Phonetic manifestation of focus particles in Korean: *hangsang* ('always') and *ocik* ('only')

Abstract

There is a long-standing debate about whether focus particles function as focus-sensitive operators or focus inducers. However, it has not yet been established which perspective is correct. The current study investigates the effect of focus particles, conducting a production experiment. The results show that the focus particles are focus-sensitive operators, since the focused words are elicited by the discourse context. Moreover, prosodic differences are observed between *hangsang* and *ocik*. The former has more increased F0, duration, and intensity values. This result suggests that *hangsang* bears notable prosodic features in and of itself. As an alternative model, the Parallel Encoding and Target Approximation model is used. (105 words)

Index Terms: focus particle, focus-sensitive operator, focus inducer, PENTA, intonational function

1. Introduction

Focus particles (henceforth, FPs) are believed to be associated with information structure cross–linguistically (Beaver & Clark 2003). Yet, it is often debated whether the FPs serve as focus inducers or focus-sensitive operators. The first point of view is that the FPs behave like focus inducers (Bayer 1996, Jacobs 1986, Kadmon 2001, Taglicht 1984). In this scenario, they are assumed to assign a focus feature to some domain. However, the counterview is that the focus-background relationship cannot be built by the FPs because *focus* can be determined and elicited by the context (Büring & Hartmann 2001, Sudhoff 2010). For this reason, the focused element is not influenced by the FPs. Rather, it is independent of the presence of the FPs.

Given the disagreement about the nature of the FPs, the goal of the present study is to investigate which claim is more adequate for Korean FPs. In addition, it is of interest to see whether or not there is a prosodic difference between the FPs.

The present study adopts the Parallel Encoding and Target Approximation (henceforth, PENTA) model to describe the pitch contour, rather than the commonly-used ToBI (Tones and break indices) system. The main reason for the use of the PENTA model is that it depicts the elements of intonation primarily in form. It is true that speech conveys several communicative functions: topic, focus, demarcation, attitudinal functions, etc. Thus, it is necessary to adopt the PENTA model because it describes that communicative elements are concurrently conveyed through speech melody by function, not by form (Xu 2004, 2005, 2009, Xu & Xu 2005). The best advantage of the PENTA model is that fluent intonational functions can be simultaneously realized through encoding schemes. Thus, we will illustrate F0 contours of the target sentences under the PENTA model. It is believed that the PENTA model will describe the sentences that we investigate, which contains multiple functions, such as a focus particle, focus, and post-focus compression in parallel.

This paper is organized as follows: in Section 2, we will overview the differences of the FPs how they have been treated in the literature. Then, the PENTA model will be introduced as an alternative to the British nuclear tone tradition and the ToBI system. Section 3 describes the experimental methodology and reports the findings of the current study. In Section 4, the results will be analyzed and the PENTA model will be applied to the Korean language. Further, unresolved issues will be discussed. Section 5 summarizes this paper.

2. Theoretical backgrounds

2.1. Focus inducer

As mentioned previously, FPs are known to be associated with information structure (Beaver & Clark 2003). Therefore, as in (1)[1], they are believed to assign a focus feature to the following constituent.

(1) a. I *always* play [basketball]$_{FOC}$.

b. I *only* play [basketball]$_{FOC}$.

As in (1), *basketball* receives a focus by virtue of the FPs. The FP is called a *focus inducer* since it literally induces a focus. The same goes for Korean, as in (2).

(2) a. Minswu–nun *hangsang* [nongkwu–lul]$_{FOC}$ ha–n–ta.

Minswu–Top always basketball–Acc play–Prs–Decl

'Minswu always plays basketball.'

b. Minswu–nun *ocik* [nongkwu–lul]$_{FOC}$ ha–n–ta.

Minswu–Top only basketball–Acc play–Prs–Decl

'Minswu only plays basketball.'

As in (2), the focus on the element, *nongkwu–lul* is triggered by the FPs. Even though *hangsang* and *ocik* are assumed to equally elicit a focused element, they show different semantic aspects. This is also exemplified in English 'always' and 'only' in (3).

[1] Abbreviations used: Acc (accusative), Decl (declarative), FOC (focused element), Prs (Present), Top(Topic)

(3) Does Sandy feed Nutrapup to her dogs? (Beaver & Clark 2003: 327)

a. Yes, Sandy *always* feeds Nurapup to $[\text{Fido}]_{\text{FOC}}$, and she also *always* feeds Nutrapup to $[\text{Butch}]_{\text{FOC}}$.

b. *Yes, Sandy *only* feeds Nurapup to $[\text{Fido}]_{\text{FOC}}$, and she also *only* feeds Nutrapup to $[\text{Butch}]_{\text{FOC}}$.

As is shown, (3a) and (3b) are identical except the FPs. But, *only* is not appropriate for the situation in (3b). According to Beaver and Clark (2003), *only* is an exhaustive focus particle because it must denote a unique entity. Conversely, non-exhaustive interpretations are allowed for *always*. The two FPs are also distinguished in a way that *always* can interact with a reduced form, whereas *only* cannot.

(4) How often was Fred the person you talked about? (Beaver & Clark 2003: 343)

a. I $[\text{always}]_{\text{FOC}}$ disccused'im with Sandy.

b. #I $[\text{only}]_{\text{FOC}}$ disccused'im with Sandy.

Unlike *always, only* must interact with an R–expression or a full pronoun (Beaver & Clark 2003). To state it differently, *only* should not be associated with an unstressed or unfocused element.

2.2. Focus-sensitive operator

Sudhoff (2010) argues that focus is not determined or induced by an FP. Rather, it is elicited by a discourse context. See (5)[2].

[2] In (5), the focused element is underlined and the sentence with the focus is italicized.

(5) a. Contrastive focus with an FP. (Sudhoof 2010:5)

Three pupils of class 10a earned some money in the last week of the school vacation by now and then cleaning machines in the BMW factory. Unfortunately, Friday's attendance list went missing. The secretary told the head of department that all three pupils had worked on that day. But she was wrong. *On Friday*, *only* <u>*Sabine*</u> *came*. The others weren't in the mood for working any longer.

b. Contrastive focus without an FP. (Sudhoof 2010: 6)

The whole family wanted to spend Christmas together, and in the week before Christmas Eve, one by one arrived at home in Regensburg. On Friday, there was a big surprise. Andreas had told us that Franziska would arrive on that day. But he had completely mixed it up. *On Friday,* <u>*Sabine*</u> *came*. Franziska still had to work.

As in (5), the only difference between them is that (5a) bears an FP; (5b) does not. The contrastive focus both in (5a) and (5b) are triggered by the discourse context even though (5b) has the FP, *only*. According to Sudhoff (2010), there is no prosodic difference between the two conditions. This is because the focus on *Sabine* is already elicited by the discourse function. Accordingly, Sudhoff maintains that the occurrence of an FP in a sentence does not give rise to a prosodic difference. (6) is also a counter-example to the view of *focus inducers*.

(6) How often do you play basketball?

a. I [always]$_{FOC}$ play basketball.

b. #I always play [basketball]$_{FOC}$.

In (6a), a contrastive focus falls on *always* and becomes emphatic. Yet, (6b) is inadequate because *basketball* does not respond to the question, *how often*. If the FP was a *focus inducer*, it should induce a focus feature on an element that relates to the FP, regardless of the context. However, in (6a) *basketball* cannot receive a focus feature by virtue of the FP, *always*. This may imply that the FP cannot always induce a focus feature to the following domain. Consequently, it is possible to speculate that the discourse context takes priority over the focus inducer.

2.2. The British nuclear tone tradition

In the British nuclear tone tradition, intonation is described as contours. The basic structure of the tone unit is as follows: *prehead*, *head*, *nucleus*, and *tail* (Crystal 1969, Ladd 1996). In this classification, a *nucleus* is required to form intonational contours, and it is often recognized as “the stressed syllable of the last prominent word in a sense group” (O’Connor & Arnold 1961: 271, cited in Xu & Xu 2005). The F0 contour after the nucleus is the *tail*. The F0 contours ahead of the nucleus are the *prehead* and *head* as illustrated in Figure 1.

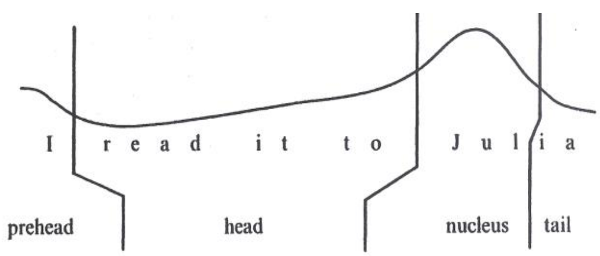


Figure 1. Division of the pitch contour in the typical analysis of the British school. (Adapted from Ladd 1996: 210)

2.3. The Autosegmental-Metrical (AM) Theory

The Autosegmental-Metrical (henceforth, AM) theory uses two levels, high (H) and low (L), to examine the intonation of an utterance (Beckman & Pierrehumbert 1986, Ladd 1996, Pierrehumbert 1980, Pierrehumbert & Hirschberg 1990). More specifically, the AM theory proposes *pitch accent* and *edge tone*. Edge tone is further subdivided into two units: *phrase accent* and *boundary tone* as in Figure 2.

The diacritic * marks a pitch accent. A phrase accent is marked by either H− or L−, whereas a boundary tone is marked by either H% or L%. In Figure 2, we see seven combinations[3] of pitch accents. According to Pierrehumbert (1980), a pitch accent can be realized individually or in combination (double-toned accents) with a phrase accent or a boundary tone.

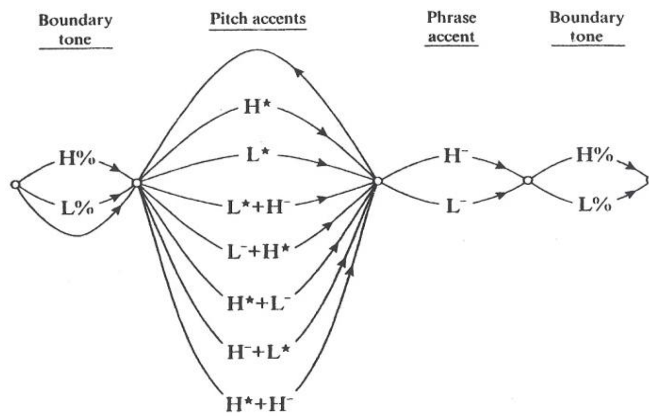


Figure 2. A finite-state grammar which generates the set of well-formed tonal sequences for an intonation phrase. (Adapted from Ladd 1996: 81, Pierrehumbert 1980: 29)

It is generally accepted that the ease of transcription is one of the main merits of the AM theory. However, despite this convenience, the AM theory has drawn several criticisms. SP Oh (2007) assumes that the AM theory does not sufficiently explain which criteria are used to divide intonation into H and/or L. Xu and Xu (2005: 163) also indicate that the theory "provides no mechanistic explanation for the exact alignment of F0 peaks and valleys with that syllable." Also, the AM theory proposes that F0 contours should have exact shapes, such as H*+L or L+H* (Xu & Xu 2005). Likewise, neither single tones nor double-toned accents of the AM theory can demonstrate a fluent repertoire of multiple intonational

[3] The interpretation of pitch accents can be altered depending upon which tone is followed by H* or L*. See Pierrehumbert and Hirschberg (1990) for more detailed information.

functions.

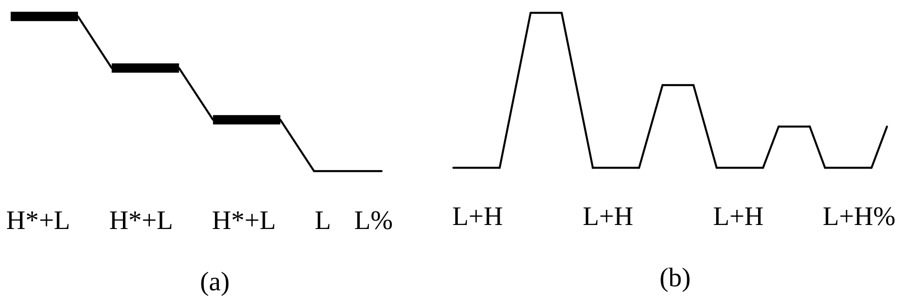


Figure 3. Schematic F0 contours. The bold lines represent stress. (Adapted from Pierrehumbert & Hirschberg 1990: 282)

As shown in Figure (3a), the sequence H*+L occurs three times in a row. However, we find a loophole in the AM theory; the second high tone (H*) is situated in a lower level than the first low tone (L). In Figure (3b), the sequence L+H appears as a high plateau followed by a drop at the end. The sequence occurs twice more. However, although each is transcribed into H, the height of the three plateaus is not the same. This is another area that the AM theory has failed to address[4].

2.4. Form vs. Function

Despite the fundamental differences, the elements of intonation in both the British nuclear tone tradition and the AM theory are defined primarily in form rather than in function (see Xu & Xu 2005). In the former system, intonation and tones are illustrated as contours. In the latter system, an H is considered as an F0 peak, whereas an L denotes an F0 valley.

The form-based systems cannot fully explain a tone language (Xu & Xu 2005). They explained that lexical tones "are defined not first by form, but by function" (Xu & Xu 2005: 161). Here, one may

[4] Xu (p.c.) states that "the most problematic aspect of AM theory is its circularity in defining the pitch accents and their component tones. These units are defined primarily by their observed forms."

point out that Korean is not a tone language. Yet, it is likely that the following communicative multiple functions in Korean can be included, as they are evident in Mandarin: lexical stress, sentence types, topic, focus, sarcasm, demarcation, etc.

2.5. Focus as a communicative function

Focus is used to emphasize a certain constituent in a discourse situation (Ladd 1996, Xu & Xu 2005, among others).

(7) a. Who ate my pizza?

b. **ANDREW** ate your pizza.

We find that *Andrew* in sentence (7b) receives a contrastive focus. In the British nuclear tone tradition, the focus is defined as a high-fall nuclear pitch accent. In the AM theory, however, it is realized with an L+H* sequence. Yet, according to the definition of both systems, if one phrase receives a nuclear pitch accent, the pitch accent will no longer be visible in the sentence (Xu & Xu 2005). This definition, however, is not consistent with the finding that lexical tones are not replaced and deleted by focus (Chen 2003, Gåring 1987, Shih 1988, S Jin 1996, Xu 1999, Xu & Xu 2005). That is, both focus and lexical tones are simultaneously conveyed through the given speech. The focused constituent neither replaces nor eliminates a lexical tone. In this sense, we do not adopt a widely used ToBI framework in which intonation and/tones are illustrated primarily by form, not by function. Therefore, as an alternative, the PENTA model proposed by Xu (2004, 2005, 2009) and Xu and Xu (2005) is used in this study.

2.6. PENTA Model (Xu 2004, 2005, 2009, Xu & Xu 2005)

Xu developed the PENTA model, which illustrates communicative elements of speech melody in terms of function, not form. This model points out the inadequacies of the British nuclear tone tradition and the AM theory, in which the elements of intonation are defined primarily in form. In both systems, semantic components are realized after the form is set up.

In a tone language, however, lexical tones are defined not by form but by function. They help one to distinguish words that are morphologically identical as (8) illustrates.

(8) [ma] (Xu 2005)

a. uttered with high-level contours ⇒ "mother"

b. uttered with rising contours ⇒ "hemp"

c. uttered with low-dipping contours ⇒ "horse"

d. uttered with falling contours ⇒ "to scold"

As can be seen, the meaning of the syllable [ma] in Mandarin is differentiated according to dissimilar pitch patterns.

Consider (9), where *focus* emphasizes some parts in the contexts and/or discourses.

(9) Who made the chair?

a. **ANDREW** made the chair.

b. *Andrew made the **CHAIR**.

In (9a), the element, *Andrew*, is focused in order to answer the question. Yet, (9b) is inadequate because the element, *chair*, does not respond to the question, *who*. Likewise, a focused constituent is more prosodically prominent than the neutral-focus one. According to the definition of the nuclear tone tradition and the AM theory, *Andrew* is assigned a high-fall nuclear pitch accent in the former and L+H* in the latter. Xu and Xu (2005) argued that if a word or a syllable is assigned the nuclear pitch accent, prominent pitch events will no longer exist in the sentence, based on both theories. However, focus never replaces the lexical tones in a Mandarin utterance. Instead, focus and tones are realized in parallel by showing different characteristics of F0 contours.

With this in mind, Xu (2004, 2005) and Xu and Xu (2005) proposed the PENTA model. It is based on function, not form. It denotes mechanisms of multiple intonational functions simultaneously conveyed through speech melody. It also shows relationships among the multiple functional elements and F0 contours.

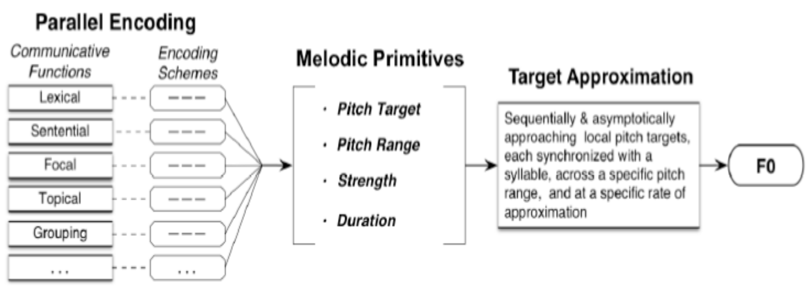


Figure 4. A schematic sketch of the PENTA model. The unnamed block at the bottom left indicates communicative functions yet to be identified. (Adapted from Xu 2005: 243)

Figure 4 shows the PENTA model, in which communicative functions on the far left are realized through speech melody. These functions then control F0 contours through encoding schemes. Following this stage, such encoding schemes determine melodic primitives (e.g. pitch target, pitch range, strength, and duration). According to Xu (2005), the encoding schemes connect communicative functions [+ meaningful] with melodic primitives [− meaningful]. We will now inspect the notations of the PENTA model.

Table 1. Possible symbolic values of the melodic primitives: local target, pitch range, and articulatory strength, which may be notationally distinguished from one another by using [ ], underline, **boldface** and *italic*, respectively[5]. (Adapted from Xu 2005: 243)

Local Target		[high], [low], [rise], [fall], [mid]
Pitch Range	Height	high, low, mid
	Span	wide, narrow, normal
Articulatory Strength		**strong**, **weak**, **normal**
Duration		*long*, *show*, *normal*

As shown in Table 1, local pitch targets can be either static ([high], [low], or [mid]) or dynamic ([rise] or [fall]). Then pitch ranges are divided into two types: *height* and *span*. Height represents the relative height of the pitch range (high, low, or mid), whereas span represents the width of the pitch range (wide, narrow, or normal). In addition, articulatory strength is related to the speed of the target being approached. <Strong> articulatory strength is in general faster than its counterpart. Lastly, duration is subdivided into three categories: *long*, *short*, and *normal*.

Now let us look at the actual transmission of the sentence, *Lee may mimic my niece*, applied to the PENTA model. In this sentence, *mimic* receives a focus feature. As control data, the same sentence with no focus on each word was also measured to acquire average F0 curves. In Xu and Xu (2005), lexical stress, sentence types, and focus were used as independent functions in order to verify whether multiple intonational functions are conveyed through speech melody in parallel rather than in alternation.

[5] According to Xu (p.c), the symbolic values of the melodic primitives might differ according to the language. Different languages might have their own inventories.

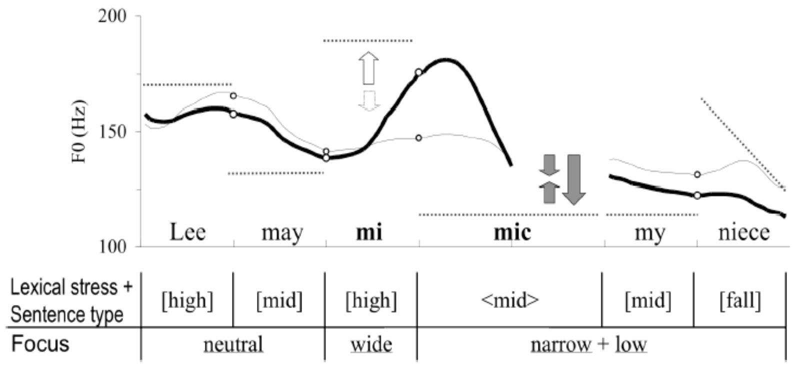


Figure 5. Decomposition of the F0 contours of “Lee may mimic my niece” according to the PENTA model. Top: Graphic decomposition. Thick solid curve: focus on “mimic”; thin solid curve: no narrow focus. Short straight lines represent hypothetical local pitch targets. Unfilled block arrows indicate on-focus pitch range expansion. Filled block arrows indicates post-focus pitch range lowering and narrowing. Bottom: symbolic decomposition. (Adapted from Xu & Xu 2005: 191)

In Figure 5, the sentence, *Lee may mimic my niece*, is described under the PENTA model. Here, the focused word, *mimic*, underwent a pitch range expansion, whereas the second unstressed syllable after focus experienced pitch range narrowing and lowering. Xu and Xu (2005: 192) attribute this to the idea that “the apparent F0 peak near the beginning of the second syllable in ‘mimic’ is likely due to inertia of the larynx whose movement cannot be fully reversed at the syllable boundary.”

Second, the F0 of *may*, *-mic*, and *my* is assigned as [mid]. This is because the height is influenced and/or readjusted by the pitch range specification under focus. The readjustment of the pitch range can also be observed as being influenced by other functions, such as topic, turn-taking, and demarcation (Xu & Xu 2005).

Last, the F0 contour in the stressed syllable is a [high] target, not a [rise] one; even though the stress falls on the syllable *mi* and the sharp rising is on the whole syllable. This is because “the F0 peaks mostly occur before the offset of the stressed syllable unless the syllable duration is very short” (Xu & Xu 2005: 192). The second likely reason is that the initial F0 valleys constantly appear close to the beginning of the stressed syllable.

In summary, the PENTA model shows that communicative elements are simultaneously conveyed through speech melody by function, not by form. Multiple intonational functions are simultaneously realized through encoding schemes. Then, the encoding schemes assign values on melodic primitives including pitch target, pitch range, strength, and duration. The PENTA model can explain a fluent repertoire of multiple intonational functions that are simultaneously conveyed through F0 contours.

3. Methodology

3.1. Stimuli

In the current study, four sets of data served as stimuli. The first three sets had two FPs: *hangsang* and *ocik*. In addition, they were paired with three different contexts. In the first set, the FPs were given without context. The corresponding items in the second set were preceded by a prompt question. In the third set, the discourse context was given in order to elicit the focus effect. Finally, the last set did not include an FP, and focus was elicited through the discourse context. Also, the number of words ranged from four to six. In total, 216 sentences were used. Among them, 126 sentences served as target sentences. The rest were used either as prompt questions or discourse contexts. See the following sample stimuli, where the target sentences are in square brackets and the FPs are in angle brackets.

(4) Prompt question + FP (*hangsang*)

Q: Hangsang mwues.ul cohahaseyyo?

A: [Nanun <hangsang> mantwulul cohahapnita].

'What do you always like? I always like dumplings.'

(5) Context + FP (*hangsang*)

Ce nun elyessul ttaypwuthe han kaci cohahanun umsiki isssupnita. Pika okena myengcel naley hokun ceyka aphul ttay celul wihayse nwunimkkeyse sonswu picecwusin mantwuka isssupnita. Kulayse, [nanun <hangsang> mantwulul cohahapnita].

'There is something I have liked since I was young. When it rained, when it was a holiday, or when I was sick, my elder sister used to make dumplings for me. For this reason, I always like dumplings.'

(6) Context + No FP

Cenun elyess.ul ttaypwuthe han kaci coh.ahanun umsiki isssupnita. Pika okena myengcel naley hokun ceyka aphul ttay celul wihayse nwunimkkeyse sonswu picecwusin mantwuka issnuntey, ilen iyulo [nanun <EMPTY> mantwulul cohahapnita].

'There is something I have liked since I was young. When it rained, when it was a holiday, or when I was sick, my elder sister used to make dumplings for me. For this reason, I <EMPTY> like dumplings.'

3.2. Subjects

Three males and three females participated in the experiment. All participants were native speakers of Korean. We recruited the subjects at the University of Pennsylvania and paid them for their participation. The participants did not exhibit problems with their speech and hearing nor noticeable accents and/dialects.

3.3. Procedure

The stimuli were recorded in a sound-proof booth in the Department of Linguistics at the University of

Pennsylvania. A head-mounted microphone was used for the experiment. The recordings were made electronically and saved directly on a computer through Praat (Boersma 2001). Before the recordings, the materials were first presented to the subjects in order for them to become familiar with the materials. During the session, the subjects were allowed to drink water and take a short break when they felt tired. However, they were asked not to pause while reading. In addition, they were instructed to repeat the stimuli when a mispronunciation or mistake was found. The stimuli were presented on a sheet of paper in a randomized order.

3.4. F0 extraction and exclusion of outliers

A Praat script was used to measure the acoustic parameters of the target sentences (Xu 2005-2009). In order to extract F0, word boundaries were marked by hand for the overall F0. Also, syllable boundaries were manually labeled for the FPs. A logarithmic algorithm was also performed in order to smooth over abrupt bumps and sharp edges. After the process of F0 extraction, all the target sentences were converted to graphs (see Figure 6). Then, time-normalized F0 curves of all the target sentences were computed, using all the tokens produced by the six subjects, as illustrated in Figure 7. Next, each F0 curve was inspected in order to exclude outliers. The criteria of eliminating the outliers were shown in (7).

(7) An F0 curve was eliminated if and only if:

a. it was definitely different from other curves.

b. subjects performed mispronunciations.

Of 756 target sentences, six F0 curves were removed. To illustrate, one F0 curve was eliminated from each subject, respectively. The eliminated curves were not used for further measurements.

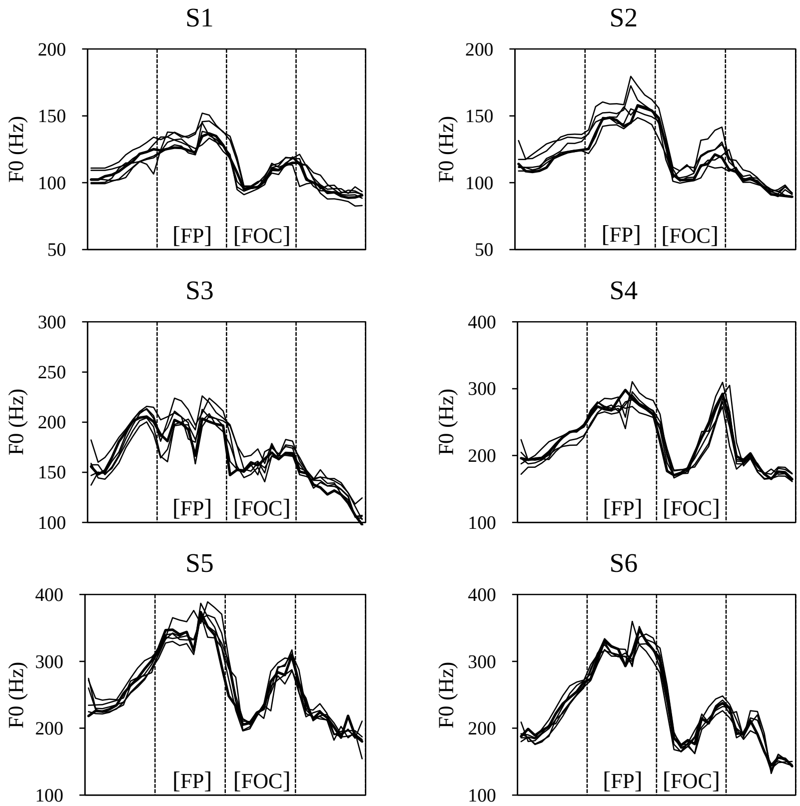


Figure 6. Each graph indicates time-normalized F0 curves of six repetitions of the sentence, *Nanun hangsang, mantwulul cohahapnita* ('I always like dumplings.') by six speakers. Each word is bordered by a vertical line. Subjects 1-3 are male speakers, and subjects 4-6 are female.

3.5. Grand mean F0 curves

After excluding the outliers, mean F0 curves of all the sentences were obtained. As mentioned previously, the method for acquiring mean F0 curves was to compute time-normalized F0 curves of all the target sentences. It is expected that grand mean F0 curves facilitate comparisons across the three FP conditions.

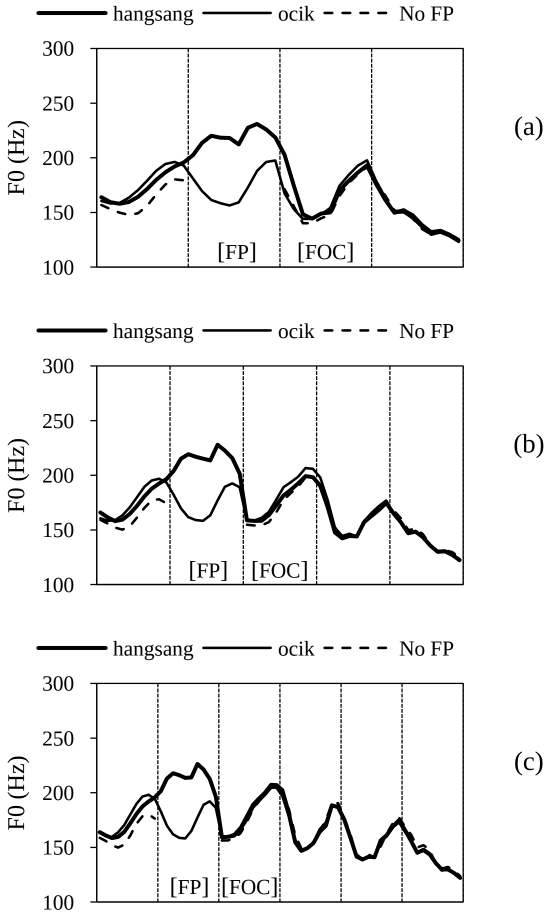


Figure 7. Mean F0 curves of all the sentences spoken by four speakers. Each F0 curve is an average of 24 repetitions: a) *Nanun* {*hangsang, ocik, EMPTY*} *mantwulul cohahapnita*; b) *Nanun* {*hangsang, ocik, EMPTY*} *nwunimuy mantwulul cohahapnita*; c) *Nanun* {*hangsang, ocik, EMPTY*} *nwunimi mantusin mantwulul cohahapnita*. (The items in braces indicate the word enclosed by the second and third areas of each figure)

In Figure 7[6], one or two peaks in each phrase are generally observed over the course of a sentence. However, noticeable differences can also be seen across the three FP conditions: i) focus preceded by

---

[6] As mentioned earlier, the stimuli were differentiated according to the number of words in order to see the effect of Sentence Length. However, we did not see any effect; thus it was excluded from the present study. Notice that the second area is at the location of the FPs. In the No FP condition, however, since it does not include an FP, the F0 contour cannot be seen in this area.

hangsang; ii) focus preceded by *ocik*; iii) focus without an FP (henceforth, the three different FP conditions). To illustrate, a large difference in F0 can be observed between the FPs, *hangsang* and *ocik*. The F0 contour of the former sharply rises, beginning with the release of the word. Afterwards, it drops sharply until it moves upward again. Conversely, *ocik* does not show a sharp increase. Rather, in the following phrase (i.e., the 3rd area), the F0 curve seems to reach the highest peak.

Upon closer examination, we see that the F0 peaks seem to become lower over the course of a sentence. This is known as F0 declination, which was first addressed by Cohen and 'tHart (1967). They observed the downward trend of F0 contours produced by many utterances. Cohen et al. (1982) report that F0 contours are gradually on the decline over the course of sentences. Also, F0 contours are known to fall rapidly at the end of sentences (Liberman et al. 1985, Umeda 1982). Further, longer utterances are known to have more gradual F0 declination trend than short sentences (Shih 2000).

Considering the F0 declination, it is important to remove the effect since the exact function of the FPs, as well as focus, can be masked. Otherwise, we will not be allowed to observe the exact intonational functions of languages. It is thus worth a try to eliminate the F0 declination and observe the F0 contours from various angles. A linear regression was thus conducted, by which a slope was measured by applying the formula in (8), to neutralize the F0 declination effect.

(8) $\hat{\beta} = (X'X)^{-1}X'y = \left(\frac{1}{n}\Sigma x_i x'_i\right)^{-1}\left(\frac{1}{n}\Sigma x_i y_i\right)$

The coefficients were tested at the 5 significance level. If not significant[7], residuals were calculated by treating the slope as 0. Afterwards, based on the residuals, the target sentences were reconstructed, as illustrated in Figure 8.

[7] Of 756 target sentences, 641 sentences experienced the F0 declination effect (approximately 84.8%).

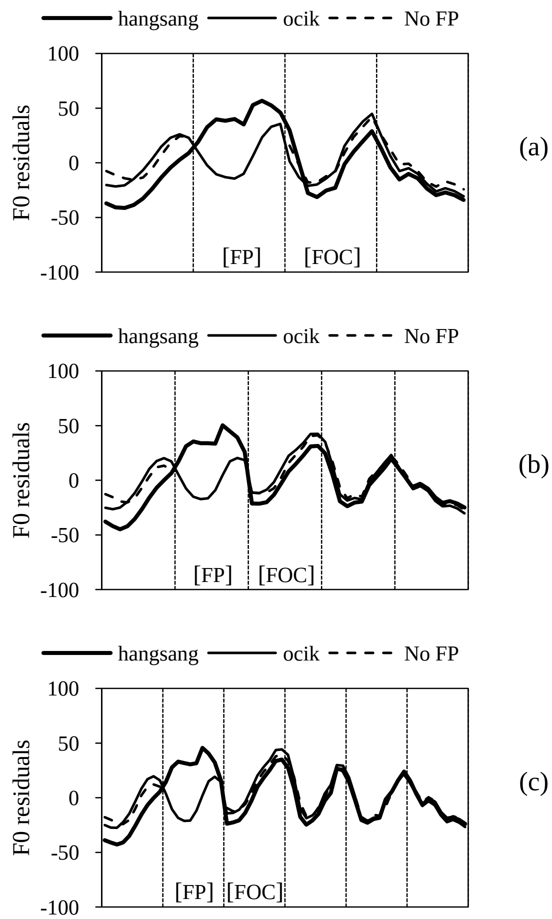


Figure 8. F0 residuals of all the sentences spoken by four speakers. Each F0 curve is an average of 24 repetitions: a) *Nanun* {*hangsang, ocik, EMPTY*} *mantwulul cohahapnita*; b) *Nanun* {*hangsang, ocik, EMPTY*} *nwunimuy mantwulul cohahapnita*; c) *Nanun* {*hangsang, ocik, EMPTY*} *nwunimi mantusin mantwulul cohahapnita*. (The items in braces indicate the word enclosed by the second and third areas of each figure)

In general, the contours in Figure 8 seem to be analogous to those in Figure 7, except the focused area. To be more specific, the focused elements in the three FP conditions seem virtually identical. However, in Figure 8, the focused element following *hangsang* seems to have lower F0 values, which presents evidence that the FPs may behave in a different way.

3.5. What to measure

Given the visual representations of the F0 contours in Figure 7, the present study measured duration, intensity, mean F0, and maximal F0 (henceforth, Max F0) across the three different FP conditions. The measurements were extracted from the 1st, and 2nd, and 3rd region. In addition, such measurements were used to examine the difference between *hangsang* and *ocik*. In the present study, the measurements were obtained directly from the Praat script (Xu 2005-2009) by using manually-labeled boundaries.

In Figure 7 and 8, we see that there are apparent differences between the FPs. For this reason, it is necessary to regard the FPs as an independent factor in order to see whether the F0 curve with *hangsang* is higher than that with *ocik*, regardless of the different contexts. In addition, the three different contexts (i.e., the context-free environment, the prompt question, the discourse context) are treated as an independent factor in order to find out the interaction effect between the contexts and the FPs. Furthermore, no apparent differences in F0 magnitude can be observed across the three conditions in the focused area (i.e., the 3rd area) in Figure 7, whereas differences are seen in Figure 8. Thus, the focused elements in the three different FP conditions are also regarded as an independent factor to examine whether the focused elements are influenced by the FPs or not.

4. Analyses and Results

4.1. On-focus area

Of the measurements, we begin with the on-focus area. As previously mentioned, it is commonly believed that *focus* is induced by the FP (Bayer 1996, Jacobs 1986, Kadmon 2001, Taglicht 1984). However, it is also argued that *focus* is elicited by the context (Büring & Hartmann 2001, Sudhoff 2010). This study focuses on five parts of the focused element (i.e., duration, mean intensity, mean F0, Max F0, mean F0 residuals) in the three different FP conditions to find out whether the FP is a focus inducer or a focus-sensitive operator.

In Figure 9, we see that there are no significant differences for all the measurements of the focused elements across the three different FP conditions except the mean F0 residuals, even though minor differences are observed. However, clear differences are observed for the mean F0 residuals.

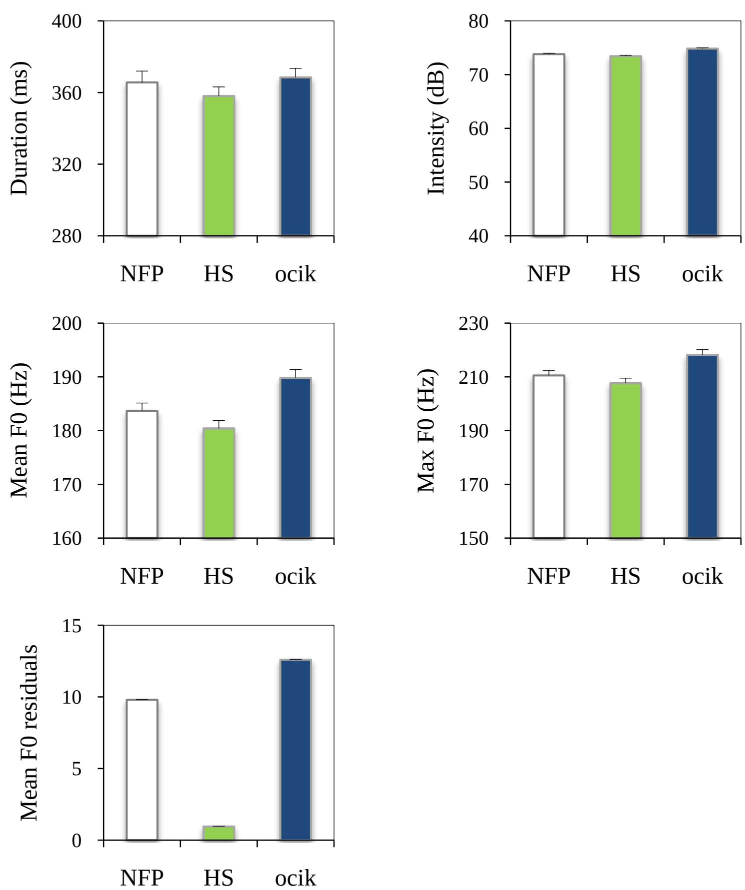


Figure 9. Means and standard errors (error bars) of duration, mean intensity, mean F0, max F0 and mean F0 residuals of the focused elements in the three different FP conditions: focus without an FP (NFP); focus preceded by *hangsang* (HS); focus preceded by *ocik*.

Table 2. Results of a one-way ANOVA for all measurements in the three FP conditions.

(df = 2)

	Focused Element
Duration	F = 0.7414, *p* = 0.4789
Mean Intensity	F = 0.5743, *p* = 0.5649
Mean F0	F = 0.2898, *p* = 0.749
Max F0	F = 0.2425, *p* = 0.7851
Mean F0 Residuals	F = 57.976, ***p* <.0001**

As can be seen in Table 2, there is no significant effect on duration: 365.7ms (NFP), 358.1 ms (*hangsang*), and 368.5 ms (*ocik*). The same is true for mean intensity: 73.817dB (NFP), 73.449dB (*hangsang*), and 74.85 dB (*ocik*). Second, both mean F0 and Max F0 are not also significantly different across the three different FP conditions: mean F0: (183.68 Hz (NFP), 180.4 Hz (*hangsang*), 189.8 Hz (*ocik*)); Max F0 (210.51 Hz (NFP), 207.69 Hz (*hangsang*), 218.18 Hz (*ocik*)). Conversely, a significant difference in mean F0 residuals is observed, which thus requires an *ad hoc* test. Table 3 shows the results of the *ad hoc* test.

Table 3. Results of a paired t-test across the three FP conditions

(df = 1079)

	NFP	hangsang
hangsang	T = -19.77 , ***p* < .0001**	
ocik	T = 6.464, ***p* < .0001**	T = -15.29, ***p* < .0001**

The results in Table 3 indicate that there are significant differences across the three different FP conditions. The residual values of *ocik* turn out to be the largest, followed by NFP and *hangsang*. Recall that no significant differences are observed in both mean F0 and Max F0 of the focused elements in the three different FP conditions. This is another issue to be raised, and we will examine this in the discussion section.

4.2. Focus particles

Through the visual inspections of Figure 7 and 8, we found that there are clear differences between the FPs: *hangsang* and *ocik*. In this section, we examine their prosodic differences through the measurements[8]: duration, mean intensity, mean F0, Max F0, and mean F0 residuals. Further, a two-way repeated ANOVA was performed to find out the interaction effect between Focus Particle and Context.

As is shown in Figure 10 and Table 4, significant differences can be found in all the measurements comparing the FPs, *hangsang* and *ocik*. That is, all the values in the region of *hangsang* are significantly higher than those with *ocik*. The duration values for *hangsang* and *ocik* are 357.9 vs. 292.2 ms. The intensity values are 72.2 vs. 69.1 dB. The mean F0 values are 219.6 vs. 176.7 Hz. The Max F0 values are 242.4 vs. 205.0 Hz. Lastly, the mean F0 residuals show clear differences between *hangsang* and *ocik* (25.5 vs. 4.90). On the other hand, significant interactions between the FPs and the three different contexts are not found, as shown in Table 4, even though there are significant differences between the FPs.

[8] The two FPs have different segments. Thus, the different productions of the FPs may be attributed the different segments. For simplicity, in the present study, we extracted two vowels of each F0 and then measured the average values.

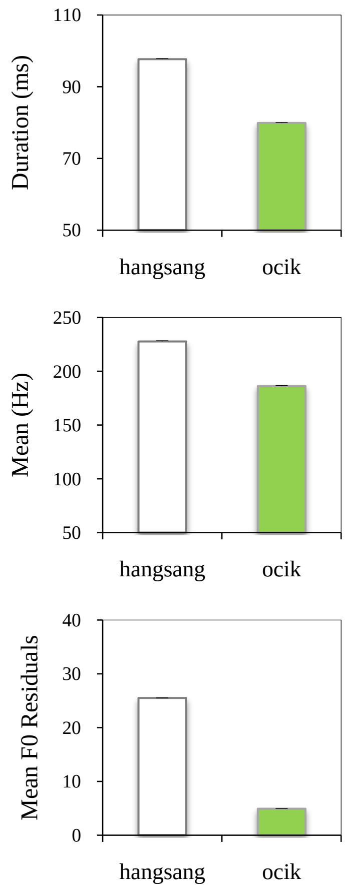


Figure 10. Means and standard errors (error bars) of duration, mean intensity, mean F0, max F0 and mean F0 residuals of the FPs *hangsang* and *ocik* in the three different contexts: the context-free environment, the prompt question, and the discourse context.

Table 4. Results of a 2-way repeated ANOVA for all measurements between the FPs *hangsang* and *ocik* in the three different contexts.

(df = 1, 2)

	FP (Focus particle)	FP * Context
Duration	F = 260.94, ***p* < .0001**	F = 0.8264, *p* = 0.438
Mean Intensity	F = 58.9729, ***p* < .0001**	F = 2.7640, *p* = 0.0638
Mean F0	F = 84.6126, ***p* < .0001**	F = 0.3772, *p* = 0.6859
Max F0	F = 71.1549, ***p* < .0001**	F = 0.2492, *p* = 0.7795
Mean F0 Residuals	F = 591.9, ***p* < .0001**	F = 0.3183, *p* = 0.7274

In this section, we found out that *hangsang* is more prosodically prominent than *ocik*. However, much is unknown about this area and the exact prosodic functions of the FPs. This will be analyzed in the discussion section.

4.3. F0 transitions

In Figure 7, we can see that the F0 curve of the sentence with *hangsang* bears the highest peak in the second region, whereas the F0 contour of the sentence with *ocik* has its highest peak in the third region. What is the crucial difference between them? To see the effect, we also measured the mean F0 of the three areas including *hangsang* and *ocik*. Figure 11 shows the transition of the F0 curves from the first to the third area, where we see that the F0 curve with *hangsang* undergoes a sharp transition (F[1] = 2690.8, $p < 0.0001$) from the first to the second area and then reaches the highest peak. Afterwards, it experiences a large drop (F[1] = 2609.4, $p < 0.0001$). Conversely, the F0 contour with *ocik* shows a smooth transition over the three areas. However, there is a slight, albeit significant, F0 decrease (F[1] = 7635.2, $p < 0.001$)

Then, the F0 begins to move upward following the second area ($F[1] = 3798$, $p < 0.0001$).

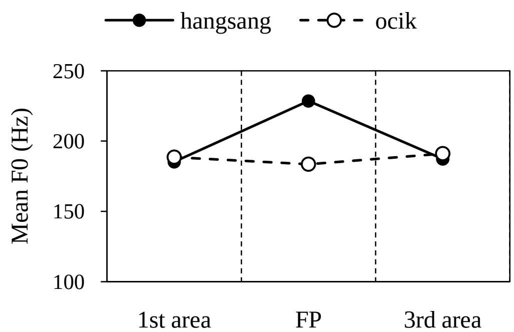


Figure 11. Differences in mean F0 between the FPs *hangsang* and *ocik* in the three regions around the FP.

However, the F0 transitions become slightly different after removing the F0 declination effect, as illustrated in Figure 12, though generally similar to Figure 11. To illustrate, the F0 curve of the sentence with *hangsang* also reaches highest peak in the second region; whereas, that of *ocik* holds its highest peak in the third region. According to the measurements of mean F0 residuals of the three areas, the F0 curve with *hangsang* experiences a steep transition ($F[1] = 1342.3$, $p < 0.0001$) from the first to the second area. Afterwards, it shows a sharp decline ($F[1] = 120.88$, $p < 0.0001$). In contrast, the F0 contour with *ocik* shows a gradual change over the three areas. For instance, there is an increase from the first to the second area, but it turns out to be statistically insignificant ($F[1] = 1.6864$, $p = 0.195$). Yet, the F0 undergoes a significant increase following the second area ($F[1] = 98.33$, $p < 0.0001$).

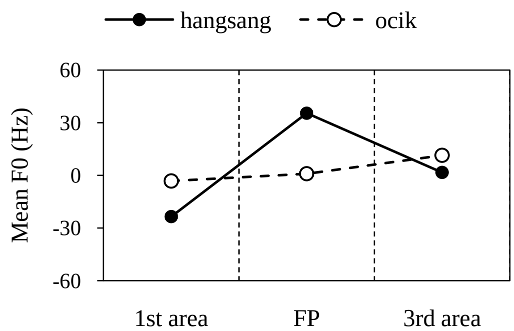


Figure 12. Differences in mean F0 residuals between the FPs *hangsang* and *ocik* in the three regions around the FP.

5. General Discussion

The analyses based on the measurements (i.e., duration, mean intensity, mean F0, Max F0, mean F0 residuals) and statistical methods played important roles in answering the main question: Is the FP a focus inducer or a focus-sensitive operator? In addition, we found that there is a clear difference between *hangsang* and *ocik*. Thus, we need to examine why they behave differently. In this section, we first focus on the key issues in the present study. Second, we describe the F0 contour of Korean, using the PENTA model. Lastly, we talk about the unresolved issues and future directions.

5.1. Focus-sensitive operator vs. focus inducer

In Figures 7 and Table 2, no prosodic differences can be observed between the three different FP conditions, where the focused items have virtually identical duration, mean intensity, mean F0, and Max F0. This finding seems to support the claim that focus is not determined by the FPs, but rather elicited by the context (Sudhoff 2010). If the FPs serve as focus inducers, the focused items without FPs should not have values that are prosodically equivalent to the ones observed with FPs.

In the present study, we performed a linear regression in order to eliminate the F0 declination effect. Then, we found that there are prosodic differences across the three FP conditions in Figure 8. First, the F0 contour of the focused element with *hangsang* turns out to be the lowest (i.e., 3rd area). This may be largely due to post-focus compression[9] (henceforth, PFC) (Y-c Lee & Xu 2010, Chen et al. 2009, W Gu & T Lee 2007), which is known to compress the F0 contour after *focus*. Languages, such as English, Mandarin, and Korean exhibit the effect whereas Cantonese and Taiwanese do not. (Chen et al. 2009, W Gu & T Lee 2007). Second, the F0 contour with *ocik* has the largest values in the focused region, which are even higher than that of No FP. Nevertheless, all things considered, it is believed that *focus* is elicited by the context because there are no significant differences in duration, mean intensity, mean F0, and Max F0, as in Table 2. Therefore, FPs need to be treated as focus-sensitive operators, rather than focus inducers.

On the other hand, an alternate proposal may arise from a different point of view. Since clear differences between *hangsang* and *ocik* were observed in the present study, we argue that the FPs should be divided into more fine-grained categories in terms of their prosodic functions.

5.2 Difference between *hangsang* and *ocik*

In Figure 7 and 8, we can see that there are differences between the FPs, *hangsang* and *ocik*. The former is realized with a longer duration, a higher pitch, a greater intensity and a larger F0 residual value than its counterpart. This provides a new perspective on FPs. Given that *hangsang* shows more dynamic, rising F0 contours than the subsequent focused item, it can be assumed that *hangsang* receives a focus feature whatever the context is (see Table 4). However, the F0 contour of *ocik* does not exhibit such prosodically prominent features.

In Figure 11, in the case of *ocik*, the F0 decrease can be found between the first and the second region. This does not imply that *ocik* is prosodically unmarked. Rather, it is likely to be influenced by F0

[9] In the case of mean F0 and Max F0, there were no significant differences of the focused elements in the three different FP conditions. We believe that the F0 contours would be masked by F0 declination.

declination, which is the tendency to lower the F0 contour at the end of sentences (Bolinger 1978, Cohen et al. 1982, among others). Thus, after removing the F0 declination, the F0 increase is found between the two regions though statistically insignificant. However, the focused item in the third area receives a focus feature, given that the F0 contour rebounds and begins to move upward.

In the case of *hangsang*, dynamic F0 transitions are observed. An exponential increase in F0 can be seen between the first and the second area. Then, the F0 decreases sharply. As is aforementioned, Korean is found to undergo PFC after *focus* (Y-c Lee & Xu 2010). However, *hangsang* does not completely delete the function of the following focused item[10] even though *hangsang* has lower mean F0 residual values than *only* and No FP. Hence, this finding supports the PENTA model (Xu 2004, 2005, 2009, Xu & Xu 2005), which is based on *function*, not *form*. As mentioned earlier, the model denotes the mechanisms of multiple intonational functions simultaneously conveyed through speech melody (Xu 2004, 2005, 2009, Xu & Xu 2005). Thus, it may be the case that two (or more) intonational functions can be simultaneously realized in one sentence.

Let us now deal with why *hangsang* and *ocik* behave differently. Beaver and Clark (2008) propose a hybrid theory of semantics and pragmatics called the Quasi/Free/Conventional theory. They claim that associations of *always* and *only* with focus are formed differently and have distinct restrictions with respect to leaners (or reduced pronouns) and extraction. In the theory, the function of *always* is categorized as free association, constructing an association with contextually salient sets of events or situations. However, *only* functions as conventional association, which constructs an association based on a lexically-encoded dependency on focus. This implies that prosodic salience of a focused element is requisite for *only* to create an association, but it is not for *always*.

Korean FPs *hangsang* and *ocik* behave in the same way as English in (9). Instead of leaners, we investigated the phenomenon using pro because reduced pronouns are not available in Korean. As the

[10] The reason is that there are no significant differences in duration, mean intensity, mean F0, and Max F0 although those values of *hangsang* are slightly smaller than those of *ocik* and No FP due to PFC.

sentence with pro show, *hangsang* is able to create an association with a pro, as opposed to *ocik*. This result is compatible with Beaver and Clark (2008)'s theory and their cross-linguistic observations, using German and other languages.

(9) Association with pro (Beaver and Clark 2003: 343)

Context: You had many discussions with Sandy, but what I want to know is the extent to which you talked about Fred. Of all the times you talked with Sandy, how often was Fred the person you talked about?

Na-nun hangsang/#ocik Sandy-wa pro tayhay tholon-ha-yess-ta. (Nambu & Y-c Lee 2010)

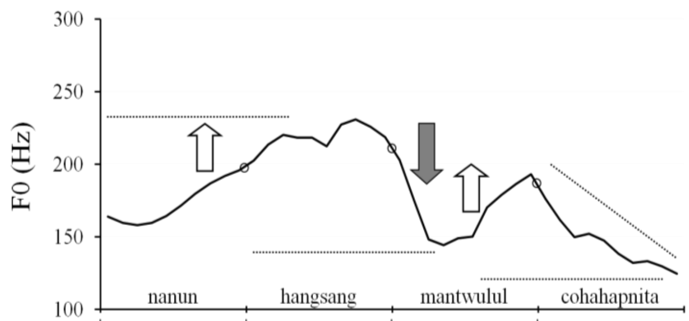
I-Top always/only Sandy-with about discuss-do-Pst-Decl

'I always/#only discussed (him) with Sandy.'

5.3. Applying PENTA to Korean

As discussed in 2.6, the present study adopted a PENTA system, rather than using a ToBI system. The Parallel Encoding and Target Approximation (PENTA) model shows communicative components of speech melody in terms of *function*, not *form*. This model points out the problems in the British nuclear tone tradition and the AM theory. Both systems depict intonation primarily in form. In this section, we will use the PENTA model to describe the pitch contour of Korean.

Let us first look at the actual transmission of the sentence, *nanun hangsang mantwulul cohahapnita*, applied to the PENTA model. In this sentence, *manturɨl* receives a focus. In the current analysis, *lexical stress*, *focus*, and *duration* were used as independent functions in order to deal with describe the target sentence. See Figure 13.



Lexical stress +	[mid]	[high]	[high]	[fall]
Focus	**neutral**	**wide**	**low + wide**	**narrow + low**
Duration	**normal**	**long**	**long**	**short**

Figure 13. The description of the F0 contours of the sentence, *nanun hangsang mantwulul cohahapnita* based on the PENTA model. Short straight dotted lines denote hypothetical local pitch targets. The unfilled block arrow shows on-focus pitch expansion. The black arrow means post-focus pitch lowering and narrowing. The table under the graph refers to symbolic analysis of the PENTA model.

As can be seen in Figure 13, the sentence is depicted under the PENTA model. The FP, *hangsang* experienced a pitch expansion. The item following the FP underwent PFC, and then it began to rebound. Then, the focused item experienced a huge drop after reaching a peak.

Second, let us examine *lexical stress*. The pitch contour of *nanun* is assigned as [mid]. The final item of the sentence is assigned as [fall]. As you probably guess, this is evident because the last item of a declarative sentence has a falling tone. The PF, *hangsang* receives a [high] lexical stress. To go into details, the high tone starts from the onset of the item. The F0 reaches a peak around the second syllable, *–sang*. The focused element, *mantwulul* is assigned as [high], in which case the last syllable *–lul* holds the highest peak.

Third, let us discuss *focus*. In the present study, the FP, *hangsang* was realized with a longer duration, higher pitch and wider amplitude than its counterpart, *ocik*. Thus, it is assigned as wide. A focus also falls on the third item of the given sentence; thus, *mantwulul* was directly influenced by the focus effect. But, because it first underwent PFC and began to rise again, the focus is characterized as both low and wide. In the sentence, *nanun* is assigned as neutral, whereas *cohahapnita* is treated as both narrow and low (i.e., narrow + low); this is because the pitch contour following a contrastive focus is compressed and narrowed.

Next, let us inspect *duration*. The pre-focus item, *nanun* is assigned as normal. However, the post-focus item, *cohahapnita* is assigned as short because the duration under the post-focus condition was found to be short due to PFC. Both *hangsang* and *mantwulul are* assigned as long because they have a focus feature.

Lastly, let us consider why the PENTA model should replace the AM theory. In the AM theory, both *hangsang* and *mantwulul* in Figure 13 will be treated as having an L+Ha, where 'a' indicates an underspecified AP-final tone (S-A Jun 2000). In the AM theory, there are no explicit mechanistic explanations of F0 peaks and valleys (see Xu & Xu 2005). Following *focus*, the post-focus area is compressed by PFC. How can the notation, 'a' describe the function? Further, the focused element, *mantwulul* shows two intonational functions at the same time, as is shown in Figure 13. It underwent PFC in the beginning. Around the second syllable, the F0 began to rebound and move upward due to the focus feature. However, the AM theory cannot explain the multiple functions in parallel. Thus, we believe that we can provide a more detailed explanation of F0 contours under the PENTA model. Further, The PENTA model is suitable to describe the multiple functions that the sentence carries; both *hangsang* and *mantwulul* are prosodically salient. Yet, *ocik* does not show a prominent intonational feature in the present study. As Beaver and Clark (2008) point out, their semantic/pragmatic functions between *hangsang* and *ocik* are different, and the prosodic difference between them in our experiment reflects the different functions.

5.4. Unresolved issues

Thus far, we have examined whether the FPs serve as focus inducers of focus-sensitive operators. In addition, the prosodic differences between *hangsang* and *ocik* have been addressed. However, we concede that the present study has several weaknesses. First, more systematic and fluent data are needed to better understand the exact functions of the FPs. Second, a difference was observed between mean F0 and mean F0 residuals of the focused elements in the three different FP conditions. The difference should also be resolved to be more aware of the F0 contours in the future research.

Considering the differences between *hangsang* and *ocik*, it is true that the different productions may be due to segmental effects. The different phonetic environments of the two FPs would give rise to dissimilar durations, intensities, and F0 contours. Also, the differences may be caused by intrinsic F0 of vowels. Whalen and Leviit (1995) report that high vowels tend to have higher F0 values than lower ones. Thus, the results would be more convincing if the normalized method of the differences would be employed in the analysis.

6. Conclusion

The findings in the present study reveal that focus can be elicited by the context. Accordingly, FPs need to be treated as focus-sensitive operators, rather than focus inducers. Yet, the FPs need to be classified by different intonation functions, given that clear differences between *hangsang* and *ocik* were observed. The former has longer duration, higher pitch, and greater intensity values. Thus, the FP, *hangsang* seems to carry a focus feature. Even though it is realized with longer duration, higher pitch, and greater intensity, it does not completely eliminate the function of the subsequent word whose *focus* is elicited by the context. This demonstrates that speech must be described in *function*, not in *form* (Xu 2004, 2005, 2009, Xu & Xu 2005), since it can bear (possibly) two or more intonational functions. Lastly, the results of the present study call for a cross-linguistic study regarding the FPs. Thus, it would be of interest to examine cross-linguistic typological patterns in order to understand the nature of FPs: i) whether FPs serve as focus-sensitive operators/focus inducers or ii) whether there are phonetic differences between FPs.

7. Reference

Avesani, Cinzia. (1987). Declination and sentence intonation in Italian. In Pier M. Bertinetto, ed, Quaderni del Laboratorio di Linguistica della Scuola Normale Superiore di Pisa 1, 8-23.

Bayer, Josef. (1996). Directionality and Logical Form. On the Scope of Focusing Particles and Wh-in-situ. Kluwer Academic Publishers.

Beaver, David and Brady Clark. (2003). Always and only: why not all focus-sensitive operators are alike. Natural Language Semantics 11, 323-362.

Beaver, David and Brady Clark. (2008). Sense and Sensitivity: How focus determines meaning. Oxford: Wiley-Blackwell.

Beckman, Mary and Janet Pierrehumbert. (1986). Intonational structure in Japanese and English, Phonology Yearbook 3, 15-70.

Boersma, Paul. (2001). Praat, a system for doing phonetics by computer. Glot International 5.9/10, 341-345.

Bolinger, Dwight. (1978). Intonation Across Languages. In Joseph H. Greenberg, ed., Universals of Human Language 2, 471-524 Phonology, Stanford University Press.

Büring, Daniel and Katharina Hartmann. (2001). The syntax and semantics of focus-sensitive particles in German. Natural Language & Linguistic Theory 19, 229–281.

Chen Szu-wei, Bei Wang, and Yi Xu. (2009). Closely related languages, different ways of realizing focus. Proceedings of Interspeech 2009, 1007-1010.

Chen, Yiya. (2006). Durational adjustment under corrective focus in standard Chinese. Journal of Phonetics 34.2, 176-201.

Cohen. Antonie., Coller, René, and ‘t Hart, Johan. (1982). Declination: construct or institute feature of speech pitch?. Phonetica 39, 254-273.

Cohen, Antonie. and Johan. 't Hart. (1967). On the anatomy of intonation. Lingua, 19: 177-92

Crystal, David. (1969). Prosodic Systems and Intonation in English. Cambridge University Press.

Gårding, Eva. (1987). Speech act and tonal pattern in standard Chinese. Phonetica 44, 13-29.

Gu, Wentao and Tan Lee. (2007). Effects of focus on prosody of Cantonese speech – A comparison of surface feature analysis and model-based analysis. In Proceedings of the International Workshop Paralinguistic Speech 2007, 59-64.

Jacobs, Joachim. (1986). The syntax of focus and adverbials in German. In Werner Abraham and Sjaak De Meij, eds., Topic, Focus, and Configurationality, 103-127. John Benjamins.

Jin, Shunde. (1996). An Acoustic Study of Sentence Stress in Mandarin Chinese. Ph.D. Dissertation. The Ohio State University.

Jun, Sun-Ah. (2000). K-ToBI (Korean ToBI) labelling conventions: Version 3. Speech Sciences 7, 143-169.

Kadmon, Nirit. (2001). Formal Pragmatics: Semantics, Pragmatics, Presupposition, and Focus. Wiley-Blackwell.

Ladd, Robert D. (1996). Intonational Phonology. Cambridge University Press.

Lieberman, Philip, William Katz, Allard Jongman, Roger Zimmerman, and Mark Miller. (1985). Measures of the sentence intonation of read and spontaneous speech in American English. Journal of the Acoustical Society of America 77, 649-657.

Lee, Yong-cheol and Yi Xu. (2010). Phonetic realization of contrastive focus in Korean. Speech Prosody 2010, 100033:1-4.

Nambu, Satoshi and Yong-cheol Lee. (2010). Phonetic realization of focus particles *always* and *only* in Korean: Theoretical implications of association with focus. ms. University of Pennsylvania.

O'Connor, Joseph D. and Gordon F. Arnold. (1961). Intonation of Colloquial English. Longmans.

Oh, S.P. (2007). A Study on the Intonational System of English and its Representation. Ph.D. Dissertation. Korea University.

Pierrehumbert, Janet and Julia Hirschberg. (1990). The Meaning of Intonation contours in the interpretation of discourse. In Philip R. Cohen, Jerry Morgan and Martha E. Pollack, eds., Intonations in Communication, 271-311. MIT Press.

Pierrehumbert, Janet. (1980). The Phonology and Phonetics of English Intonation. Ph.D. Dissertation. MIT.

Shih, Chilin. (1988). Tone and intonation in Mandarin. Working Papers, Cornell Phonetics Laboratory 3, 83-109.

Shih, Chilin. (2000). A declination model of Mandarin Chinese. In Anotonis Botinis, ed., Intonation: Analysis, Modelling and Technology, 243-268. Kluwer Academic Publishers,.

Sudhoff, Stefan. (2010). Focus particles and contrast in German. Lingua 120, 1458-1475.

Taglicht, Josef. (1984). Message and Emphasis. On Focus and Scope in English. Longman.

Umeda, Noriko. 1982. “F0 declination” is situation dependent. Journal of Phonetics 10, 279-290.

Whalen, Douglas H. and Andrea G. Levitt. (1995). The universality of intrinsic F0 of vowels. Journal of Phonetics 13, 349-366.

Xu, Yi. (2004). The PENTA model of speech melody: Transmitting multiple communicative functions in parallel. In Proceedings of From Sound to Sense: 50+ years of discoveries in speech communication, 91-96.

Xu, Yi. (2005). Speech melody as articulatorily implemented communicative functions. Speech Communication 46, 220-251.

Xu, Yi. (2009). Timing and coordination in tone and intonation -- An articulatory-functional perspective, Lingua 119, 906-927.

Xu, Yi. (2005-2009). _TimeNormalizeF0.praat. Online: http://www.phon.ucl.ac.uk/home/yi/tools.html.

Xu, Yi and Ching X. Xu. (2005). Phonetic realization of focus in English declarative intonation. Journal of Phonetics 33, 159-197.