Phonetic Realization of Prosodic Focus in Persian

Mortaza Taheri Ardali[1], *Yi Xu*[2]

[1]Department of Linguistics, Institute for Humanities and Cultural Studies, Tehran, Iran
[2]Department of Speech, Hearing and Phonetic Sciences, University College London, UK
ncmit1957@yahoo.com, yi@phon.ucl.ac.uk

Abstract

It has been repeatedly observed that focus substantially changes the sentence prosody in many languages not only by increasing F_0, duration, and intensity on the focused components but also by compressing the pitch range and intensity of post-focus elements. However, it is not yet fully clear in Persian what the main effect of focus is on pre-focus, on-focus and post-focus elements. To achieve this goal, we have embarked on a full-scale investigation of the phonetic realization of prosodic focus in Persian. The findings of this study reveal that focus dramatically changes the three regions. F_0 and duration has significantly increased in on-focus words without any significant change in intensity. Compared to their counterparts, pre-focus elements show weaker intensity but no change in duration and mean F_0. Finally post-focus words show significant lowering of F_0 and decrease of intensity. The duration of post-focus words remains intact. Thus, according to the present data, it can be concluded that Persian, like English and Mandarin, falls into the category of PFC (post-focus compression) languages.

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

1. Introduction

Prosodic focus as a core element of speech prosody has been examined in many studies. Although in some cases, there have been inconsistencies in the methodologies used, it is generally reported that the focused component is given increased F_0, duration, intensity and upper spectral energy [4, 5, 6]. In addition to the change of on-focus region, it is also observed in recent studies that many languages encode prosodic focus with a reduction of pitch range and intensity in post-focus region, which is referred to as post-focus compression [14]. PFC has been already found in languages such as English, Mandarin and Korean among many others [2, 7, 13, 15, 16]. However, the absence of PFC has been reported for Taiwanese and Cantonese, both Chinese languages closely related to Mandarin [2, 9, 12].

Persian, an SOV language, is a member of the Southwestern group within the Iranian branch of the Indo-Iranian languages [3]. The dialect under this study is Modern Conversational Persian which is spoken in Tehran, the capital city of Iran.

There has been little research on the effect of prosodic focus in Persian. Mahjani [8] says that focused word lengthens in duration considerably, while words before and after it are usually shorter. Sadat-Tehrani [10], on the other hand, has reported greater pitch excursion and longer duration of focal elements and de-accentuation of all post-focal elements. Taheri Ardali [11] argues that the only significant difference due to the effect of focus on F_0 and duration is in focused words without much difference in pre- and post-focus region.

The main aim of the current study is to investigate the phonetic realization of prosodic focus in Persian. We have made a systematic comparison of F_0, duration and intensity in different focus positions to find out whether there is PFC in Persian.

2. Method

2.1. Material and Participants

To check the effect of focus on different parts of a sentence, we measured duration, F_0 and intensity of the sentence components before, on and after focus with their non-focused counterparts. Five versions of the same sentence were used, differing only in focus location. Table 1 shows the stimulus sentence. To elicit contrastive focus on a specific word, a sentence was given in parentheses before the target sentence. This focus cueing sentence is the same as the target sentence except the focused word, the verb and ending 'but'. In both sentences, the word in contrast is marked in bold. For example: (**una** babaye niliro lændæn nædidæn, bælke) **maha** babaye niliro lændæn didim '(**they** didn't see Nili's father in London, but) **we** saw Nili's father in London.'

Table 1.*The stimulus sentence of the experiment*

Maha	baba-ye	nili-ro	lændæn	did-im
We-PL	father-EZ	Nili-DO	London	see.PST-1PL

'we saw Nili's father in London.'

Five male speakers with the age range of 18-29 (mean = 25 years) participated in this experiment. None of the participants reported any speech or hearing impediments and all were native speakers of Persian.

2.2. Procedure

The data were recorded in a very quiet room. Before sitting in front of a computer, the speakers were instructed how to read the sentences naturally and at a normal rate. A microphone was placed to the right of the monitor. For each focus location, each speaker repeated the sentences five times in blocked random order. They also uttered the non-focused sentence five times. Using Praat [1] the data were recorded with a sampling rate of 22050 in mono wave format. The speakers were asked to repeat a sentence if they made a mistake or if we found their reading unnatural. Ultimately we obtained 30 sentences from each speaker.

2.3. F_0 extraction

A script [17] written for the software Praat was used to extract all the measurements including mean F_0, max F_0, duration, and mean intensity for the data analysis. At first we hand-labeled the word boundaries. Then with the help of this script we did the rectification manually for missed or double marked vocal cycles in the wave form. After obtaining the F_0 values, we converted all the sentences to graphs to check for differences

due to focus. Figure 1 displays time-normalized mean F_0 curves averaged across all tokens produced by the speakers. As it is evident, in the focused words, which are separated by solid vertical lines, the pitch range is robustly raised and suddenly drops afterwards. We can also see the substantial lowering of the pitch range in all post-focus regions. In contrast, the pitch range of pre-focus words is more or less the same as that of the neutral-focus counterparts.

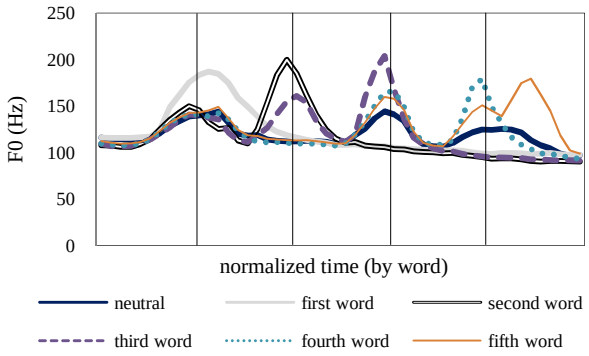

Figure 1. *Time-normalized mean F_0 contours of all the sentences uttered by the speakers. Each curve is an average of 25 repetitions by five speakers.*

2.4. Analyses and Results

In order to make comparison between focused and neutral-focus sentences in three different regions, we considered mean F_0, max F_0, duration and mean intensity as dependent variables and focus (focus, neutral) and focus position (w1, w2, w3, w4, w5) as independent variables. For any focus position, the values of the dependent variables are the averages of all the words that are in the appropriate region. For example, when focus is on w1, on-focus values is only from the on-focus word, post-focus values are the averages of the second to the fifth words, but there are no pre-focus values. To avoid the confound of carryover influence of the preceding word, the measurements of mean F_0, max F_0 and mean intensity were taken only from the second half of each word.

The results of 2-way repeated measures ANOVAs on mean F_0, max F_0, duration and mean intensity are shown in Table 2, and the mean values are shown in Figure 2.

For on-focus words, mean F_0 and max F_0 are both significantly higher and duration is longer than their neutral-focus counterparts. However, there is no on-focus difference in mean intensity. For post-focus words, mean F_0, max F_0 and mean intensity are significantly lower than the neutral-focus counterparts. However, there is no significant difference in duration. For pre-focus words, max F_0 is higher but mean intensity is lower than the neutral focus counterparts. There is no difference in mean F_0 or duration.

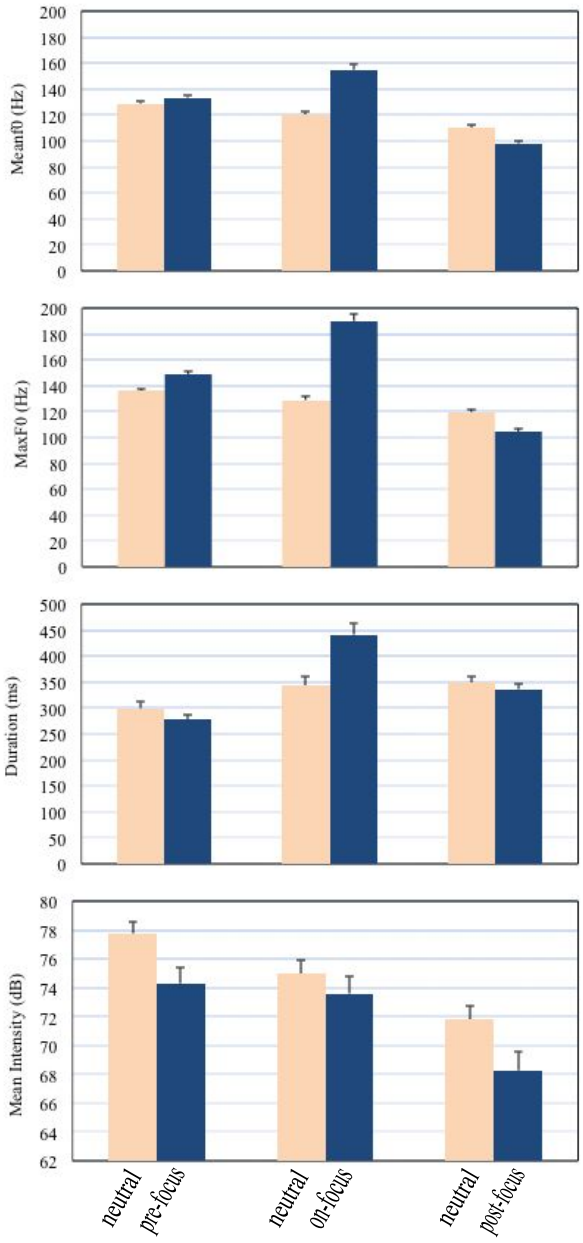

Figure 2. *Means of max F_0, mean F_0, duration and mean intensity in pre-focus, on-focus, and post-focus areas and corresponding neutral-focus areas.*

Table 2. *Results of 2-way repeated measures ANOVAs for all measurements in three different regions.*

Focus (df = 1, 4)			
	Pre-focus	On-focus	Post-focus
Mean F_0	F = 2.561 *P* = 0.1848	F = 100.358 ***P* = 0.0006**	F = 20.473 ***P* = 0.0106**
Max F_0	F = 19.315 ***P* = 0.0117**	F = 127.28 ***P* = 0.0004**	F = 21.297 ***P* = 0.0099**
Duration	F = 7.351 *P* = 0.0535	F = 63.648 **P = 0.0013**	F = 5.596 *P* = 0.0772
Mean Intensity	F = 18.343 ***P* = 0.0128**	F = 2.073 *P* = 0.2233	F = 8.189 ***P* = 0.0459**
Focus Position (df = 3,12; 4,16; 3,12)			
	Pre-focus	On-focus	Post-focus
Mean F_0	F = 6.687 ***P* = 0.0066**	F = 27.783 ***P*< .0001**	F = 45.905 ***P*< .0001**
Max F_0	F = 2.653 *P* = 0.0962	F = 9.061 ***P* = 0.0005**	F = 20.17 ***P*< .0001**
Duration	F = 99.549 ***P*< .0001**	F = 61.523 ***P*< .0001**	F = 114.955 ***P*< .0001**
Mean Intensity	F = 3.448 *P* = 0.516	F = 122.438 ***P*< .0001**	F = 32.318 ***P*< .0001**
Focus Position*Focus (df = 3,12; 4,16; 3,12)			
	Pre-focus	On-focus	Post-focus
Mean F_0	F = 1.171 *P* = 0.3613	F = 4.301 ***P* = 0.015**	F = 14.206 ***P* = 0.0003**
Max F_0	F = 4.181 ***P* = 0.0305**	F = 3.182 ***P* = 0.0421**	F = 15.616 ***P* = 0.0002**
Duration	F = 1.396 *P* = 0.2917	F = 6.165 ***P* = 0.0034**	F = 4.064 ***P* = 0.0331**
Mean Intensity	F = 1.58 *P* = 0.2456	F = 4.608 ***P* = 0.0115**	F = 4.401 ***P* = 0.0263**

The middle panel in Table 2 shows the effect of focus position on the four measurements. The corresponding mean values are plotted in Figure 3. As can be seen in Figure 3, the highly significant positional differences in mean F_0 and max F_0 in pre-focus and post-focus regions are mostly due to a trend for F_0 to become lower over the course of a sentence. However, Bonferroni/Dunn post-hoc tests showed only some significant cross-position differences (pre-focus mean F_0: w2-w4, w2-w5, post-focus mean F_0: most position pairs except w3-w4, post-focus max F_0: w1-w3, w1-w4, w2-w4). In the on-focus region, the sentence medial focus, again, only some of the differences are significant in the post-hoc tests (on-focus mean F_0: w1-w5, w2-w5, w3-w4, w3-w5, w4-w5; max F_0: w1-w5, w2-w5, w3-w5).

For duration, the general trend is an increase over the course of the sentence, with some exceptions (pre-focus: all pairs; on-focus: all except w1-w5, w2-w3, w3-w4; post-focus: w1-w2, w1-w3, w2-w3).

For mean intensity, the general trend is similar to that of F_0, decreasing over the course of the sentence. However, post-hoc tests (Bonferroni/Dunn) showed no difference in the pre-focus regions. For on-focus region, the differences in all pairs are significant except w3-w4. For post-focus region, all pairs are significant except w2-w3.

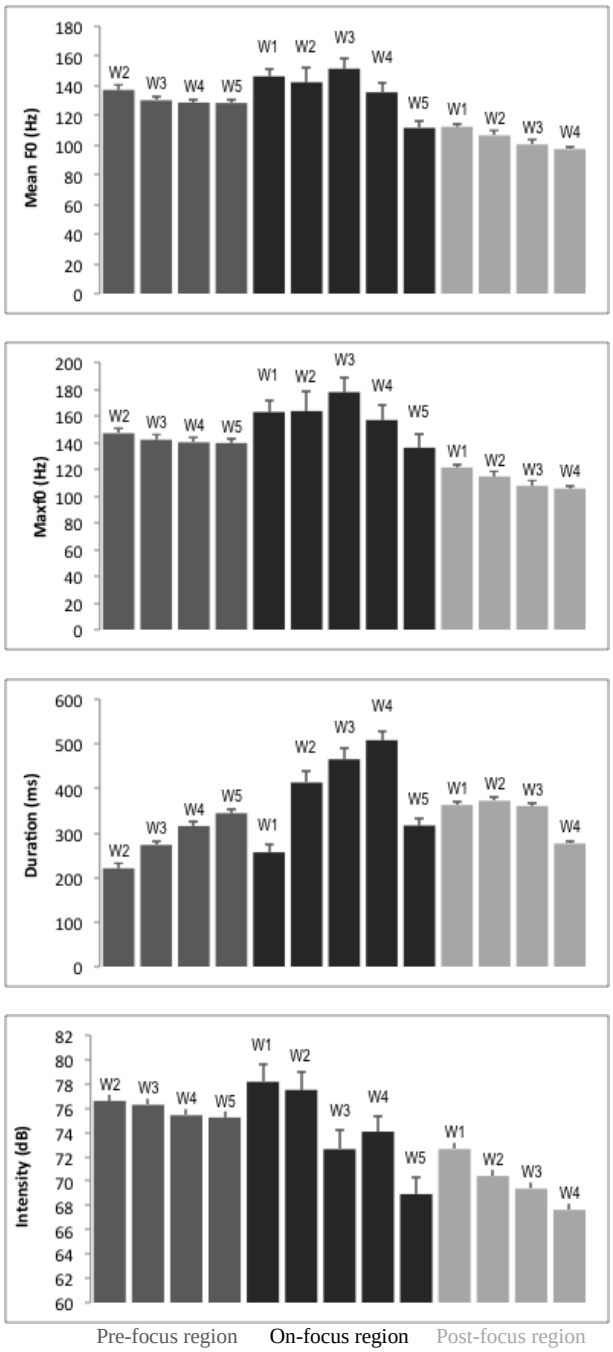

Figure 3. *Effects of focus position. Each bar is an average of the measurements from all the words in the respective region. Pre-focus: w2 = w1, w3 = average (w1, w2), w4 = average (w1, w2, w3), w5 = average (w1, w2, w3, w4). Post-focus: w1 = average (w2, w3, w4, w5), w2 = average (w3, w4, w5), w3 = average (w4, w5), w4 = w5. On-focus: value of each focused word.*

The bottom panel in Table 2 shows the interactions of focus and focus position for all the measurements. Worth particular mentioning are the large effects on max and mean F_0 in the post-focus region. This is mainly due to much smaller focus effects when focus is on w1 and w4 than on w2 and w4. Judging from Figure 1, this is likely due to the fact that in

these two focus positions, some remaining carryover effect must have been included even when the measurements are taken in the later half of the first post-focus word. When focus is on w2 or w3, all carryover F_0 raising seem to have disappeared by the middle of the following word. This shows the importance of avoiding the confound of carryover effects in assessing the existence of PFC in a language.

Another large interaction between focus and focus location occurred in the on-focus duration. This is mainly due to a much larger focus effect when focus is on w2, w3 and w4.

3. Discussion

According to the current data it seems that the words directly under focus have higher pitch and longer duration than the same words in neutral-focus condition. Intensity, on the other hand, does not show any significant difference in this region. Pre-focus elements were realized with weaker intensity compared to non-focused counterparts. Although the effect of focus on mean F_0 of pre-focus region was not significant, max F_0 is higher pre-focally. In the post-focus region, max F_0, mean F_0 and mean intensity of post-focus words were significantly decreased in comparison with the corresponding neutral focus ones. There was, however, no difference in duration in the post-focus region.

With regard to the mixture of results in different zones, it could be argued, as suggested by Xu et al. [14], that the domain of a single focus consists of three temporal zones, with distinct pitch range, duration and intensity adjustments for pre-, on- and post-focus components. It is also found that for Persian speakers both F_0 and duration are the acoustic correlates of on-focus elements. Both are considerably increased in all focused words. In contrast, intensity did not show significant difference in that region.

4. Conclusions

The results of the current study reveal that F_0 and duration are the main acoustic correlates of prosodic focus in Persian. Following [14] we conclude that the temporal domain of focus is much wider than that of the focused item itself. Comparing with other languages, we can group Persian with English, Mandarin and many others as PFC languages, which are in contrast with Taiwanese, Cantonese and many others where PFC is found to be absent [16]. Future collaborative research, with similar methodology, will help to find out whether PFC exists in languages where this property has not yet been explored.

5. References

[1] Boersma, P., "Praat, a system for doing phonetics by computer". Glot International 5:9/10: 341-345, 1992-2011.

[2] 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.

[3] Comrie, B. *The world's major languages*, 2ed edition, Routledge, London, 2009.

[4] Cooper, W., S. Eady, & P. Mueller. "Acoustical aspects of contrastive stress in question-answer contexts", JASA 77(6): 2142-2156, 1985.

[5] De Jong, K., "Stress, lexical focus, and segmental focus in English: patterns of variation in vowel duration", *Journal of Phonetics* 32: 493-516, 2004.

[6] Heldner, M., "On the reliability of overall intensity and spectral emphasis as acoustic correlates of focal accents in Swedish". *Journal of Phonetics* 31: 39-62, 2003.

[7] Lee, Y.-c,. Xu, Y., Phonetic realization of contrastive focus in Korean, In Proceeding of speech Prosody 2010, Chicago, 2010.

[8] Mahjani, B., "An Instrumental study of prosodic features and intonation in Modern Farsi (Persian)", M.Sc. thesis, retrieved from: http://www.ling.ed.ac.uk/teaching/postgrad/mscslp/archive/dissertations/2002-3/behzad_mahjani.pdf. 2003.

[9] Pan, H., "Focus and Taiwanese unchecked tones", In Topic and Focus: Cross-linguistic Perspectives on Meaning and Intonation. C. Lee, M. Gordon and D. Büring. Springer: 195-213, 2007.

[10] Sadat-Tehrani, N., "Intonational grammar of Persian", Doctoral dissertation. Manitoba: University of Manitoba, 2007.

[11] Taheri Ardali, M., "The intonation of focus in declarative sentences in Persian", M.A. thesis, Allameh Tabataba'i University, Tehran, 2010.

[12] Wu, W. L., Xu, Y., "Prosodic focus in Hong Kong Cantonese without post-focus compression", In Proceedings of Speech Prosody 2010, Chicago, 2010.

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

[14] Xu, Y., Xu, C. X. & Sun, X., "On the temporal domain of focus", International Conference on Speech Prosody 2004, Nara, 2004.

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

[16] Xu, Y., "Post-focus compression: Cross-linguistic distribution and historical origin", In Proceedings of the 17th International Congress of Phonetic Sciences, Hong Kong, 2011.

[17] Xu, Y., "ProsodyPro.praat", Available from: http://www.phon.ucl.ac.uk/home/yi/ProsodyPro/., 2005-2011.