Phonetic Realization of Second Occurrence Focus in Japanese

Satoshi Nambu[1], *Yong-cheol Lee*[2]

[1,2]Department of Linguistics, University of Pennsylvania, U.S.A.

satoshin@babel.ling.upenn.edu, leyong@babel.ling.upenn.edu

Abstract

Previous studies have recently agreed that second occurrence focus is phonetically realized as prosodic prominence. What has been missing in the previous studies, however, is a comparison with neutral-focus, in addition to main focus and pre/post-focus, which is necessary to elucidate a precise phonetic status of second occurrence focus. Using evidence from Japanese, this study shows that second occurrence focus in the pre/post-focus position is realized with high pitch less salient than main focus but more than pre/post-focus. Compared with neutral-focus, the pitch of second occurrence focus is higher in the pre-focus position but lower in the post-focus position due to post-focus compression. Furthermore, this study provides a cross-linguistic insight of focus realization. The result suggests that Japanese focus experiences pre-focus compression, in addition to post-focus compression, which is different from Korean, English, and Mandarin.

Index Terms: prosody, focus, second occurrence focus, pre-focus compression, post-focus compression

1. Introduction

The phonetic realization of second occurrence focus (SOF) has been raised as an issue in the literature [1, 2, 3] due to a lack of distinct prosodic prominence. SOF is a focused element that is given from the discourse context (old information). This contrasts with main focus (or first occurrence focus) which is contextually new as shown in example 1. A focused element in English is canonically realized with a pitch accent whose acoustic correlates usually are higher F0, longer duration, and greater intensity. Thus, a correspondence between semantic focus and prosodic prominence is generally presumed. However, it is not straightforward with respect to SOF. For pitch, [4] investigated English SOF in a post-focus position and found that there is no difference between SOF and post-focus. [5] conducted an experiment using German and showed that SOF in a pre-focus position has a higher pitch than pre-focus but lower than main focus. The findings about the pitch of SOF from the previous studies are summarized in 2.

(1) Second Occurrence Focus ([2])

A: Everyone already knew that Mary only eats [vegetables]$_{\mathrm{MF}}$.

B: If even [Paul]$_{\mathrm{MF}}$ knew that Mary only eats [vegetables]$_{\mathrm{SOF}}$, then he should have suggested a different restaurant.

(2) a. Pitch of SOF in a pre-focus position ([5])

Main Focus $>$ SOF $>$ Pre-Focus

b. Pitch of SOF in a post-focus position ([4])

Main Focus $>$ SOF = Post-Focus

The hierarchy in 2 shows that the pitch of SOF is not as salient as main focus, and it is realized higher than pre-focus in a pre-focus position but equal to post-focus in a post-focus position. [4, 5] attributed the unmarked pitch of SOF in a post-focus position to deaccenting (or post-focus compression), and when SOF appears in a pre-focus position, it has pitch prominence.

What has been missing in the previous studies, however, is a comparison of SOF with neutral-focus. Contrasting with neutral-focus, the phonetic nature of SOF can be elucidated more clearly. The prediction posited about the pitch of SOF relative to neutral-focus is as follows. SOF in a pre-focus position should be higher than neutral-focus because it is focused, and SOF in a post-focus position should be lower than neutral-focus due to deaccenting. Thus, this study provides a more precise status of SOF through a comparison with neutral-focus. To investigate this, we have examined SOF in Japanese; and thus, this study also provides a cross-linguistic perspective of SOF.

2. Experiment

2.1. Stimuli

Four sets of sentences served as stimuli for four conditions: neutral-focus, main focus, SOF, and pre/post-focus. To verify the effect of pre/post-focus position on SOF, the condition was divided into two subsets, depending on whether it occurs in a pre-focus or post-focus position. For comparison, the remaining three conditions also contained the same subsets, and we inclusively labeled them as an initial position and a final position for ease. 28 target sentences were used, and each sentence contained 6 words. To serve as the target sentences for naturally occurring SOF, we used a variety of words. In addition, if the measurements are consistent across the conditions, this raises the credibility of phonetic realization in each condition.

For the target sentence with main focus, one sentence was followed as discourse. The sentence with SOF was preceded by two sentences in order to introduce SOF naturally. For target sentences with neutral-focus that do not follow any discourse contexts, 8 fillers were presented. To elicit pre/post-focus, each target sentence was preceded by a prompt question.

To introduce main focus and SOF in the stimuli, we used focus particles *dake* 'only' and *mo* 'also'. The focused element in a pre/post-focus condition corresponds to the answer to the prompt question. We illustrate some examples as stimuli using *dake* 'only' in 3 below. The target constituents are underlined.

(3) **Main Focus in an initial position**

A. Gakkai-de hotondo-no gakusee-wa happyoo-o shinkenni kikimashita.
'Most students listened to presentations seriously at the conference.'

B. Demo [Masaya]$_{\mathrm{F}}$-**dake** sokode happyoo-o shinkenni kikimasendeshita.

'But only Masaya did not listen to presentations seriously there.'

Main Focus in a final position

A. Kanakotachi-wa kinoo pasokon-o sonoomise-ni kainiikimashita.
'Kanako went to the store to buy a pc yesterday.'

B. Demo Kanako-wa kinoo [mausu]$_{F}$-**dake** sokode kaimashita.
'But Kanako only bought a mouse there.'

SOF in an initial position

A. Kinoo hotondo-no gakusee-wa ie-de shikenbenkyoo-o shimashita.
'Yesterday most students studied for the exam at home.'

B. Demo [Naoko]$_{F}$-**dake** ie-de dorama-o chakkari mimashita.
'But only Naoko watched a drama at home.'

C. Soreni [Naoko]$_{SOF}$-**dake** ie-de [anime]$_{F}$-**mo** chakkari mimashita.
'And only Naoko also watched an anime at home.'

SOF in a final position

A. Doyoobi-ni gakuseetachi-wa toshokan-de hon-o yomimashita.
'The students read books at the library on Saturday.'

B. Sonohi Makoto-wa toshokan-de [shosetsu]$_{F}$-**dake** nesshinni yomimashita.
'Makoto only read novels devoutly at the library on the day.'

C. Sonohi [Yuuka]$_{F}$-mo sokode [shosetsu]$_{SOF}$-**dake** nesshinni yomimashita.
'Yuka, also, only read novels devoutly at the library on the day.'

Pre-Focus

A. Kinoo Nanako-wa Kyoto-de dare-to hisashiburini aimashitaka?
'Who did Nanako meet in Kyoto yesterday after a long time?'

B. Kinoo Nanako-wa Kyoto-de [Masaya]$_{F}$-to hisashiburini aimashita.
'Nanako met Masaya in Kyoto yesterday after a long time.'

As seen in the examples, *dake* 'only' and *mo* 'also' are realized as bound morpheme and only modify the constituent to which they attach. We measured the entire noun phrase including the focus particles. The length of target constituents are different because of the length of focus particles. Thus, this study focuses on the pitch realization of each target constituent.

2.2. Subjects

5 native speakers of Tokyo Japanese participated in the experiment, ranging in age from 25-34 (3 males and 2 females). They were recruited at the University of Pennsylvania and signed a consent form. They were paid for their participation. None of the participants reported any problems with their speech or hearing.

2.3. Procedure

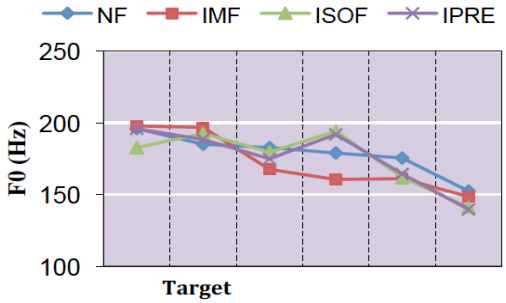
Recordings were made in a sound-proof booth in the Department of Linguistics at University of Pennsylvania. The reading materials were presented on a sheet of paper, and stimuli were randomized. The subjects were asked to read the sentences as naturally as possible. Stimuli for pre/post-focus were elicited in a communicative task. The test assigner read the prompt question and the subjects read the answer. All sentences were recorded electronically and saved on a computer as *wav* files, using Praat [6]. They were also asked to repeat the sentences if they made a mistake or if we found their reading to be unnatural.

2.4. F0 extraction and the Overall Picture

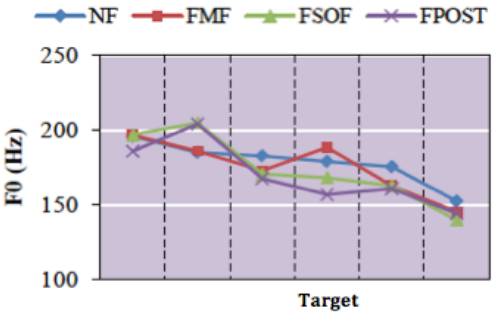
A Praat script was used to extract F0 of the target sentences [7]. The word boundaries were hand-labeled. The script also generated additional measurements to be used in the data analysis, such as mean F0 and maximal F0 (Max F0). After completing the process of F0 extraction, all the target sentences were converted to graphs to identify inconsistencies.

Figure 1 displays the mean F0 and maximal F0 of all the target sentences with five boundaries for the six words in each sentence. Each figure has F0 lines for the four conditions, and the target constituents are located in the initial position or the final position. The figures provide an overview of F0 trends in fluctuation. In the initial position (marked as 'Target' in Figure1), the mean F0 of the target constituent (Figure 1a) shows that main focus has the highest pitch, and that SOF is the second highest. The neutral-focus and pre-focus are low and show similar patterns for the target constituents. As for Max F0, results for the initial position (Figure 1c) show a pattern similar to the mean F0 except that Max F0 of neutral-focus is relatively high. In the final position, the mean F0 of the target constituent (Figure 1b) shows a highest to lowest pitch order of main focus $>$ neutral-focus $>$ SOF $>$ post-focus. Here, SOF is lower than neutral-focus because it undergoes post-focus compression (PFC) after main focus [8, 9, 10], as we can clearly observe in Figure 1b. SOF is still higher than post-focus whose pitch is also suppressed by PFC. The Max F0 (Figure 1d) represents the same pattern.

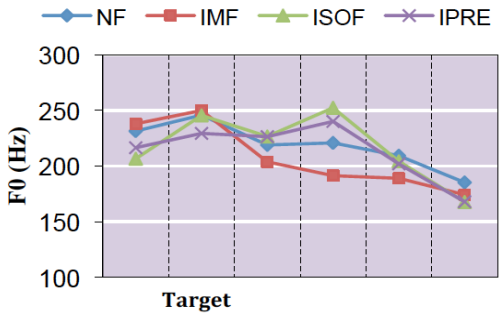
In addition to the pitch of SOF, we observe a general phenomenon related to a phonetic realization of focus in Figure 1. The main foci show a high F0 peak, which is followed by a decrease in the F0, which contrasts with the F0 of neutral-focus. This indicates an effect of PFC. Furthermore, Figure 1 suggests an intriguing characteristic of focus effect in Japanese. When we look at the main focus located in the final position, which occurs in the target sentences for main focus, SOF in the initial position, and pre-focus, the figures show that the element immediately before the main focus experiences F0 compression. Contrasted with the F0 of neutral-focus, the pitch preceding the main focus is lowered.



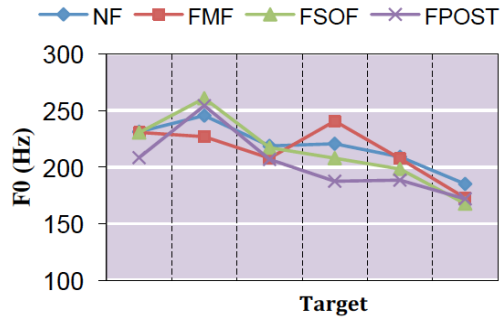
a) Mean F0, Target in an initial position



b) Mean F0, Target in a final position



c) Max F0, Target in an initial position

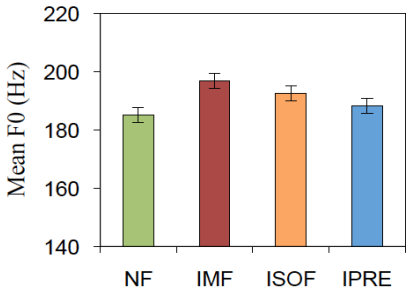


d) Max F0, Target in a final position

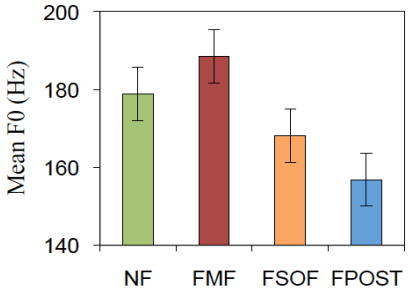
Figure 1: *Mean F0 and Max F0 of all the target sentences spoken by five speakers (averaged in Hz) (NF=Neutral-Focus, IMF=Initial Main Focus, ISOF=Initial Second Occurrence Focus, IPRE=Initial Pre-Focus, FMF=Final Main Focus, FSOF=Final Second Occurrence Focus, FPOST=Final Post-Focus).*

2.5. Analysis and Results

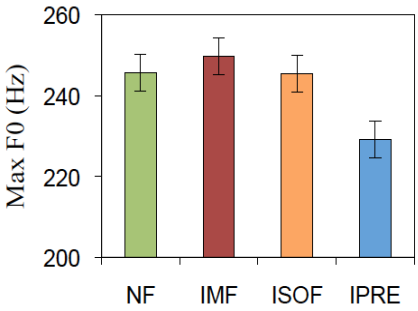
To make a direct comparison between the four conditions, we take a closer look at the mean F0 and Max F0 of the target constituents with main focus, SOF, neutral-focus, and pre/post-focus. Figure 2 provides the mean F0 and Max F0 with standard errors for the initial and final positions, and shows that main focus always has the highest F0 with SOF second in the initial position. SOF is lower than neutral-focus in the final position. The results of paired t-tests on the mean F0 across conditions are shown in Table 1. For the initial position (Table 1a), the mean F0 of SOF (192.5 Hz) is significantly higher than pre-focus (188.4 Hz) and neutral-focus (185.0 Hz). There is no significant difference between neutral-focus and pre-focus. The difference between main focus (196.7 Hz) and SOF is marginally significant. As for the final position (Table 1b), the mean F0 of main focus (188.4 Hz) is significantly higher than neutral-focus (178.9 Hz), and neutral-focus is significantly higher than SOF (168.1 Hz). The difference between SOF and post-focus (156.86 Hz) is also significant.



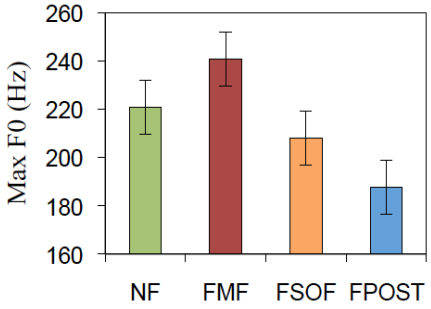
a) Mean F0, Target in an initial position



b) Mean F0, Target in a final position



c) Max F0, Target in an initial position



d) Max F0, Target in a final position

Figure 2: *Means and standard errors (error bars) of mean F0 and Max F0 in the four conditions with an initial/final position (NF=Neutral-Focus, IMF=Initial Main Focus, ISOF=Initial*

Second Occurrence Focus, IPRE=Initial Pre-Focus, FMF=Final Main Focus, FSOF=Final Second Occurrence Focus, FPOST=Final Post-Focus).

Table 1: *Results of paired t-test for mean F0 in four conditions*

(df=119)

	INF	IMF	ISOF
IMF	T = -4.14 $p < 0.001$	-	-
ISOF	T = -2.37, $p < 0.05$	T = 1.81 $p = 0.071$	-
IPRE	T= -0.98 $p = 0.327$	T = 2.89 $p < 0.01$	T = 1.78 $p = 0.077$

a) Target in the initial position
Ranking: IMF > ISOF > INF, IPRE

(df=119)

	FNF	FMF	FSOF
FMF	T = -3.71 $p < 0.001$	-	-
FSOF	T = 3.19 $p < 0.01$	T = 8.11 $p < 0.001$	-
FPOST	T = 6.09 $p < 0.001$	T = 12.74 $p < 0.001$	T = 5.70 $p < 0.001$

b) Target in the final position
Ranking: FMF > FNF > FSOF > FPOST

In relation to the effect of focus in the pre/post-focus position, we conducted paired-t tests to confirm whether pre/post-focus elements experience pitch compression. In terms of PFC, the difference in pitch between the neutral-focus element (mean = 175.4 Hz) and the post-focus element (mean = 163.5 Hz) is statistically significant (T[119] = 3.199, $p < 0.01$). For the pre-focus element, the paired t-test shows a statistically significant difference between the pre-focus element (mean = 173.8 Hz) and the neutral-focus element (mean = 182.7 Hz) (T[119] = -3.7174, $p < 0.001$). Thus, Japanese focus exhibits an effect of PFC and pre-focus compression.

3. Discussion

From the results of paired t-tests, we found evidence of pitch prominence with SOF in the pre-focus position, which is the same as German. The idiosyncratic status of Japanese SOF also has pitch prominence when it occurs in a post-focus position that has PFC. Although the pitch prominence of SOF in a post-focus position is less salient than main focus, it is still significantly higher than post-focus, which is different from English and German. In addition, we found that the pitch of SOF in a pre-focus position is higher than neutral-focus, and lower in a post-focus position than neutral-focus. This clearly represents that SOF has prosodic prominence, and that SOF in a post-focus position experiences PFC that masks the prominence of SOF, though not completely in the case of Japanese. The results are summarized in 4.

(4) a. Pitch of four conditions in the initial position
Main Focus > SOF > Neutral-Focus = Pre-Focus

b. Pitch of four conditions in the final position
Main Focus > Neutral-Focus > SOF > Post-Focus

In reference to the effect of focus in the pre/post-focus position, Japanese not only has PFC but also has pre-focus compression. With respect to PFC, Japanese can be grouped with languages such as Korean, English, and Mandarin, which also have the effect of PFC, as contrasted with Cantonese and Taiwanese Chinese that do not [10]. In addition, the results show that focus in Japanese also represents pitch compression immediately preceding the focused element. Korean and English, however, do not show this effect. Thus, the existence of pre-focus compression in Japanese gives us a more sophisticated insight of typological focus realization.

4. Conclusion

We clarified the status of the prosodic prominence of SOF by incorporating neutral-focus into the hierarchy of pitch prominence. This study presented an intermediate level of pitch realization of SOF in Japanese, showing that it is always less prominent than main focus. The comparison of SOF with neutral-focus indicates that SOF in a pre-focus position has pitch prominence but that SOF in a post-focus position is lowered due to the effect of PFC. The pitch of SOF is higher than pre-focus and post-focus even when it is under PFC. The finding of pitch prominence for SOF in a post-focus position in Japanese proves the direct correlate of semantic focus and prosodic prominence. Furthermore, this study found that Japanese focus has not only PFC but also pre-focus compression, which is different from English, Korean, and Mandarin [11]. To ascertain more precise characteristics of pre-focus elements in Japanese, further research is needed.

5. References

[1] Rooth, M., "On the interface principles for intonational focus.", in T. Galloway and J. Spence [Ed], SALT VI, 202-226, 1996.

[2] Partee, B., "Focus, quantification, and semantics-pragmatics issues.", in P. Bosch and R. van der Sandt [Eds], Focus: Linguistics, cognitive, and computational perspectives, Cambridge University Press, 213-231, 1999.

[3] Krifka, M., "Focus and/or context: A second look at second occurrence expressions.", in H. Kamp and B. Partee [Eds], Context-Dependence in the Analysis of Linguistic Meaning, Elsevier Science, 187-207, 2004.

[4] Beaver, D., Clark, B. Z., Flemming, E., Jaeger T. F., and Wolters, M., "When semantics meets phonetics: Acoustic studies of second occurrence focus.", Language, 83:245-276, 2007.

[5] Féry, C. and Ishihara, S., "The phonology of second occurrence focus.", Journal of Linguistics, 45:285-313.

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

[7] Xu, Y., "_TimeNormalizeF0.praat. Online: <http://www.phon.ucl.ac.uk/home/yi/tools.html>," 2005-2010.

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

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

[10] Xu, Y., Chen, S., and Wang, B., "Prosodic focus with post-focus compression: Single or Multiple Origin?", The Second Workshop on Evolutionary Linguistics, 2010

[11] Lee, Y.-C., and Xu, Y., "Phonetic realization of contrastive focus in Korean.," Speech Prosody 2010, 100033:1-4, ISBN:978-0-557-51931-6.