Early peak: a case in Persian

Nasimeh Bahmanian[1], Moharram Eslami[2]

[1]Sharif University of Technology, Iran
[2]University of Zanjan, Iran
n.bahmanian@gmail.com, meslami@znu.ac.ir

Abstract

This study investigates the intonational properties of vocative structure in standard Persian in the framework of Auto-segmental-metrical (AM) phonology. In Persian, all lexical units are stressed in their final syllable and there is a language-specific feature in Persian according to which pitch accents are aligned with lexically stressed syllables. Contrary to the said feature, all previous studies claim that in vocative structure, it is the first syllable that is accented. The present study suggests the intonational pattern of ^H+H*H-(L/H)% for vocative structure in Persian, using Persian Tone and Break Indices (PToBI) annotations. Here ^H stands for early peak aligned with the second syllable, and H* stands for the pitch accent aligned with the lexically stressed, here the last syllable and H- stands for the phrase tone which signals the non-finality in the vocative utterance. Perceptual experiments support the hypotheses behind the study.

Index Terms: early peak, pitch accent, intonation, vocative structure, Persian.

1. Introduction

In Persian, lexical stress aligns with the last syllable, with only a few exceptions. Pitch accents, on the other hand, align with the lexically stressed syllables [1]. In spite of this fact, previous studies have introduced the first syllable of the nouns as the place of pitch accent in vocatives.

Different scholars have discussed briefly the issue of accent/stress in Persian vocatives ([2], [3], [4], [5], [6], and [7]). Fergusen believes that Persian is a language in which stress always falls on the last syllable with a few exceptions, and one of them is the vocative structure that is initially stressed/accented [2]. Vahidian-Kamyar states that noun stress moves to the first syllable, accompanied by a specific intonation [5]. Sadat-Tehrani has determined the first syllable as the most prominent syllable in Persian vocatives. Through a perceptual test, he has asked a few Persian native speakers to determine the most prominent syllable after hearing some recorded nouns in vocative and citation forms [6]. He claims that all the participants in the test have reported the first syllable in vocatives, and the last syllable in the citation forms as the most prominent. In his research, Sadat-Tehrani has examined three types of normal, angry, and surprised vocatives, and introduced L*+H h L% for normal and angry and L*+h l L% for surprised calls, in all of them L* aligns the first syllable. There is a contradiction between the result of above-mentioned perceptual test and the intonational pattern by Sadat-Tehrani for Persian vocatives, while the pattern starts with a low pitch. Eslami, however, believes that in vocatives, no change in the place of pitch accent occurs [7]. He suggests that there is a language-specific feature in Persian according to which, pitch accent

always aligns the lexically stressed syllable. He believes that "the only intonational difference between a vocative noun and its citation form is that the early non-stressed syllables of a vocative noun have higher pitch compared to its citation form". He does not explain more on this issue.

2. Theoretical Background

2.1. Stress vs. Pitch Accent

None of the researches prior to [7] distinguish the two abstract (stress) and concrete (pitch accent) concepts of prominence in Persian. In previous studies on Persian prosody such as [8], [2], [3], [5], [9], [4], and [10] some ideas like stress weakening, deletion, movement, and increase have been introduced, in order to explain the prominence in the level of utterance compared to the word level. However, one should clarify what is meant by stress in the first place, the abstract or the acoustic sense of prominence. Lexical stress never changes, but lexemes/word may or may not be accented in an utterance due to the context. "If a word in an utterance is accented, the prominence is always aligned with the lexically stressed syllable" [7].

Although Sadat-Tehrani has truly distinguished the above-mentioned concepts on prominence [6], he does not explain why the first unstressed syllable could carry pitch accent in vocatives. He solely relies on the native speakers' judgement of the most prominent syllable, although the place of pitch accent is part of unconscious knowledge of native speakers, and it is difficult to locate it consciously. In spite of the fact that in controversial cases speakers' judgment could not be the only evidence, we asked several native Persian speakers, chosen among Persian teachers to foreigners, to determine the most prominent syllable after hearing some vocative utterances, and their choices were not just the first syllable, but also the last syllable. The observation by Sadat-Tehrani might be just by coincidence or due to the small number of participants in his examination or even the data and the non-standard type of examination he has used.

2.2. Early Peak

Phonological approaches to intonation begun with [11], [12], and [13]. These scholars used abstract tones to represent speech intonation for the first time. Ladd, for the first time, called this theory the AM phonology [14]. Eslami applied this theory to Persian [7] and used a labelling system which he called PToBI [15], based on studies by [13], [16], [17], [18], and [19].

F0 peak alignment, which refers to relative distance of the accented syllable and the peak in F0 contour, is of an importance in phonological studies based on fundamental frequency. F0 peak sometimes does not align with the accented syllable. "Peak delay", for example, is a phenomenon that refers

to the occurrence of F0 peak after the corresponding syllable. This phenomenon has been reported in Persian clitic groups in [20]. Another similar phenomenon is "early peak" in which F0 peak precedes the associated syllable, and it has been reported in some languages such as German in [21], [22], and [23]. This study suggests that the F0 peak in Persian vocatives on the second syllable is a case of early peak, which is reported here for the first time. Using PToBI [15] the proposed transcription of pitch accent for Persian vocatives would be ^H+H* and the intonational pattern for vocatives would be ^H+H*H-(L/H)%.

3. Methodology

In order to find out the intonational pattern of Persian vocatives, six native speakers, three males and three females, were asked to utter 20 Iranian proper names in two citation and vocative forms. 15 out of 20 names were of 2 to 4 syllables words, and all composed of sonorant consonants in their segmental string, to avoid devoicing in the F0 contour and to minimize the consonantal effects on the stressed syllable. Finding one-syllabic names based on such criteria was impossible, and we used them only to explain the findings in other parts of the investigation. The proper nouns used in the study can be seen in Table 1. As for the vocative form, speakers were asked to read pre-written interrogative or imperative sentences with the selected nouns in the beginning, so that they produce vocative nouns. For citation forms, they were asked to read each noun as if they were answering to the questions like "what is your name?" or "who was on the phone?".

Table 1: *Proper nouns used in the dataset*

One-syllabic	Two-syllabic	Three-syllabic	Four-syllabic
giv	royâ	narimân	?amirali
nâz	lâle	mâniyâ	narimâni
râd	râyân	?âriyâ	?imâniân
roz	minâ	?âlâle	ruyâniân
râz	maryam	?âriyan	?amirâni

Since vocative feature is a linguistic category, if we agree on the categorical function of tonal elements as well, then vocative structure should have its own acoustic correlates. In the framework of AM phonology, an intonation phrase is composed of three main categorical elements: pitch accent(s), phrase tone, and boundary tone. In other words, what makes a noun in a vocative mood to be different from that of statement mood should be a set of specific intonational features. Any change, in the F0 contour of an utterance that gives to categorical change in the contextual meaning, is taken as a phonological feature; otherwise is not of phonological importance. The perceptual experiments support the hypotheses behind this study in two ways. First, by substituting F0 values of the assumed accented syllable by the values in its citation form pair in two-syllabic words. Second, by making gradual changes in F0 contour in a set of stimuli which was created in initial low pitch in the first syllable.

Praat (version 6.0.23) [23] was used as a complete phonetic laboratory tool in all stages.

4. Analyzing the Data

Persian speakers produce nouns in a prominently higher pitch in a vocative utterance rather than to their citation forms. The recorded data in this study show that F0 mean value of a vocative noun is approximately 20% higher than when that noun is uttered in its citation form (see Table 2). In addition, F0 changes are of greater variations in the case of vocatives. Figure 1 shows an example of this systematic variation that is observed in all the data recorded in the study. As it could be seen in Figure 1, a curve with the slightly highest F0 on the final syllable in the citation form, while in the vocative noun F0 value is rising at end of first syllable that reaches to highest point on the first half of the final syllable.

The general intonational pattern of vocative data shows that F0 peak always occur at the beginning of second syllable. Besides, F0 curve remains in the high range of speakers' pitch domain after the peak. Therefore, assuming the final syllable as the accented syllable, the peak at the beginning of second syllable could be described by the "early peak" phenomenon. In the framework of PToBI [15], we represent the early peak in vocatives by a high preceding tone (H) in second syllable which is followed by a high main tone (H*) aligned with the last stressed syllable. Since there is an upstepped preceding peak, we use ^H for the preceding tone. So, the pitch accent in Persian vocatives could be shown by ^H+H* tone.

It should be noted that pitch accent is followed by a high phrase tone which signals the non-finality of the utterance, and is shown by H-. The vocative data used show that boundary tone could be either L% or H% depending on the type of vocative, which is not the main focus of this study.

Table 2: *Percent of increase in F0 mean values in vocatives with respect to the citation forms*

Speaker	Number of Syllables	Mean of F0 increase
Female	2	18%
	3	20%
	4	19%
	Mean	19%
Male	2	18%
	3	23%
	4	22%
	Mean	21%
Overall Mean		20%

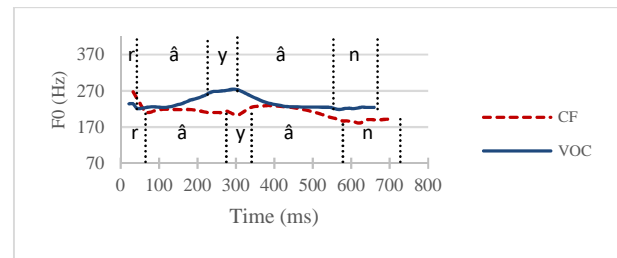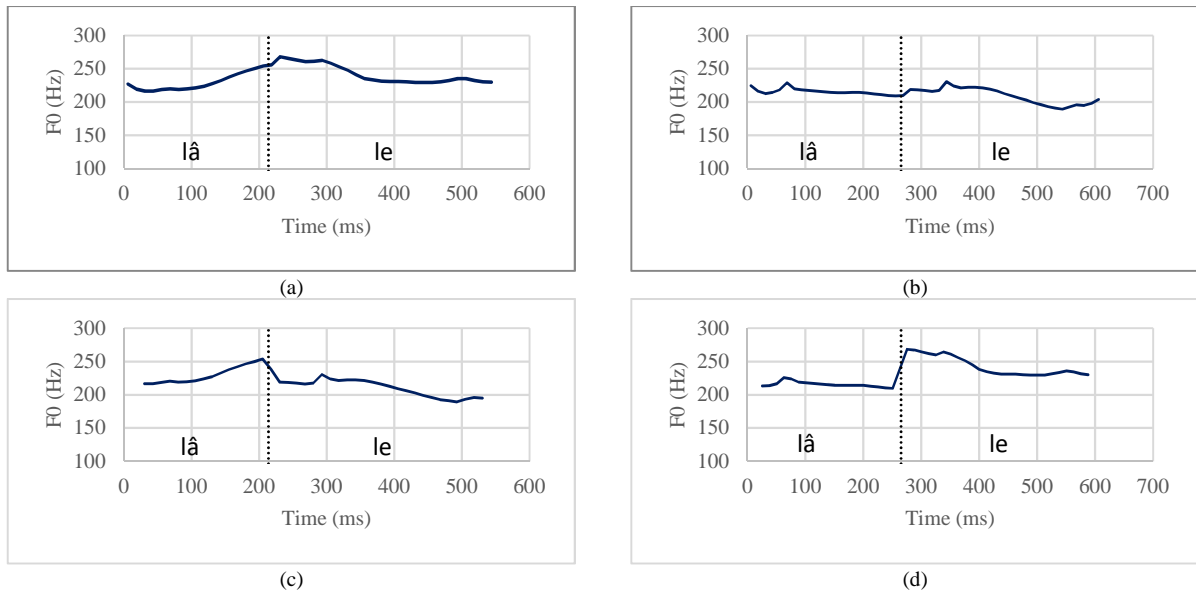Figure 1: *F0 contour of the two-syllabic noun /râyân/ in the vocative (VF) and citation form (CF) uttered by female speaker F2*

Figure 2: *Coupling the CF-VF and VF-CF tokens from two-syllabic noun /lâle/ uttered by female speaker F2 in Experiment 1: (a) vocative, (b) citation form, (c) coupled VF-CF token, (d) coupled CF-VF token*

5. Perceptual Experiments

5.1. Experiment 1

In the first experiment, we used all two-syllabic nouns in section 3. Each noun was recorded in the two vocative and citation forms. For each pair uttered by a single speaker we created two new two-syllabic tokens as shown in Figure 2, one token by coupling the first syllable of vocative noun with the second syllable of the citation form (VF-CF token), and another one by coupling the first syllable of citation form with the second syllable of vocative form (CF-VF token). The motivation behind this experiment was to find out which of the first and final syllables' F0 contour plays an important role in making an utterance into a vocative mood. In other words, we examined the quality of vocative-ness by eliminating a part of F0 contour of vocatives in each stimulus.

5.1.1. Stimuli

We had five nouns uttered by six speakers in two forms. Therefore, we obtained thirty VF-CF tokens and thirty CF-VF tokens as the stimuli in this experiment.

5.1.2. Procedures

Two subjects, one male and one female, participated in this experiment. They were both linguists and native Persian speakers. The sixty obtained tokens together with the ten original utterances (CF and VF) were randomly played for subjects. They were told to choose one of the CF/VF options on the screen for each utterance they were listening to, and then pick a goodness score of 1 (worst) to 5 (best) based on their choice. For this experiment, we used Experiment MFC, the perception experiment module of Praat [23]. Using this tool, we took the advantages of randomly playing sounds, assigning a number of occurrence for each sound, and measuring reaction time which is the time measured after a sound is played and before the listener choses an option on the screen.

Figure 3: *Number of CF vs. VF choices in experiment 1*

	CF	VF	CF	VF	CF	VF
	CF-VF		VF-CF		Unmanipulated	
Total	4	116	119	1	19	21

Figure 4: *Mean values of CF and VF scores and the average of response reaction times in Experiment 1*

	CF	VF	CF	VF	CF	VF
	CF-VF		VF-CF		Unmanipulated	
Average of goodness	3.75	4.39	4.09	5.00	4.58	4.81
Average of reactionTime	2.06	1.63	1.90	2.77	1.50	1.22

5.1.3. Results

In this experiment, subjects recognized 97% of CF-VF tokens as vocative with the score of 4.39, and 99% of VF-CF tokens as citation form with the score of 4.09. As it could be seen in Figure 3, one of the original CF utterances has been chosen as vocative. Besides, the average of goodness scores for the not manipulated utterances in Figure 4 shows that subjects' responses were conservative to some extent. This is also

reflected in the response reaction times, which is longer for the manipulated stimuli than the original utterances.

Experiment 2

In the second experiment, the main motivation was to understand the importance of low tone in the first syllable. We used the four-syllabic vocative noun /ʔamirali/ uttered by female speaker F1 in section 3, to obtain four stimuli by making gradual shifts in F0 contour up to the peak value in the first syllable.

5.1.4. Stimuli

Figure 5 shows the original F0 curve together with the four shifted curves of four stimuli.
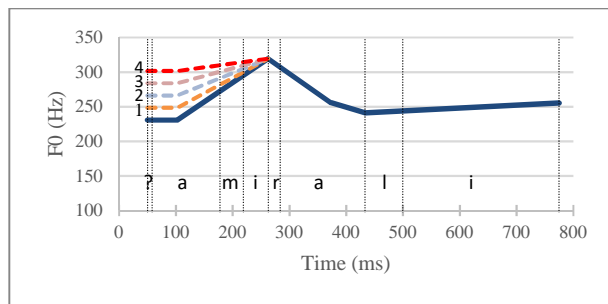
Figure 5: *Gradual shifting of F0 contour to develop stimuli in Experiment 2*

5.1.5. Procedures

Five subjects, three males and two females, participated in this experiment. They were all students of linguistics and native Persian speakers. The four obtained stimuli together with the original utterance, with the number of occurrence three were randomly played for subjects. They were told to choose one of the Natural/Unnatural options on the screen for each utterance they were listening to, and then give a goodness score of 1 (worst) to 5 (best) as a degree of vocative-ness. For this experiment, we used Experiment MFC tool from Praat which is introduced in section 5.1.2.

5.1.6. Results

The average score of 4.07 for the original vocative utterance shows that all scores are conservative to a certain extent (See Figure 6). Although it was not synthesized, Subjects' score to the original not-manipulated sound is 4.07, which is too close to their scores to the manipulated sounds. It means that they could not recognize a significant change in the vocarive-ness of the sound when the initial low tone changes into a high tone. Figure 7, on the other hand, represents the cognition of subjects on the naturalness of the utterances. As it could be seen all responses have been highly conservative, so that even the not-manipulated sound has been recognized as unnatural nine times out of fifteen occurrences. It could be inferred that speakers' judgement of the degree of naturalness is similar for the manipulated tokens and otherwise, and this shows that changes were made in the first syllable in this experiment did not have a major effects on the vocative characteristic perceived by speakers. This conclusion, together with the average score of vocative-ness, rejects the phonological value of the L tone preceding the peak in the second syllable, although it has been considered a categorically important element in Persian vocatives by Sadat-Tehrani [6].
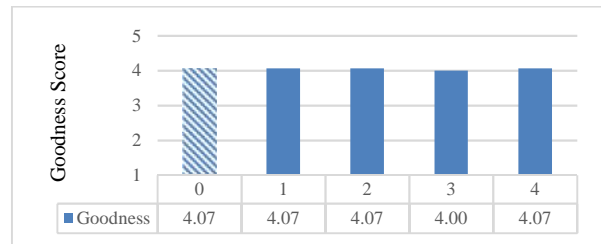

Figure 6: *Mean values of vocative-ness scores of the responses in Experiment 2 for the utterances No. 0 (original not-manipulated sound) and No. 1 to 4 (manipulated sounds) as shown in Figure 5.*
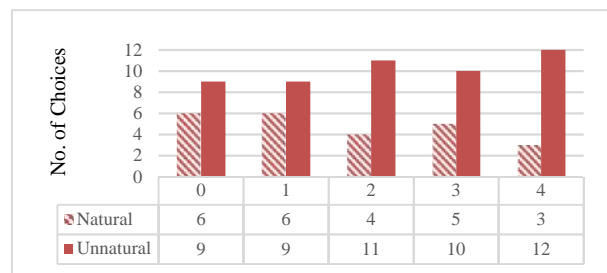

Figure 7: *Number of natural vs. unnatural choices in experiment 2 for the utterances No. 0 (original not-manipulated sound) and No. 1 to 4 (manipulated sounds) as shown in Figure 5.*

6. Discussion

Since the intonation and pitch accent convey a pragmatic and post lexical meaning, the results of above experiments raise serious doubts about the first syllable as the accented syllable in Persian vocatives. If a post lexical meaning, here vocative-ness, is not affected by changing part of the F0 contour, it could be concluded that there is no intonation element in that part of the utterance. On the contrary, when post lexical meaning is affected by changing part of the intonation contour, we should look for the intonation elements through that part of the utterance. Experiment 1 shows that all the information about the vocative category of two-syllabic nouns are in the last syllable preceded by an early peak in the first part of second syllable. This is the evidence for the fact that lexically stressed syllable is the only place of pitch accent in Persian vocatives. Experiment 2 refuses to accept the L+H* alignment which is proposed by Sadat-Tehrani [6], and proves that the assumed L tone is not a valid element in Persian vocatives.

7. Conclusions

In conclusion, the perceptual experiments confirm the hypothesis proposed in section 4, where we suggested that the phenomenon of proceeding "early peak" exists in Persian vocatives. This characteristic can be seen through a flat F0 contour in one-syllabic vocative nouns, and this shows that tonal alignment of the two to four-syllabic nouns is done properly in this research. This study rejects the previous claims introducing the first syllable as the accented syllable in Persian vocatives, and demonstrates that the pitch accent aligns with the lexically stressed syllable in such structures as proposed by Eslami [7].

8. References

[1]   M. Eslami, *vaj-shenasi: tæhlil-e nezam-e ahæng-e zæban-e farsi,* *[Phonology: Analysis of Persin intonation]*. Tehran: SMT press, 2005.

[2]   C. A. Ferguson, "Word stress in Persian," *Language,* vol. 33, no. 2, pp. 123-135, 1957.

[3]   C. T. Hodge, "Some aspects of Persian style," *Language,* vol. 33, no. 3, pp. 355-369, 1957.

[4]   Y. Samareh, *A course in Colloquial Persian*. Tehran: Tehran University Press, 1977.

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

[6]   N. Sadat-Tehrani, "The intonational grammar of Persian," PhD, University of Manitoba, Doctoral dissertation, 2007.

[7]   M. Eslami, "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]," Tehran: Tehran University PhD thesis*,* 2000.

[8]   K. S. Lambton Ann, *Persian Grammar,* ed: Cambridge University Press, 1953.

[9]   J. Towhidi, *Studies in the phonetics and phonology of modern Persian: intonation and related features*. Buske, 1974.

[10]   A. Kahnemuyipour, "Syntactic categories and Persian stress," *Natural Language & Linguistic Theory,* vol. 21, no. 2, pp. 333-379, 2003.

[11]   M. Y. Liberman, "The intonational system of English," Massachusetts Institute of Technology, 1975.

[12]   G. Bruce, *Swedish word accents in sentence perspective*. LiberLäromedel/Gleerup Malmo, 1977.

[13]   J. B. Pierrehumbert, "The phonology and phonetics of English intonation," Massachusetts Institute of Technology, Doctoral dissertation, 1980.

[14]   D. R. Ladd, *Intonational Phonology*. Cambridge University Press, 1996.

[15]   M. Eslami, "PToBI: A phonological system in transcribing the intonation of Persian," in R. Hoffmann (Ed.) Elektroniche Sprachsignalverarbeiung, *Proceedings of the 17th Conference of Electronic Speech Signal Processing (ESSV2006)* Freiberg, 2006, pp. 45-53: TUpress.

[16]   D. R. Ladd, "Phonological features of intonational peaks," *Language,* pp. 721-759, 1983.

[17]   E. Beckman Mary and B. Pierrehumbert Janet, "Intonational structure in English and Japanese," *Phonology Yearbook,* vol. 3, pp. 255-310, 1986.

[18]   P. J. Price, M. Ostendorf, S. Shattuck Hufnagel, and C. Fong, "The use of prosody in syntactic disambiguation," *the Journal of the Acoustical Society of America,* vol. 90, no. 6, pp. 2956-2970, 1991.

[19]   C. W. Wightman, S. Shattuck Hufnagel, M. Ostendorf, and P. J. Price, "Segmental durations in the vicinity of prosodic phrase boundaries," *The Journal of the Acoustical Society of America,* vol. 91, no. 3, pp. 1707-1717, 1992.

[20]   M. Bijankhan and V. Abolhasanizadeh, "dirkærde qolleye heja dær saxte næva'i-ye vajebæst-ha-ye farsi [peak delay in Persian clitiques]," *Alzahra University journal of linguistics,* vol. 5, 1390.

[21]   K. J. Kohler, *Einführung in die Phonetik des Deutschen,* E. Schmidt, 1995.

[22]   T. Rathcke and J. Harrington, "The variability of early accent peaks in Standard German," *Laboratory Phonology,* vol. 10, pp. 533-555, 2010.

[23]   S. Baumann, M. Grice, and R. Benzmüller, "GToBI-a phonological system for the transcription of German intonation," 2000, pp. 21-28.