The Perception of Prosodic Focus in Persian

Mortaza Taheri-Ardali[1], *Hamed Rahmani*[2], *Yi Xu*[3]

[1]Department of Linguistics, Institute for Humanities and Cultural Studies, Tehran, Iran
[2]Department of Linguistics, Radboud University, Nijmegen, Netherlands
[3]Department of Speech, Hearing and Phonetic Sciences, University College London, UK
m.taheri@ihcs.ac.ir, h.rahmani@let.ru.nl, yi.xu@ucl.ac.uk

Abstract

In a previous production experiment, post-focus compression (PFC) of F_0 and intensity were found to be present in Persian. It was also shown that F_0 and duration were the main correlates of prosodic focus in Persian. However, the perceptual relevance of PFC in Persian was not yet clear. The present paper reports the findings of an experiment on focus perception in Persian. Native speakers of Persian listened to sentences produced with focus in different positions as well as the neutral-focus sentence, and judged the presence and location of focus. Results show that final focus is identified much less well than other types of focus, and most of its confusion is with neutral focus. This shows that the presence of PFC is a main factor in recognizing prosodic focus in Persian.

Index Terms: prosodic focus, PFC, perception, post-focus, pre-focus, on-focus, F_0, intensity, Persian.

1. Introduction

There are different strategies to catch the attention of a listener to a particular portion of an utterance, i.e., to mark focus. This can be done both by syntactic and morphological means and by prosodic devices. Prosodically, many languages use phonetic variation in F_0, duration and intensity to mark focus. In particular, prosodic focus is realized not only by increasing F_0, duration and intensity of the on-focus component itself, but also by compressing the pitch range and intensity of the post-focus elements [5, 7, 16, 27, 28, 31]. There is also increasing evidence that post-focus compression as a perceptual cue plays a pivotal role in focus perception [11, 29]. It has been reported that if there are no F_0 peaks after an earlier peak in a sentence, it would be easier for listeners to perceive a non-final focus; otherwise listeners are prone to hear an additional late focus or no focus [15, 18]. It is also shown that when focus is not utterance-final, i.e., when PFC is applicable, its perceptual recognition is much easier, whereas final focus is often confused with neutral focus [3, 5, 13, 18]. Hence, findings from focus perception seem to lend further support to the importance of PFC [29]. However, the perceptual importance of PFC is not widely accepted, and much of the research on focus perception is still mainly concentrating on the focused words only [e.g., 20, 25]. There is therefore a need to explore further evidence for the importance of PFC in focus perception. In this paper, we present data of a perception experiment on Persian, for which PFC has been found in a recent study [22]. Before probing the perceptual effectiveness of PFC, a brief background of Persian prosody is provided in the next section.

2. Persian prosody

Persian, an SOV language with fairly free word order, is an Iranian language within the Indo-Iranian branch of the Indo-European family. Regionally, Persian has three major varieties: (1) the Persian of Iran (2) the Persian of Afghanistan, now called Dari and (3) the Persian spoken in Tajikistan in Central Asia. The variant used in this study is Iranian Persian, the official language of Iran and the mother tongue of about 60% (42 million) of Iran's population. It is worth noting that bilingualism and multilingualism are widely found in Iran [6].

There have been three central issues about Persian prosody. The first concerns word prosody. Persian has traditionally been described as a stress language. Abolhasanizade et al [1] recently found no marked phonetic difference between stressed and unstressed syllables independently of the presence of intonational pitch accents. The authors conclude that Persian word prosody involves a lexically-sensitive pitch accent assignment system that is more like Tokyo Japanese, which has no stress in the phonetic sense, than West Germanic, where stressed and unstressed syllables differ in durational and spectral properties.

The second issue relates to sentence prosody. Initial works, mostly based on the British tradition, have carefully documented the intonational patterns for various sentence types [24, 23]. In more recent literature, there has been a tendency toward the framework of autosegmental-metrical and intonational phonology [8, 14, 19]. A remarkable development to emerge from these studies concerns the issue of main stress in the sentence, generally referred to as the nuclear stress. Eslami [8] and Kahnemuyipour [12] have proposed a number of syntax-based rules to identify the location of the main stress within sentence.

The notion of syntax-based sentence stress, however, is contrasted by the findings related to the third issue, which concerns prosodic focus [1, 10, 19, 21]. According to these findings, any constituent in Persian can be contrastively focused, the only constraint being that focused elements cannot appear post-verbally. The phonetic properties of focus has been addressed in a number of experimental studies. The evidence provided by Sadat-Tehrani [19] points to greater pitch excursion and longer duration of focused elements. In a more detailed experiment, Abolhasanizade et al [1] show that focus has no significant durational and spectral effect and is expressed only by the differences in F_0. That the effect of duration and intensity is negligible is also maintained by Hosseini [10], who claims that F_0 is the only robust acoustic correlate of focus. As for the post-focal region, Abolhasanizade et al [1] provide evidence for PFC of F_0 and intensity. According to these authors, while the pitch range of the post-focal elements is phonetically reduced, the pitch accents are not deleted after the focus. This contradicts earlier report of complete post-focal de-accentuation [19, 21, 10]. The

existence of PFC in Persian is also found by Taheri-Ardali and Xu [22], who showed that focus not only increased the F_0 and duration of words under focus, but also decreased F_0 and intensity of post-focus words. In contrast, they did not find significant changes in on-focus intensity or post-focus duration.

Previous research has therefore established the existence of PFC in Persian. But there have been no empirical studies on the importance of PFC in the perception of focus in Persian, a gap the present study aims to fill.

3. Method

3.1. Stimuli

The sentence used in the perception experiment, as shown in Table 1, was taken from the previous production experiment [22]. The key words in the sentences consisted of mostly sonorant sounds to make sure that the F_0 contours were as smooth and connected as possible. The sentences were produced by five male speakers, who repeated each of the sentences five times in blocked random order. There were a total of 30 utterances from each speaker.

Table 1. *The target sentence of the experiment*

W1	W2	W3	W4	W5
Maha	baba-ye	nili-ro	lændæn	didim
we-PL	father-EZ	Nili-DO	London	see.PST-1PL

To elicit focus on a specific word in the production experiment, the target sentence was preceded by a sentence in parentheses. This focus cueing sentence was the same as the target sentence except the focused word, the verb and the ending word 'but' (Table 2).

Table 2. *The focus cueing sentences*

Focus	Focus cueing sentence (plus 'but')
W1	**They** didn't see Nili's father in London
W2	We didn't see Nili's **uncle** in London
W3	We didn't see **Amini's** father in London
W4	We didn't see Nili's father in **Tehran**
W5	We didn't **take** Nili's father to London

The stimuli for the current perceptual study were selected using mean standard deviation of F_0 across all repetitions of each speaker as arbitrary criteria, a method previously used in [5]. Speakers with minimum, maximum and median standard deviations were chosen. Then all tokens from these three speakers were used stimuli. In total, there were 6 foci x 5 repetitions x 3 speakers = 90 tokens.

3.2. Subjects

Five males and five females participated as subjects. All were native speakers of Persian with average age of 26.1 which is comparable to the age range of those who took part in production experiment. Each subject was paid in exchange for his/her participation in the test. They have also reported no hearing or speech disorders.

3.3. Listening Procedure

The experiment was conducted using ExperimentMFC in Praat software [2]. The listeners were instructed on how to choose the emphasized word. They listened to the stimuli once and then judged which word was focused. They were also told if none of the words was focused, the neutral focus choice must be selected. Before the start of the experimental trials, listeners had five practice trials without any feedback on the correctness of the answers.

3.4. Results

Figure 1 shows focus recognition rates of all six focus conditions, in percentage of correct identification. The overall rate of focus identification is fairly high. Compared to each other, the identification rate for the final focus (last word) is much lower than other focus conditions.

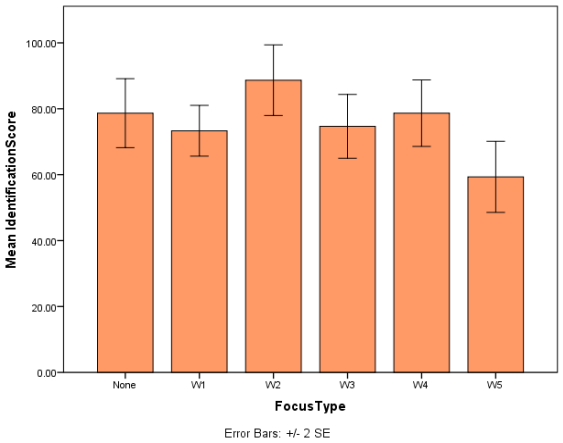

Figure 1. *Percentage of correct identification of neutral focus and focus on word 1-5. The error bars represent standard errors.*

Table 3 shows a confusion matrix of the focus perception experiment. It can be seen that when listeners identified focus location wrongly, they were prone to hearing no focus in the utterance. For example, 29.3% of the final foci that the listeners heard wrongly were heard as neutral focus. In contrast, there were almost no wrongly identified cases that were recognized as final focus.

Table 3. *Confusion matrix of focus perception (percent). Bold face indicates correct focus identification.*

heard as / original	none	W1	W2	W3	W4	W5
none	**78.6**	9.3	4.6	3.3	4	0
W1	24	**73.3**	2	0	0.6	0
W2	6.6	4	**88.6**	0.6	0	0
W3	16.6	2	4.6	**74.6**	2	0
W4	16	0.6	2	2	**78.6**	0.6
W5	29.3	2.6	0.6	3.3	4.6	**59.3**

Table 4 shows results of post-hoc pairwise comparisons between the focus conditions with Bonferroni adjustments. It can be seen that focus on Word 5 has significantly worse recognition rate than neutral, focus on Word 2 and focus on Word 4. In contrast, there is no significant difference between any other two focus conditions.

Table 4. *Results of post-hoc pairwise comparisons. The mean difference is significant at the .05 level.*

Focus Type (I)	Focus Type (J)	Mean Difference (I-J)	Sig.
None	W1	5.334	1.000
	W2	-10.000	.713
	W3	4.001	1.000
	W4	-.001	1.000
	W5	19.335*	**.009**
W1	W2	-15.334	.177
	W3	-1.333	1.000
	W4	-5.335	1.000
	W5	14.001	.598
W2	W3	14.001	.315
	W4	9.999	1.000
	W5	29.335*	**.004**
W3	W4	-4.002	1.000
	W5	15.334	.258
W4	W5	19.336*	**.015**

4. Discussion

Results from Table 3 show that recognition rates of all focus positions were high except for final focus. That the lowest recognition rate is for final focus is in line with many other studies where the final focus had the lowest rate of identification compared to other positions [3]. The most likely reason is that the lack of PFC impedes the easy recognition of focus in this position [13]. This is supported by Figure 2, which displays time-normalized mean F_0 contours of the three speakers whose utterances were used as the perception stimuli in the present study. It can be seen that the F_0 increase by final focus relative to neutral focus is just as substantial as in other focus positions. The only thing missing compared to the non-final focus is the compressed F_0 (and intensity) after the focused word, since there are no words following the final word. The second word has the highest rate of focus identification, followed by the fourth word and neutral focus.

The high recognition rate of focus on the second word can be explained from a syntactic perspective. The basic pattern of Persian sentence prosody is that every major class word receives an accent. From the neutral-focus contour in Figure 2 we can see that all words are accented except the second word, i.e. *babaye* 'father-EZ'. Generally speaking, under some discourse conditions, the head noun of a definite noun phrase may be unaccented when post-modified, as is the case with the head noun *babaye* inside the noun phrase *babaye nili* 'Nili's father' [19]. This usual lack of pitch accent on the second word apparently had a consequence on the focus perception. That is, when asked to put focus on the second word, speakers in fact accented a constituent that would receive no accent in the neutral pronunciation of the sentence. Thus, the focus on the second word involves one extra structural manipulation (i.e. accenting a structurally unaccented word) compared to that on the other words in the sentence which are structurally accented words. This may have made the second word even more salient than the other focused words, and so helped the subjects to more efficiently identify focus on this word.

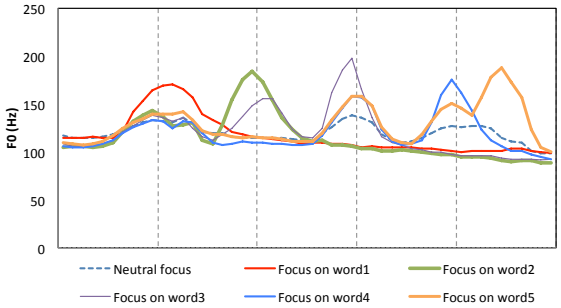

Figure 2. *Time-normalized mean F_0 contours of all the sentences uttered by speakers with* minimum*, maximum and median standard deviations. Each curve is an average of 25 repetitions.*

In Persian, like Uygur [26], initial focus was mostly confused with neutral focus (24%). But in terms of confusion of final focus with neutral focus, with 29.3% wrong identification, Persian acts like Beijing Mandarin [5].

Speakers of Iranian and Turkic languages (both PFC languages) have been in contact since pre-Islamic times [17]. However, to compare with each other, phonetic realization of prosodic focus in Persian and Turkish as two exemplars of Iranian and Turkic languages is not similar in pre-, on- and post-focus regions. To name a few, contrary to Persian, Turkish does not show significant changes in F_0 in on-focus region. And it was also observed that in Turkish pre-focus region is raised in F_0 while in Persian there is no change in mean F_0 in this region. Furthermore, in Turkish, unlike Persian, the acoustic differences in the three above-mentioned regions are not independent of the position of focus. It is worth noting that Azeri language as a Turkic language which is also spoken in Iran is heavily influenced by Persian as an Iranian language. Thus, since Azeri has a common proto language with Turkish, on the one hand, and its close contact with Persian on the other hand, the prosodic focus of Azeri and its cross-comparison with Persian and Turkish are worth investigating.

In addition, since Persian has been in close contact with Arabic and even some part of the population (2%) in Iran is Arab, findings from the study of prosodic focus for Arab-speaking areas and its comparison with Persian and the other dialects of Arabic like Hijazi, Lebanese [4] and Egyptian [9] might reveal further implications.

5. Conclusions

Considering the results of the present perception experiment, it can be concluded that post-focus compression of F_0 and intensity is a highly important acoustic cue for the recognition of prosodic focus in Persian [30]. Its presence in non-final position leads to an average of over 78% focus recognition, whereas its absence in final position leads to less than 60% of focus recognition. Therefore, the overall recognition rate is 75% somewhere between the PFC languages like Beijing

Mandarin [5] and Uygur [26] with 90%, and languages without PFC like Taiwanese with 60% [5]. This perception experiment thus provides another piece of evidence that Persian, like English, Beijing Mandarin, Japanese, Turkish and Tibetan, can be categorized as a PFC language. This finding therefore adds yet another piece to the overall picture of focus typology.

6. Acknowledgements

We thank all subjects who helped us to carry out this research.

7. References

[1] Abolhasanizadeh, V., Bijankhan, M. and Gussenhoven, C., "The Persian pitch accent and its retention after the focus". Lingua 122, 1380-1394, 2012.

[2] Boersma, P., "Praat, a system for doing phonetics by computer". Glot International 5:9/10: 341-345, 1992-2013.

[3] Botinis, A., Fourakis, M. and Gawronska, B., "Focus identification in English, Greek and Swedish". In Proceedings of the 14th International Congress of Phonetic Sciences, San Francisco, 1557-1560, 1999.

[4] Chahal, D., "Phonetic cues to prominence in Lebanese Arabic". In Proceedings of the 15th International Congress of Phonetic Sciences, Barcelona. pp. 2067-2070, 2003.

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

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

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

[8] Eslami, M., Šenaxt-e næva-ye goftar-e zæban-e farsi væ karbord-e an dær bazsazi væ bazšenasi-ye rayane'i-ye goftar [The prosody of the Persian language and its application in computer-aided speech recognition]. PhD thesis, Tehran University, 2000.

[9] Hellmuth, S., "Focus-related pitch range manipulation (and peak alignment effects) in Egyptian Arabic". In Proceedings of Speech Prosody 2006, Dresden, Germany. pp. PS4-12-164, 2006.

[10] Hosseini, A., "L1 interference in L2 prosody: Contrastive focus in Japanese and Persian". Journal of Logic Language and Information 03/2013; 11:55-67, 2013.

[11] Ipek, C., "Phonetic realization of focus with no on-Focus pitch range expansion in Turkish". In Proceedings of the 17th International Congress of Phonetic Sciences, Hong Kong, 2011.

[12] Kahnemuyipour, A., *The syntax of sentential stress*. Oxford: OUP, 2009.

[13] Liu, F. and Xu, Y., "Parallel encoding of focus and interrogative meaning in Mandarin intonation", Phonetica 62: 70-87, 2005.

[14] 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.

[15] Mixdorff, H., "Quantitative tone and intonation modeling across languages". In proceedings of International Symposium on Tonal Languages, Beijing, 137-142, 2004.

[16] Pell, M. D., "Influence of emotion and focus on prosody in matched statements and questions", Journal of the Acoustical Society of America 109: 1668-1680, 2001.

[17] Perry, J., "Turkic-Iranian contacts: linguistic contacts," Encyclopedia Iranica, online edition, 2013, available at http://www.iranicaonline.org/articles/turkic-iranian-contacts-i-linguistic (accessed on 27 December 2013).

[18] Rump, H. H. and Collier, R., "Focus conditions and the prominence of pitch-accented syllables", Language and Speech 39: 1-17, 1996.

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

[20] Sahkai, H., Kalvik, M.-L. and Mihkla, M., "Prosody of contrastive focus in Estonian," in Proceedings of Interspeech 2013, Lyon, France, 315-319, 2013.

[21] Scarborough, R., "The intonation of focus in Farsi". UCLA working papers in phonetics, No. 105, pp. 19-34, 2007.

[22] Taheri-Ardali, M. and Xu, Y., "Phonetic realization of prosodic focus in Persian", in Speech Prosody 2012, Shanghai, 326-329, 2012.

[23] Towhidi, J., "Studies in the phonetics of Modern Persian. Intonation and related features" (=Forum Phoneticum 2). Hamburg: Helmut Buske [originally dissertation, London 1973], 1974.

[24] Vahidian-Kamyar, T., *Næva-ye goftar dær farsi [Melody of speech in Persian]*. Mashhad, Ferdowsi University Press, 2001.

[25] van Heuven, V. J., and de Jonge, M., "Spectral and temporal reduction as stress cues in Dutch," Phonetica 68, no. 3:120-132, 2011.

[26] Wang, B., Qadir, T. and Xu, Y., "Prosodic encoding and perception of focus in Uygur". Chinese Journal of Acoustics. 2013. (in Chinese)

[27] Xu, Y., "Effects of tone and focus on the formation and alignment of F_0 contours", Journal of Phonetics 27, 55-105, 1999.

[28] Xu, Y., "Speech melody as articulatory implemented communicative functions", Speech Communication 46, 220-251, 2005.

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

[30] Xu, Y., Chen, S.-w. and Wang, B., "Prosodic focus with and without post-focus compression (PFC): A typological divide within the same language family?" The Linguistic Review, 29, 2012.

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