Phonetic Realization of Contrastive Focus in Korean

Yong-cheol Lee[1], *Yi Xu*[2]

[1] Department of Linguistics, University of Pennsylvania, USA
[2] Department of Speech, Hearing and Phonetic Sciences, University College London, UK
leyong@ling.upenn.edu, yi@phon.ucl.ac.uk

Abstract

It has been observed that a focused item in Korean is by and large realized with higher pitch, longer duration and greater intensity. But it is not fully clear how much focus affects the pre-focus and post-focus regions. The present study examines the full scale of focus effect in Korean in pre-focus, on-focus and post-focus words. Results show that focused words have significantly increased F_0, duration and intensity, post-focus words have significantly reduced F_0, duration and intensity, but pre-focus words lack systematic changes in any of these parameters. Thus Korean seems to resemble languages like English and Mandarin that exhibit post-focus compression (PFC), but differ from Taiwanese and Cantonese where PFC has been found to be absent.

Index Terms: pre-focus, on-focus, post-focus, duration, F_0, intensity, post-focus compression, PFC

1. Introduction

Crosslinguistically, a focused item is known to have higher pitch, longer duration and greater amplitude than its unfocused counterpart. Furthermore, the constituents outside the focus domain are also known to be affected by focus. The pre- and post-focus constituents are believed to experience F_0 and durational changes [5, 9, 13, 19]. In the case of edge-prominence languages (e.g., Korean, Japanese, Bengali, French), the areas preceding and following focus are often said to be dephrased [12, 13]. In head-prominence languages (e.g., English, German, Dutch), however, the pre- and post-focus areas are said to be deaccented [1, 11, 14]. Whatever the terminology, the pre- and the post-focus regions are believed to be affected by focus. But there has not been agreement about the exact nature of the focus effect in those regions. Different changes have been reported for different languages, or even for the same languages [6, 8, 10, 15].

An important source of the discrepancy is likely to be in the methodology. In many cases, no direct comparisons are made between the same word or phrase in different focus conditions, e.g., when it is pre-focus or post-focus and when it is neutral-focus. In cases where such direct comparisons are made, it is found that some languages, e.g. English and Mandarin, consistently compress the pitch range of post-focus constituents, hence, post-focus compression (PFC), but make no systematic changes in the pre-focus constituents [4, 16, 19]. In other languages, such as Taiwanese and Cantonese, no PFC is found when such direct comparisons are made [3, 7].

The objective of the present study was to investigate how focus is phonetically realized in Korean, for which previous research has generated mixed results [8, 9, 10, 15]. This was done through systematic comparison of F_0, duration and intensity in different focus conditions. In particular, the constituents before and after focus were closely examined to see if there is clear evidence of PFC and whether there is any systematic pre-focus effects.

2. Methodology

2.1. Stimuli

Four sets of sentences served as stimuli. Among them, the first two sets were used as control data and the other two as test data. Forty sentences were used: 16 target sentences and 24 fillers. The 16 target sentences were divided into two sets, which were differentiated according to focus location and the number of syllables before and after focus. The number of syllables ranged from 3 to 13 syllables in the pre- and post-focus region.

To elicit contrastive focus on a specific word, each target sentence was preceded by a prompt question, as has been successfully done before [4, 16]. The target sentences and the prompt questions are shown in Table 1.

Table 1. *Left column: Prompt sentences. Right column: target sentences. The focused constituents are in boldface.*

Prompt	Target
Minsuka **muəsɨl** məknɨntako? ‘What is Minsu eating?’	minsuka **mantuɾɨl** məknɨnta ‘Minsu is eating potstickers’
Minsuka ʧənjəke **muəsɨl** məknɨntako? ‘What is Minsu eating in the evening?’	Minsuka ʧənjəke **mantuɾɨl** məknɨnta ‘Minsu is eating potstickers in the evening’
Minsuka nonsanesə ʧənjəke **muəsɨl** məknɨntako? ‘What is Minsu eating in the evening in Nonsan?’	Minsuka nonsanesə ʧənjəke **mantuɾɨl** məknɨnta ‘Minsu is eating potstickers in the evening in Nonsan’

2.2. Subjects

Six native speakers of Korean participated in the experiment. Three of them were males and the other three females, ranging in age from 22-31 (mean = 28.1 years). They were recruited from the Chungnam National University and were paid for their participation. No participants reported any problems with their speech or hearing.

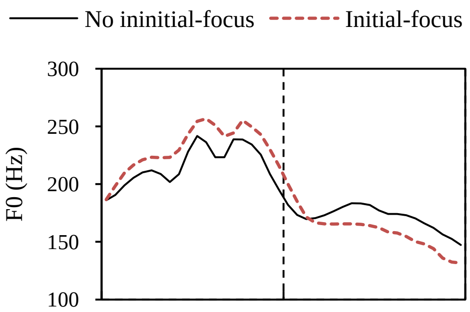
2.3. Procedure

The subjects were recorded in a sound-attenuated booth in the Department of Linguistics at Chungnam National University. They were seated in front of a computer monitor. A microphone was placed to the left of the monitor. All sentences were recorded electronically and saved on a computer as *wav* files, using Praat [2]. Before each recording session, the materials were given to subjects to allow their readings to be as natural as possible. Subjects were asked not to pause during each trial. They were also asked to repeat the sentences if they made a mistake or if we found their reading unnatural. The reading materials were presented on a paper sheet in random order.

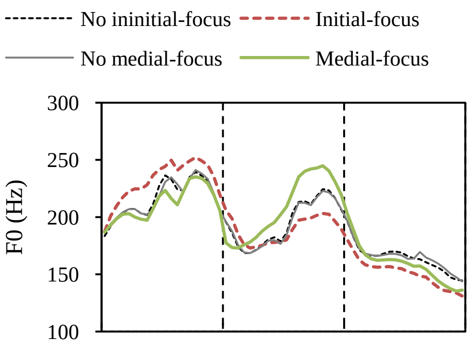
2.4. F_0 extraction

A Praat script was used to extract F_0 contours of the target sentences [18]. The syllable boundaries were hand-labeled. The script also generated various measurements to be used in the data analysis. After the process of F_0 extraction, all the target sentences were converted to graphs to make it easy to spot inconsistencies. Each F_0 curve was inspected in order to exclude apparent outlier sentences. Figure 1 displays mean time-normalized F_0 curves of all the target sentences (excluding the outliers) averaged over all tokens produced by all six speakers.

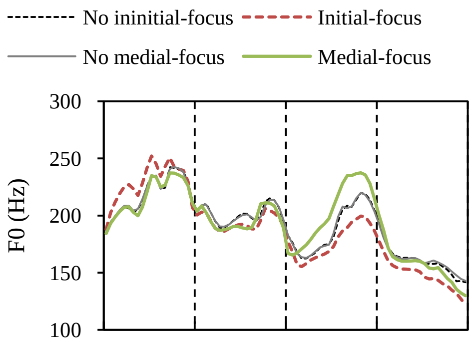
From Figure 1 we can see that each phrase (except the sentence-final one), as bordered by the vertical lines, has a rising-falling F_0 contour with one or two peaks, regardless of the focus condition. At the same time, visible differences can be seen across the focus conditions. Compared to the neutral-focus contour, the on-focus contour rises more sharply toward a higher peak (or two peaks) and then drops more quickly afterwards. After the focused word F_0 continues to drop until it reaches a lower level than that of the neutral-focus contour. Such differences can be potentially captured by comparing either maximum F_0 (henceforth Max F_0) or mean F_0 in each phrase. Figure 1 also shows that there is an overall trend for the F_0 peaks to become lower over the course of a sentence. Thus there is a need to treat focus location and sentence length as independent factors.



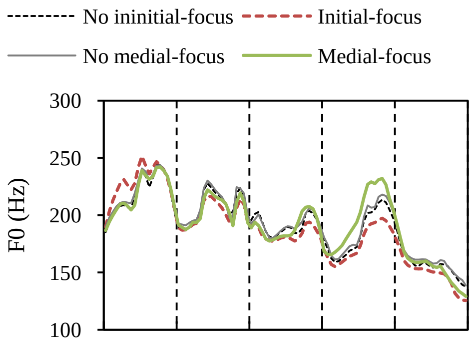
a)



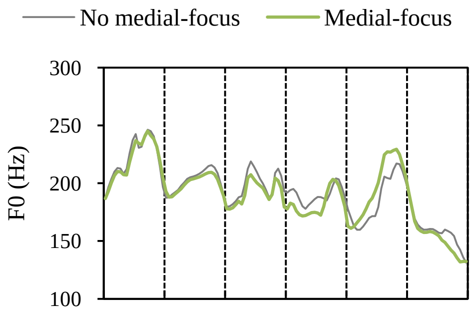
b)



c)



d)



e)

Figure 1: *Mean F_0 curves of all the sentences spoken by six speakers (averaged in Hz): a) minsu-ka məknɨnta; b) minsu-ka mantu-ɾɨl məknɨnta; c) minsu-ka ʧənjək-e mantu-ɾɨl məknɨnta; d) minsu-ka nonsan-esə ʧənjək-e mantu-ɾɨl məknɨnta; and e) minsu-ka nuna-wa nonsan-esə ʧənjək-e mantu-ɾɨl məknɨnta.*

2.5. F_0 Measurements

Based on visual inspection of plots like those in Figure 1, the following measurements are taken from pre-focus, on-focus and post-focus areas: mean F_0, maximum F_0, duration and intensity of the target words. The measurements were automatically generated by the Praat script based on the hand-labeled syllable boundaries.

2.6. Analyses and Results

The basic analysis strategy is to make comparisons between focused and neutral-focus sentences in three separate regions: on-focus, pre-focus and post-focus. The dependent variables are mean F_0, max F_0, duration and mean intensity in each region. There are three independent variables: focus (focus, neutral), focus location (early, late) and sentence length (2-3 words, 3-4 words, 4-5 words, 5-6 words).

The means and standard errors of the four measurements as a function of focus in the pre-focus, on-focus and post-focus regions are displayed in Figure 2. The results of 3-way repeated measures ANOVAs (or 2-way ANOVA for pre-focus regions) on mean F_0, max F_0, duration and mean intensity are shown in Table 2.

For both mean F_0 and max F_0 (Figures 2a, b), on-focus words are significantly higher than their neutral-focus counterparts, and post-focus words are significantly lower than neutral-focus words. However, there are no significant differences between pre-focus and neutral focus words. For duration (Figure 2c), on-focus words are significantly longer than the same words in the neutral focus condition, while post-focus words are significantly shorter than their neutral-focus counterparts. The difference in the pre-focus region is again not significant. A similar tendency is also seen in mean intensity (Figure 2d). Focused words have greater intensity than neutral-focus words in the same region, while post-focus words have lower intensity than their neutral-focus counterparts. In contrast, there is no significant focus effect in the pre-focus region.

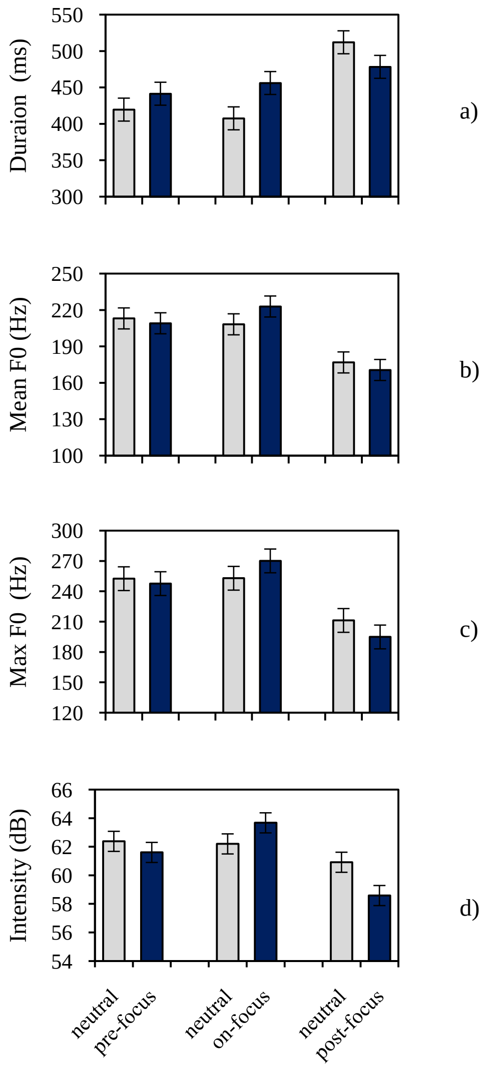


Figure 2: *Means and standard errors (error bars) of duration, mean F_0, max F_0 and mean intensity in pre-focus, on-focus, and post-focus areas and corresponding neutral-focus areas.*

Table 2. *Results of 2-way and 3-way repeated measures ANOVAs for all measurements in pre-focus, focus, and post-focus regions.*

	Focus (df = 1, 5)		
	Pre-focus	On-focus	Post-focus
Mean F_0	F = 1.758, p = 0.2422	F = 10.35, $\boldsymbol{p}$ **= 0.0235**	F = 9.542, $\boldsymbol{p}$ **= 0.0272**
Max F_0	F = 0.761, p = 0.4229	F = 8.167, $\boldsymbol{p}$ **= 0.0355**	F = 21.409, $\boldsymbol{p}$ **= 0.0057**
Duration	F = 2.256, p = 0.1934	F = 11.45, $\boldsymbol{p}$ **= 0.0196**	F = 7.078, $\boldsymbol{p}$ **= 0.0449**
Mean Intensity	F = 4.133, p = 0.0978	F = 8.699, $\boldsymbol{p}$ **= 0.0319**	F = 44.27, $\boldsymbol{p}$ **= 0.0012**
	Focus location (df = 1, 5)		
		On-focus	Post-focus
Mean F_0		F = 45.125, $\boldsymbol{p}$ **= 0.0011**	F = 76.216, $\boldsymbol{p}$ **= 0.0003**
Max F_0		F = 28.08, $\boldsymbol{p}$ **= 0.0032**	F = 94.725, $\boldsymbol{p}$ **= 0.0002**
Duration		F = 3.045, p = 0.1414	F = 31.295, $\boldsymbol{p}$ **= 0.0025**
Mean Intensity		F = 12.392, $\boldsymbol{p}$ **= 0.0169**	F = 18.464, $\boldsymbol{p}$ **= 0.0077**
	Sentence length (df = 3, 15)		
	Pre-focus	On-focus	Post-focus
Max F_0	F = 7.611, $\boldsymbol{p}$ **= 0.0025**	F = 1.564, p = 0.2396	F = 23.06, $\boldsymbol{p}$ **< .0001**
Mean F_0	F = 24.14, $\boldsymbol{p}$ **< .0001**	F = *11.52*, $\boldsymbol{p}$ **= 0.0004**	F = 1.58, p = 0.2357
Duration	F = 38.638, $\boldsymbol{p}$ **< .0001**	F = 2.778, p = 0.0773	F = 36.535, $\boldsymbol{p}$ **< .0001**
Mean Intensity	F = 9.726, $\boldsymbol{p}$ **= 0.0008**	F = 18.52, $\boldsymbol{p}$ **=< .0001**	F = 6.654, $\boldsymbol{p}$ **= 0.0045**

The middle panel of Table 2 shows that the effect of focus location is highly significant for most measurements. In the on-focus region all the measurements have larger values in early focus than in later focus: mean F_0: 229.2 vs. 201.8 Hz, max F_0: 261.6 vs. 235.7 Hz, duration: 438.9 vs. 424.7 ms (but not significant), and mean intensity: 63.7 vs. 62.1 dB. The same is true in the post-focus region with the exception of duration, mean F_0 (180.4 vs. 167.0 Hz), max F_0 (203.1 vs. 166.7 Hz) and mean intensity (60.1 vs. 59.4 dB). For duration, post-focus words are shorter in early focus than in late focus: 475.1 vs. 515.2 ms. There are only two moderately significant interactions between focus location and focus: on-focus duration ($F[1,5] = 7.919$, p = 0.0374) and post-focus max F_0 ($F[1,5] = 7.121$, p = 0.0444).

The effect of sentence length is highly significant for most of the measurements. However, the directions of the differences are rather mixed and their interpretations are not straightforward as many of them seem to be related to the structural differences that happen to co-occur across the sentences. There are a number of significant interactions between sentence length and focus for duration in all regions (pre-focus: $F[3,15] = 18.14$, $p <$ 0.001, on-focus: $F[3,15] = 14.15$, $p < 0.0001$, post-focus: $F[3,15] = 7.916$, $p = 0.0021$) and mean intensity in the pre-focus region ($F[3,15] = 5.435$, $p = 0.0099$). Again, the directions of the interactions are mixed.

3. Discussion

In both Figures 1 and 2 we can clearly see the F_0 differences due to contrastive focus in Korean. And the differences in the on-focus and post-focus regions are all significant as shown in Table 2. The focused words were generally realized with higher pitch than the same words in the neutral-focus condition. Moreover, the focused words had longer duration and greater intensity than their neutral-focus counterparts. Words in post-focus regions were realized with lower pitch, shorter duration, and weaker intensity than the same words in the neutral post-focus condition. But no significant differences in any measurement were found in the pre-focus words. Thus Korean seems to exhibit similar focus effects as in English and Mandarin when direct comparisons are made between sentences that differ only in focus conditions, i.e., a pattern of three-zone pitch range adjustments: expansion under focus, compression after focus, and little or no change before focus [16, 19]. But additionally, duration and intensity also show similar patterns.

From Figure 1 we can also see that, despite the significant differences in Table 2, the overall F_0 contours of the same sentences in different focus conditions appear very similar to each other. In fact the overall focus-related differences are smaller in magnitude than those reported for Mandarin [16]. Thus there is no clear evidence of major changes in intonational structure due to focus. This casts doubts on previous proposals that in Korean focus involves dephrasing [12, 13].

Finally, the present data seem to group Korean with languages like English and Mandarin where prosodic focus is realized with PFC [16, 19], rather than with languages like Taiwanese and Cantonese where PFC is found to be absent when a similar methodology as used in the present study is applied [3]. It is therefore possible to find out for other languages whether PFC is present with similar methodologies.

4. Conclusion

The findings of the present study demonstrate that in Korean, focused words have longer duration, higher pitch, and greater intensity than the corresponding neutral-focus words. In addition, post-focus words have lower pitch, shorter duration and lower intensity than the corresponding neutral-focus words. In contrast, there are no consistent focus effects on the pre-focus words. Therefore, Korean seems to exhibit the same tri-zone adjustments for focus as seen in English and Mandarin: on-focus expansion, post-focus compression and no pre-focus adjustment [16, 19]. More importantly, since on-focus expansion of F_0, duration and intensity is quite widely observed [17] (in fact most studies have concentrated *only* on on-focus changes), Korean, like English and Mandarin, seems to contrast with languages like Taiwanese and Cantonese only in terms of presence/absent of PFC. The distribution of such typological patterns among the worlds' languages thus seem to be worth pursuing in future research.

5. References

[1] Avesani, C. and Vayra M. "Accenting Deaccenting and Information Structure in Italian Dialogues", In L. Dybkjaer and W. Minker eds., Proceedings of the 6th DIGdial Workshop on Discourse and Dialogue, 19-24, Lisbona, 2005.

[2] Boersma, P., 2001. Praat, a system for doing phonetics by computer. Glot International 5:9/10: 341-345.

[3] Chen, S.-w., Wang, B. and Xu, Y, "Closely related languages, different ways of realizing focus", Proceedings of Interspeech 2009, Brighton, UK, 1007-1010, 2009.

[4] Cooper, W. E., Eady, S. J., & Mueller, P. R.., "Acoustical aspects of contrastive stress in question-answer contexts", Journal of the Acoustical Society of America, 77, 2142-2156, 1985.

[5] Dohen, M. and Lœvenbruck H., "Pre-focal rephrasing, focal enhancement and post-focal deaccentuation in French", Proceedings of the 8th International Conference on Spoken Language Processing, 1313-1316, Jeju, Korea, 2004.

[6] Féry, Caroline, "Focus in French", In Féry, C. and Wolfgang Sternefeld eds., Audiatur Vox Sapientiae. A Festschrift for Arnim von Stechow, 153-181. Berlin: Akademie-Verlag, 2001.

[7] Gu, W.; Lee, T., 2007. Effects of focus on prosody of Cantonese speech – A comparison of surface feature analysis and model-based analysis. In Proceedings of Paraling'07, Saarbrücken. pp.

[8] Jun, Sun-Ah, "The phonetics and phonology of Korean prosody", Ph.D. Dissertation, The Ohio State University, 1993.

[9] Jun, S.A & H.J. Lee, "Phonetics and phonological markers of contrastive focus in Korean", Proceedings of the 5th International Conference on Spoken Language Processing, 4, 1295–1298, Sydney, Australia, 1998.

[10] Kenstowicz, Michael and H. S. Sohn, "Focus and phrasing in northern Kyungsang Korean", In P-M. Bertinetto, ed., Certamen Phonologicum, 3, 137-156. Torino: Rosenberg & Sellier, 1997.

[11] Krifka, Manfred. "Non-novel Indefinites in Adverbial Quantification", In C. Condoravdi and G. Renardel de Lavalette eds., Logical Perspectives on Language and Information, 1-40. Stanford: CSLI Publications, 2001.

[12] Maekawa, Kikuo, "Is there 'Dephrasing' of the accentual phrase in Japanese?", Working Papers in Linguistics 44, 146-165, Ohio State University, 1994.

[13] Oh, Mira, "Focus and prosodic structure, Speech Sciences 8 (1), 21–32, 2001.

[14] Swerts, M., Krahmer E. and Avesani C. "Prosodic Marking of Information Status in Dutch and Italian: A Comparative Analysis", Journal of Phonetics 30, 629-654, 2002.

[15] Ueyama, Motoko and Sun-Ah Jun, "Focus realization in Japanese English and Korean English intonation", vol. 7, 629-645, CSLI, 1998.

[16] Xu, Y., "Effects of tone and focus on the formation and alignment of F0 contours", Journal of Phonetics 27, 55-105, 1999.

[17] Xu, Y., 2005. Speech melody as articulatorily implemented communicative functions. Speech Communication 46: 220-251.

[18] Xu, Y., "_TimeNormalizeF0.praat. Online: <http://www.phon.ucl.ac.uk/home/yi/tools.html>,"2005-2009.

[19] Xu, Y. and Xu, C. X., "Phonetic realization of focus in English declarative Intonation", Journal of Phonetics 33, 159-197, 2005.