeners were explicitly asked to emphasise the target constituent as a correction of what was previously said. These differences may be related to different focus types. The experimental techniques in the present study would elicit corrective focus (also in Jun and Lee, 1998), unlike those in Lee and Xu (2010) which may correspond to presentational (Gussenhoven, 2007) or informational focus (Féry, 2013). Third, the differences in the

structure of the experimental materials may have led to different results. For instance, in Lee and Xu (2010), the target word seemed to be either utterance-initial as a subject of the sentence or utterance-penultimate as a direct object of the final predicate, although the full materials were not provided. However, in the present study, the defocused constituent was within the target between carrier phrases forming a number sequence with the preceding focused constituent and the target was never utterance-initial or -penultimate. Lee and Xu (2010) actually found a statistically significant effect of the position of the constituent (utterance initial vs. medial), and it is possible that the post-focus compression effect is more easily observed towards the end of the utterance. All in all, there is no strong evidence showing the reduction in prosodic parameters in defocused constituent in comparison to the neutrally spoken counterpart in Korean.

5. Conclusions

Speech data from four Seoul Korean speakers revealed that the constituent under narrow (corrective) focus tends to be acoustically more salient than its neutrally-spoken or defocused counterparts. The focused constituent always begins a new prosodic phrase and the hierarchical level of the prosodic phrase at its right edge shows variation, and speakers' production of the focused constituent boundaries can be exaggerated at both edges. The focal structure did not affect the F_0 contour shapes of the APs and there was no significant effect of the type of phrase-initial segment which tends to be associated with the AP-initial tone type on the measured prosodic parameters in relation to focus.

Under narrow focus, speakers were likely to employ an IP boundary tone compared to when speaking neutrally. Although IP boundary tones with a rise (e.g., H%, LH%) were frequently observed across experimental conditions, the bitonal boundaries (LH%, HL%) were more likely to occur than the monotonal ones (L%, H%) under narrow focus. The formation of IPs

was also affected by how the target was phrased (e.g., 2 + 3, 3 + 2), the number of syllables in the target, and the location of the target in the utterance. Focused constituents also tended to have wider F_0 span and there were cases showing higher phrase-initial F_0 compared to neutrally spoken counterparts.

In addition, focused constituents tended to be longer than neutrally spoken or defocused counterparts. Phrase-initial or -final syllables were lengthened to a greater magnitude than phrase-medial syllables. Focus-related lengthening at the left edge may be related to the articulatory strengthening. The right edge seems to be related to the fact that complex pitch movement frequently occurs in focused constituents delivering semantic or pragmatic meaning.

On the other hand, the defocused constituent did not seem to be distinct from the neutrally spoken counterpart in general. There was no significant difference in the constituent duration and F_0 span between defocused PWs and neutrally spoken PWs. All in all, the results suggest that the consistent marker for focus in Korean is prosodic disjuncture at the left edge, but variations in the phonetic parameters (e.g., increase in F_0 span) or formation of a higher-level prosodic phrase for the focused constituent can be complementary cues to listeners. Unlike in West Germanic languages with lexical stress, the phonetic events related to focus marking tend to be concentrated near prosodic boundaries in Korean.

6. Notes

1. Intonation here refers to the pitch contour excluding collaborative phonetic dimensions such as duration and intensity. This is a more restrictive definition than that in Ladd (2008), where intonation is defined as "the use

of suprasegmental phonetic features to convey 'post-lexical' or sentence-level pragmatic meanings in a linguistically structured way (Ladd, 2008, p. 4)."

2. In Jun (2006), the ip is defined in relation to focal or syntactic structure, i.e., 1) a focused AP which is realised with expanded pitch range, and with a larger magnitude of disjuncture from the preceding AP compared with the AP sequences in an unfocused condition, forms an ip, or 2) an ip is defined when a syntactic clause boundary or a complex syntactic phrase is marked by a larger boundary than that of an AP. Chung and Kenstowicz (1997) also considered the ip in relation to narrow focus. The larger degree of perceivable disjuncture than that between canonical APs would be created before the focused AP when the pitch downtrend is interrupted (Chung and Kenstowicz, 1997; Jun, 2007), or when there is a higher AP-initial tone or AP-final H than that of the canonical AP (Jun 2006, 2007). However, using pitch range or phonetic downstep as a marker to define the ip is problematic as discussed in Jun (2006). The fact that the type of phrase-initial segment (strong vs. weak) also affects the phrasal tones makes the picture even more complicated.

3. A word here includes a lexical item and a following particle or particles (e.g. case marker, delimiters, conjunctive particle, etc.). Although the particle can be considered an independent word (Sohn, 1999, sect. 8.1.), the definition of 'word' adopted here has been commonly used in prosodic analyses in Korean.

4. The definition of pitch accent, deaccenting and dephrasing in French is not consistent across studies. For example, Jun and Fougeron (2000) assume that the Accentual Phrase has the underlying tonal pattern LHiLH*. Hi is an optional initial H which can be promoted to be a pitch accent and the final H* is referred to as a demarcative pitch accent, and the post-focal constituent is deaccented (i.e., realised as a plateau) but not dephrased (i.e., there is no decrease in its duration). On the other hand Féry (2001) refers to the deletion of tones as a consequence of dephrasing.

5. Morpho-syntactic marking of narrow focus is possible but not always transparent in Korean. The focused constituent can be marked by attaching a particle or particles, by changing the word order (i.e., the focused constituent tends to be at the beginning of an utterance) or by clefting (see Sohn, 1999).

6. As in French, the distinction between dephrasing and deaccenting is unclear in Korean. Despite the elusiveness of pitch accents in Korean, Jun and Fougeron (2000) comment that post-focus constituents in Korean are deaccented.

7. The traditional analysis of the Korean vowel system provides ten monophthongs including /y, ø/. However, in contemporary Seoul Korean, these are realised as /wi, wɛ/ respectively and analysed as diphthongs (see Shin, 2015).

8. TOP (topic particle), NM (nominative particle), ENDER (sentence ender)

References

Baumann, S., Grice, M., & Steindamm, S. (2006). Prosodic marking of focus domains: Categorical or gradient?. In *Proceedings of the Speech Prosody 2006*, 301–304. Dresden, Germany, May 2–5.

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2014). *lme4: Linear mixed-effects models using Eigen and S4*. R package version 1.1-7, URL: http://CRAN.R-project.org/package=lme4.

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language, 59*, 390–412. doi: 10.1016/j.jml.2007.12.005

Beckman, M., & Pierrehumbert, J. (1986). Intonational structure in Japanese and English. *Phonology Yearbook, 3*, 225–309. doi: 10.1017/S095267570000066X

Boersma, P., & Weenink, D. (2010). *Praat: Doing Phonetics By Computer* [Computer program]. Version 5.1.31. retrieved 4 April 2010 from http://www.praat.org/.

Cambier-Langeveld, T., & Turk, A. E. (1999). A cross-linguistic study of accentual lengthening: Dutch vs. English. *Journal of Phonetics, 27*, 255–280. doi:10.1006/jpho.1999.0096

Chen, A. (2012). The prosodic investigation of information structure. In M. Krifka, & R. Musan (eds.), *The Expression of Information Structure* (pp. 249–286). Berlin: de Gruyter.

Chen, Y. (2006). Durational adjustment under corrective focus in Standard Chinese. *Journal of Phonetics, 34*, 176–201. doi: 10.1016/j.wocn.2005.05.002

Chen, Szu-wei., Wang, B., & Xu, Y. (2009). Closely related languages, different ways of realizing focus. In *Proceedings of the Interspeech*, 1007–1010. Brighton, UK, 6–10 September.

Cho, T., & Keating, P. (2001). Articulatory and acoustic studies on domain-initial strengthening in Korean. *Journal of Phonetics, 29*, 155–190. doi:10.1006/jpho.2001.0131

Cho, T., Lee, Y., & Kim, S. (2011). Communicatively driven versus prosodically driven hyper-articulation in Korean. *Journal of Phonetics, 39*, 344–361. doi:10.1016/j.wocn.2011.02.005

Chung, S., & Kenstowicz, M. (1997). Focus expression in Seoul Korean. *Harvard Studies in Korean Linguistics, 12*, 93–105. Montréal, 2–7 June.

Dimitrova, S., & Turk, A. (2012). Patterns of accentual lengthening in English four-syllable words. *Journal of Phonetics, 40*, 403–418. doi:10.1016/j.wocn.2012.02.008

Eady, S., & Cooper, W. (1986). Speech intonation and focus location in matched statements and questions. *Journal of the Acoustical Society of America, 80*, 402–415.

Eady, S. J., Cooper, W. E., Klouda, G. V., Mueller, P. R., & Lotts, D. W. (1986). Acoustical characteristics of sentential focus: Narrow vs. broad and single vs. dual focus environments. *Language and Speech, 29*, 233–251. doi:10.1177/002383098602900304

Féry, C. (2001). Focus and phrasing in French. In C. Féry & W. Sternefeld (eds.), *Audiatur vox Sapientiae: A Festschrift for Arnim von Stechow* (pp. 153–181). Berlin: Akademie Verlag.

Féry, C. (2013). Focus as prosodic alignment. *Natural Language and Linguistic Theory, 31*, 683–734. doi:10.1007/s11049-013-9195-7

Féry, C. & Kügler, F. (2008). Pitch accent scaling on given, new and focused constituents in German. *Journal of Phonetics, 36*, 680–703. doi:10.1016/j.wocn.2008.05.001

Gussenhoven, C. (2002). Intonation and interpretation: Phonetics and phonology. In *Proceedings of the Speech Prosody 2002*, 47–57. Aix-en-Provence, France, 11–13 April.

Gussenhoven, C. (2004). *The Phonology of Tone and Intonation*. Cambridge, UK: Cambridge University Press.

Gussenhoven, C. (2007). Types of focus in English. In C. M. Lee, M. Gordon, & D. Büring (eds.), *Topic and Focus: Cross-linguistic Perspectives on Meaning and Intonation* (pp. 83–100). Dordrecht: Springer.

Hanssen, J., Peters, J., & Gussenhoven, C. (2008). Prosodic effects of focus in Dutch declaratives. In *Proceedings of the Speech Prosody 2008*, 609–612. Campinas, Brazil, 6–9 May.

Hothorn, T., Frank Bretz, F. and Peter Westfall, P. (2008). Simultaneous inference in general parametric models. *Biometrical Journal, 50(3)*, 346–363. doi: 10.1002/bimj.200810425

Jeon, H.-S. (2011). *Prosodic Phrasing in Seoul Korean: The Role of Pitch and Timing Cues. Ph.D. dissertation*. Department of Theoretical and Applied Linguistics, University of Cambridge.

Jeon, H.-S. (2015). Prosody. In L. Brown & J. Yeon. *Handbook of Korean Linguistics* (pp. 41–58). New West Sussex, UK: Wiley-Blackwell.

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

Jun, S.-A. (1996b). Influence of microprosody on macroprosody: A case of phrase initial strengthening. *UCLA Working Papers in Phonetics, 92,* 97–116.

Jun, S.-A. (1998). The Accentual Phrase in the Korean prosodic hierarchy. *Phonology, 15,* 189–226.

Jun, S.-A. (2000). K-ToBI (Korean ToBI) labelling conventions (Ver. 3.1). *UCLA Working Papers in Phonetics, 99,* 149–173.

Jun, S.-A. (2005). Korean intonational phonology and prosodic transcription. In S.-A. Jun (ed.) *Prosodic Typology* (pp. 201–229). Oxford: Oxford University Press.

Jun, S.-A. (2006). Intonational phonology of Seoul Korean revisited. *UCLA Working Papers in Phonetics, 104,* 14–25.

Jun, S.-A. (2007). The intermediate phrase in Korean: Evidence from sentence processing. In C. Gussenhoven & T. Riad (eds.), *Tones and Tunes: Experimental Studies in Word and Sentence Prosody,* 143–170. Berlin, New York: Mouton de Gruyter.

Jun, S.-A. (2012). Prosodic typology revisited: Adding macro-rhythm. In *Proceedings of Speech Prosody 2012,* 535–538. Shanghai, China, 22–25 May.

Jun, S.-A., & Cha, J. (2011). High-toned [il] in Seoul Korean intonation, In *Proceedings of the International Congress of Phonetic Sciences XVII,* 990–993. Hong Kong, 17–21 August.

Jun, S.-A. & Cha, J. (2015). High-toned [il] in Korean: Phonetics, intonational phonology, and sound change. *Journal of Phonetics, 51,* 93–108. doi:10.1016/j.wocn.2015.05.002

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

Jun, S.-A., & Kim, H.-S. (2007). VP focus and narrow focus in Korean. In *Proceedings of the XVIth International Congress of Phonetic Sciences,* 1277–1280. Saarbrücken, Germany, 6–10 August.

Jun, S.-A., & Kim, S. (2004). Default phrasing attachment preference in Korean. In *Proceedings of Interspeech 2004- International Conference on Spoken Language Processing,* 3009–3012. Jeju, Korea, 4–8 October.

Jun, S.-A., & Lee, H.-J. (1998). Phonetic and phonological markers of contrastive focus in Korean. In *Proceedings of the 5th International Conference on Spoken Language Processing,* 1295–1298. Sydney, Australia, 30 November–4 December.

Kim, Sahyang. (2004). *The Role of Prosodic Phrasing in Korean Word Segmentation*. Ph.D. dissertation, Department of Linguistics, UCLA.

Kim, Sunhee. (2009). Intonation patterns of Korean spontaneous speech. *Malsoliwa Umsengkwahak* [Journal of the Korean Society of Speech Sciences], *1*, 85–96.

Kim, S., Yoo, H., Hong, H., & Lee, H.-Y. (2007). Momelul iyonghan hankwukeuy ekyang yenkwu [A study on Korean intonation using Momel]. *Malsori* [Journal of the Phonetic Society of Korea] *63*, 85–100.

Kügler, F. (2008). The role of duration as a phonetic correlate of focus. In *Proceedings of Speech Prosody 2008*, 591–594. Campinas, Brazil, May 6–9.

Ladd, D. R. (2008). *Intonational Phonology* (2nd ed.). Cambridge, UK: Cambridge University Press.

Lee, A., & Xu, Y. (2012). Revisiting focus prosody in Japanese. In *Proceedings of Speech Prosody 2012*, 274–277. Shanghi, China, 22–25 May.

Lee, Y., & Xu, Y. (2010). Phonetic realization of contrastive focus in Korean. In *Proceedings of Speech Prosody* 2010, 100033, 1–4. Chicago, 11–14 May.

Nespor, M. & Vogel, I. (1986). *Prosodic Phonology*. Dordrecht: Fortis.

Nolan, F., & Jeon, H.-S. (2014). Speech rhythm: a metaphor?. *Philosophical Transactions of the Royal Society B: Biological Sciences, 396*, 20130396. doi:10.1098/rstb.2013.0396

Park, M.-J. (2012). *The Meaning of Korean Prosodic Boundary Tones*. Leiden, The Netherlands: Brill.

Pinheiro, J. C. & Bates, D. M. (2000). *Mixed-Effects Models in S and S-PLUS*. New York: Springer-Verlag.

Peters, J., Hanssen, J., & Gussenhoven, C. (2014). The phonetic realization of focus in West Frisian, Dutch Low Saxon, German Low Saxon, High German, and three varieties of Dutch. *Journal of Phonetics, 46*, 185–209. doi:10.1016/j.wocn.2014.07.004

Peterson, G., & Lehiste, I. (1960). Duration of syllable nuclei in English. *Journal of the Acoustical Society of America, 32*, 693–703.

R Development Team. (2015). *R: A Language and Environment for Statistical Computing, ver. 3.20* [computer software].

Schafer, A. & Jun, S. (2002). Effects of accentual phrasing on adjective interpretation in Korean. In M. Nakayama (ed.), *East Asian Language Processing* (pp. 223–255). Stanford: Center for the Study of Language and Information Publications.

Selkirk, E. O. (1984). *Phonology and Syntax: The Relation between Sound and Structure*. Cambridge, MA, and London: MIT Press.

Selkirk, E. O. (1986). On derived domains in sentence phonology. *Phonology Yearbook, 3*, 371–405.

Sohn, Ho-Min. (1999). *The Korean Language*. Cambridge, UK: Cambridge University Press.

Shin, J. (2015). Vowels and consonants. In Brown, L. & Yeon, J. (eds.). *The Handbook of Korean Linguistics* (pp. 3–21). West Sussex, UK: Wiley-Blackwell.

Turk, A. E., & White, L. (1999). Structural influences on accentual lengthening in English. *Journal of Phonetics, 27*, 171–206. doi:10.1006/jpho.1999.0093

Turk, A., Nakai, S., & Sugahara, M. (2006). Acoustic segment durations in prosodic research: A practical guide, In S. Sudhoff, D. Lenertová, R. Meyer, S. Pappert, P. Augurzky, I. Mleinek, N. Richter, & J. Schließer (eds.), *Methods in Empirical Prosody Research* (pp. 1–28). Berlin/New York: De Gruyter.

Venditti, J. J. (2005). The J_ToBI model of Japanese Intonation, In S.-A. Jun (ed.), *Prosodic Typology* (pp. 172–200). Oxford: Oxford University Press.

Venditti, J. J., Maekawa, K., & Beckman, M. E. (2008). Prominence marking in the Japanese intonation system, In S. Miyagawa & M. Saito (eds.), *The Oxford Handbook of Japanese Linguistics* (pp. 456–512). Oxford: Oxford University Press.

Wang, B., & Xu, Y. (2011). Differential prosodic encoding of topic and focus in sentence-initial position in Mandarin Chinese. *Journal of Phonetics, 39*, 595–611. doi:10.1016/j.wocn.2011.03.006

Xu, Y. (1999). Effects of tone and focus on the formation and alignment of f0 contours. *Journal of Phonetics, 27*, 55–105. doi:10.1006/jpho.1999.0086

Xu, Y., Chen, S.-W., & Wang, B. (2012). Prosodic focus with and without post-focus compression (PFC): A typological divide within the same language family?. *The Linguistic Review, 29*, 131–147. doi:10.1515/tlr-2012-0006

Xu, Y., & Xu, C. X. (2005). Phonetic realization of focus in English declarative intonation. *Journal of Phonetics, 33,* 159–197. doi:10.1016/j.wocn.2004.11.001

Yang, A., Cho, T., Kim, S., & Chen, A. (2015). Phonetic focus-marking in Korean-speaking 7- to 8-year olds and adults. In *Proceedings of the 18th International Congress of Phonetic Sciences*, 673, 1–9, Glasgow, UK, 10–14 August.

Welby, P. (2002). The realization of early and late rises in French: A production study. In Proceedings of the Speech Prosody 2002, 695–698. Aix-en-Provence, France, 11–13 April.

Welby, P. (2006). French intonational structure: Evidence from tonal alignment. *Journal of Phonetics, 34*, 343–371. doi:10.1016/j.wocn.2005.09.001