The Structure of Persian Intonation

Nima Sadat-Tehrani

Department of Linguistics
University of Manitoba, Canada
nisate@yahoo.com

Abstract

This paper is a detailed investigation of the phonology and phonetics of the intonation of Persian carried out in the framework of the AM theory of intonational phonology ([7], [5]). Based on 2112 utterances read by a total of 8 native speakers, the work, on the one hand, presents a phonological account of the prosodic structure of this language, a structure that consists of the level Accentual Phrase (AP) with the pitch accent (L+)H*, immediately dominated by the level Intonational Phrase (IP), each level being marked by a low or high right boundary tone. This system is less complicated than previous proposals which consider an additional level between IP and AP. On the other hand, this work scrutinizes the phonetic implementation of tones with regard to segments. It is shown that the L of an AP is aligned with the consonant preceding the stressed vowel, and the H with the consonant following this vowel in nuclear APs and with the next vowel in non-nuclear ones. Focused APs have more duration than non-focused ones, and non-nuclear APs show more pitch excursion than nuclear ones.

1. Introduction

[4] classifies Persian as a 'stress-accent' language, i.e., a language in which a certain syllable in a word is more prominent than other syllables by phonetic factors, showing syntagmatic contrast. Pitch accents are associated with the stressed syllable ([3]), which is the final syllable for nouns (*šuné* 'comb'), adjectives (*kutáh* 'short'), and most adverbs (*yæváš* 'slowly'), and the final syllable of the main constituent for verbs (*xæríd-æm* 'I bought'), with verbal prefixes attracting the stress (*mí-xærid-æm* 'I'd buy'). Previous studies ([6], [8]) suggest that the prosodic structure of Persian consists of three levels: the Accentual Phrase, the Intermediate Phrase, and the Intonational Phrase. In the present paper, I propose a new model of Persian intonation and I argue that the level Intermediate Phrase is unwarranted for Persian and the Accentual Phrase (AP) and the Intonational Phrase (IP) are sufficient to account for the intonational structure of this language.

In the second part of the paper (Section 3), it is shown how the phonology of Persian contours is phonetically realized. The phonetic implementation concerns the alignment of the valley and the peak, the duration, and the pitch excursion of an AP.

2. Persian prosodic structure

Based on the recordings done for this paper, there is a recurrent tonal/accentual pattern for all utterances in Persian. The pattern, which following [6] I will call the Persian Accentual Phrase, or AP, consists of a low tone (L) followed by a high tone (H) forming the pitch accent L+H*, which is associated with the stressed syllable. The valley and the peak of the Persian AP are aligned before and after the stressed vowel (see Section 3 for details).

There are two allophones for this pitch accent: L+H* and H*, each of which occurs in a particular context. The most common allophone is L+H* which is for words or phrases with final stress, e.g., nouns and adjectives, longer than one syllable, and also for vocatives. Initially-stressed words, e.g., most verb forms, and monosyllabic content words have the allophone H*. Utterance initial APs usually take the form of the first allophone regardless of their stress pattern due to the occurrence of an utterance-initial rise in Persian.

The part of an Accentual Phrase between the pitch accent and the AP end is handled by a boundary tone, which can be high or low, named here as h and l. Thus, the boundary tone marks the right edge of an AP. This part can consist of zero syllables (when the stressed syllable of an AP is its final syllable), in which case the boundary tone is realized on the stressed syllable itself. It can also consist of several syllables, in which case the boundary tone is realized on all these syllables up to the AP end. The motivation for the existence of the AP boundary tone comes from the comparison of APs that have the nuclear pitch accent (NPA) and those that do not: in most types of simplex unmarked sentences, the NPA AP, which is the last AP, takes the l boundary tone, and other APs (which are not nuclear) take the h counterpart.

One or more Accentual Phrases are immediately dominated by an Intonational Phrase, which corresponds to an utterance for simplex sentences. An IP is phonologically marked by a boundary tone L% or H% on the final syllable or part of the final syllable. Phonetically, IPs are usually accompanied by pitch resetting at the beginning and a pause and sometimes vowel lengthening at the end. There is usually one nuclear pitch accent in every IP. L% is used for declaratives (SOV or scrambled), leading yes/no questions, WH-questions, alternative questions, imperatives, and vocatives. H% is used for yes/no questions, tag questions, echo questions, coordinate structures, and IP-forming subordinate clauses.

Persian prosodic structure is exemplified below. The acute accent marks the stressed syllable of an AP and the NPA AP is underlined. The voice analysis software used is Praat [2].

miná	milán-æm	<u>mí-mun-e</u>	čænd	ruz.	(1)
Mina	Milan-too	PRFX-stay-3SG	a few	day	

'Mina stays a few days in Milan too.'

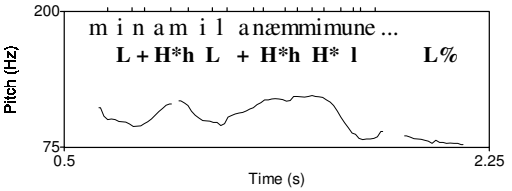

Figure 1: *The utterance 'Mina stays a few days in Milan too.'*

In (1), there are three APs corresponding to the noun Mina, the adverb Milan and its clitic *–æm*, and the verb. The first two carry the pitch accent L+H*, and the initially-stressed verb (*mi-mun-e*) bears the H* allophone. The first two APs are high-boundary-toned and the last is the nuclear AP and low-boundary-toned. The declarative is realized as one IP (and one utterance) marked by L%, indicating that it ends low with no pitch change involved. Everything following the NPA is deaccented up to the IP end, so there is no tonal event in the phrase *čænd ruz*.

An Accentual Phrase normally consists of one content word with its possible clitic(s). The clitic usually carries the AP boundary tone, and in cases where this boundary tone is high, the clitic is realized at a higher pitch than the previous H. The configuration of an AP may be affected by factors such as speech rate, focus, length, subordination, and information structure. For instance, to see the effect of the latter, consider the utterance in (2) and the two possible realizations of its subject noun phrase in Figure 2.

mærdom-e	inja	mehræbun-æn.	(2)
people-LINKER	here	kind-are	

'The people here are very kind.'

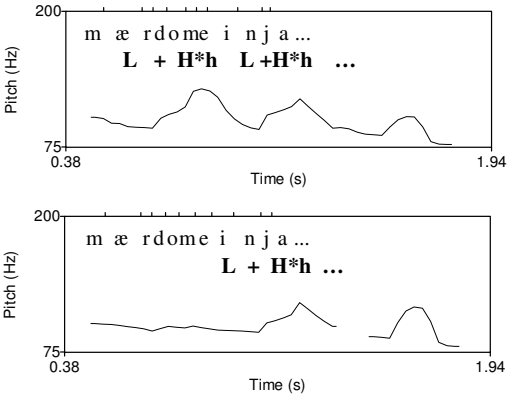

Figure 2: *The two possible realizations of 'the people here.'*

In the first version, the noun phrase is new information and shows two rises. In the second, it is given information (e.g., used in response to the question 'How do you find the people here?'), hence realized as one AP. In such cases, i.e., where the L of an AP is realized on several syllables before the H*, a low plateau is formed (on *mærdom-e in-* in the above example) which is the result of the spreading of the L to the left up to the beginning of the AP.

A contrastive focus element forms its own AP which behaves phonologically in the same way as an ordinary AP, i.e., it has the tonal pattern of (L+)H*. It is low-boundary-toned and causes deaccentuation up to the utterance end. Phonetically, a focused AP has more pitch excursion and longer duration than the non-focused counterpart (see Section 3).

The proposed prosodic system is less complicated than previous proposals which suggest an additional phrase accent for Persian ([6], [8]). I argue that the AP boundary tone can in all instances account for the part of the F0 contour following the NPA, i.e., the part that the phrase accent is supposed to associate with. Thus, the following two situations do not occur after the NPA AP in Persian: a low AP boundary tone followed by a high stretch and a high AP boundary tone followed by a low plateau. In other words, whatever boundary tone the NPA AP has will persist up to the IP end, i.e., if it is low, it will stay low (Example (1) above), and if it is high, it will remain high (Example (3) below).

arezú	čí-ro	ru	miz	gozašt?	(3)
Arezu	what-OBJ	on	table	put.PAST.3SG	

'Arezu put what on the table?' [echo question]

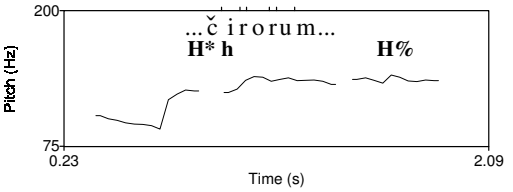

Figure 3: *The echo question 'Arezu put what on the table?'*

3. Phonetic implementation experiment

This section contains the report of an experiment carried out to determine the phonetic characteristics of Persian intonation and also to see the impact of contrastive focus in this regard. To this end, three different types of cliticized APs were compared. The first type had a high boundary tone and was not the nuclear pitch accent of the utterance. The second *was* the nuclear pitch accent marked by a low boundary tone. In the third type, the same AP as in the other two types was contrastively focused and ended in a low boundary tone. In what follows, we refer to the above three types as Non-nuclear, Nuclear, and Focus types respectively. The comparison concerned the alignment of L and H relative to certain segmental landmarks, the pitch difference of the valley and the peak, and the duration of the segments and of the whole AP. The APs under investigation excluded those that have a low plateau caused by a leftward spreading rule (as in the second pitch track of Figure 2) and so are all APs realized with an observable valley.

3.1. Method

3.1.1. Design and materials

30 utterances and 4 speakers were used in this experiment, which made a total of 120 samples. The 30 utterances consisted of 10 utterances from each type. A sample set of utterances containing the Non-nuclear, Nuclear, and Focus

types is given in (4), (5) and (6) respectively, in which the test AP is italicized and contrastive focus is shown with capitalization.

dirúz *namæ-mun* umæd. (4)
yesterday letter-our come.PAST.3SG
'Yesterday our letter arrived.'

mál-e *namæ-mun* bud. (5)
property-LINKER letter-our be.PAST.3SG
'It was for our letter.'

mál-e *NAMÆ-MUN* bud. (6)
property-LINKER letter-our be.PAST.3SG
'It was for OUR LETTER.'

The elicitation of the utterances was done by showing them to speakers on cue cards in random order. The speakers were asked to use a normal speech rate, neither fast nor slow, and to read each utterance once. For the third type, a prompter question was asked to render the test Accentual Phrase contrastively focused. Thus, every time a contrastive focus utterance came up on a card, the experimenter asked the speaker a question whose answer was the utterance on the card. For example, the question for the focus utterance in (6) was (7).

mál-e ketáb-etun bud? (7)
property-LINKER book-your be.PAST.3SG
'Was it for your book?'

The test AP (e.g., *namæ-mun* 'our letter' above) was designed to be always preceded and followed by other syllables in the utterance in order to reduce the unwanted utterance-initial and utterance-final effects. The consonants on the sides of the stressed vowel in the test word were always the sonorants /l/, /m/, /n/ in order for the F0 track of the test word to be smooth and connected.

3.1.2. Speakers

Four speakers, two female and two male, took part in this experiment, one of the males being the author. Their age ranged from 27 to 41, all educated native speakers of Persian who spoke Tehrani dialect and who had lived in Iran all their lives before moving to Canada 3 to 6 years ago. They had been using Persian in some of their daily communications since they left Iran.

3.1.3. Procedure

The productions of the speakers were recorded by a Marantz PMD660 professional digital voice recorder using a Shure KSM109 cardioid condenser microphone placed at a fixed distance of about 40 cm from the speaker. The recordings were input to the Praat software. The test Accentual Phrases always had the pattern L+H*. The recordings of those utterances whose valley and peak were not fully observable (about 7.5 % of the recordings) were discarded and such utterances were re-recorded in a second session. The measurement methodology was basically that of [1]. The following eight landmarks were selected in each test AP:

- V1: The beginning of the vowel preceding the stressed vowel (V_{pre})
- C1: The beginning of the consonant preceding the stressed vowel (C_{pre})
- V2: The beginning of the stressed vowel (V_{str})
- C2: The beginning of the consonant following the stressed vowel (C_{post})
- V3: The beginning of the vowel of the clitic (V_{cli})
- C3: The end of V_{cli}
- L: The location of F0 minimum
- H: The location of F0 maximum

These landmarks help us determine the location of the maximum and the minimum of the test AP with regard to the segments. Figure 4 shows an example of the location of the above landmarks for the Nuclear type sentence in (5).

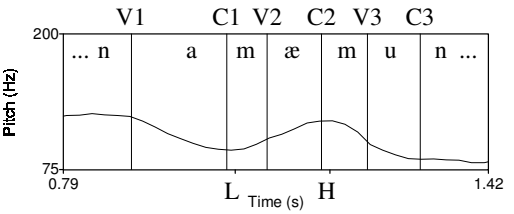

Figure 4: *Landmarks for* namæ-mun *'our letter' as a nuclear AP.*

3.2. Results and discussion

The mean amounts of AP duration, alignment values, and normalized pitch excursion taken among all speakers are given in Table 1.

AP type	AP dur. (ms)	LV1 (ms)	LC1 (ms)	HC2 (ms)	HC3 (ms)	H-L norm.
Non-nuc	443.8	129.0	27.2	78.0	-32.1	0.202
Nuclear	483.8	124.7	15.1	12.6	-109.3	0.129
Focus	544.1	131.9	6.3	7.7	-122.5	0.210

Table 1: *Mean AP duration, alignment values, and normalized pitch excursion.*

A repeated-measures ANOVA verified significant difference in the AP durations ($F(2,6) = 11.65$, $p = 0.009$). In order to detect where the difference was, a Tukey-Kramer multiple comparison test was performed and it was revealed that it is between the Non-nuclear and the Focus AP. This means that the duration of a focused Accentual Phrase is significantly longer than that of the Non-nuclear counterpart. The duration of the Nuclear type is between those of the other two types but not significantly different from either. The difference between the duration of AP types is caused by the duration of vowels (V_{pre}: $F(2,6) = 8.70$, $p = 0.02$; V_{str}: $F(2,6) = 12.75$, $p = 0.007$; V_{cli}: $F(2,6) = 31.16$, $p = 0.0007$) and not consonants (C_{pre}: $F(2,6) = 1.47$, $p = 0.30$; C_{post}: $F(2,6) = 0.16$, $p = 0.86$).

The alignment of L and H was measured with regard to two reference points: V1 and C1 for L, and C2 and C3 for H. The alignment of L occurs in the consonant preceding the stressed vowel in all three types. A repeated-measures ANOVA determined significant difference ($F(2,6) = 7.25$, $p =$

0.03) relative to the reference point C1, and Tukey-Kramer multiple comparison tests showed that the difference lies between the Non-nuclear and the Focus types: in the former the valley is halfway through the consonant and in the latter at its onset. With regard to the reference point V1, the alignment difference of L does not reach significance level ($F(2,6) = 0.42$, $p = 0.68$). This is due to the fact that V1 is a farther point to L than C1. Based on [9], who suggest that it is preferable to measure alignment relative to a nearer variable, the results from C1 reference point can reflect the reality more than those from V1. The alignments of H are significantly different relative to both points (HC2: $F(2,6) = 93.11$, $p = 0.00003$, HC3: $F(2,6) = 61.51$, $p = 0.0001$). Tukey-Kramer tests revealed that the Non-nuclear type is different from the other two. Segmentally, H is aligned with the middle of the vowel of the clitic in the Non-nuclear type but in the beginning of the preceding consonant in the other two types. In other words, the AP type with a high boundary tone has a later alignment of H. Figure 5 contains a schematic representation of length and alignment properties of the three AP types.

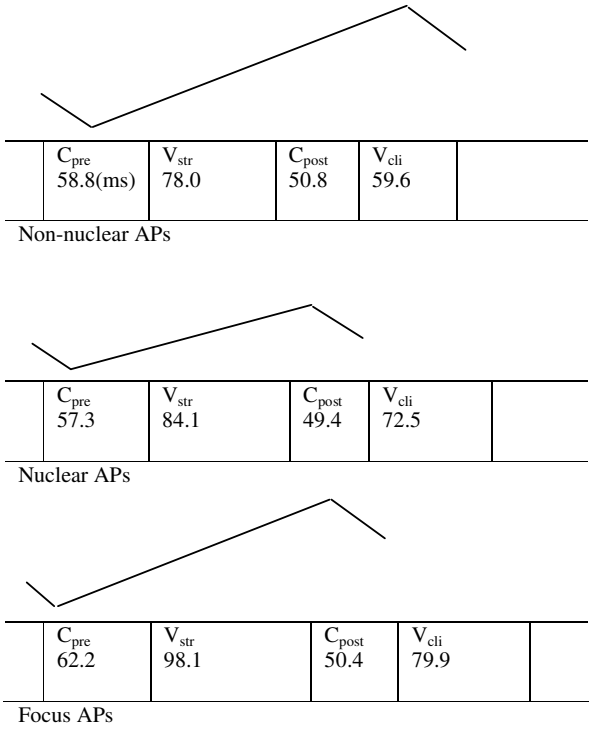

Figure 5: *Comparison of alignment in the three AP types*

The mean Ls are quite close to one another (164, 163, 162 Hz) and are not significantly different ($F(2,6) = 0.30$, $p = 0.75$). The Hs show more variation: 199, 184, 200 Hz. The variations do not reach significance level but approach it ($F(2,6) = 4.37$, $p = 0.067$). The difference in the three types becomes significant in the normalized excursion, that is, the difference between H and L divided by their mean ($F(2,6) = 14.71$, $p = 0.005$). The difference is between the Nuclear type and the other two types (Tukey-Kramer test). The higher excursion of the Focus and the Non-nuclear types is justified: in the former the more change in pitch fluctuation is a way to make the contrastively focused AP more prominent, and in the latter the higher excursion is caused by the high AP boundary tone.

4. Conclusions

This paper proposed a prosodic structure for Persian. This structure consists of Accentual Phrase as the smallest unit of prosody for this language, characterized by the pitch accent L+H*, and immediately dominated by Intonational Phrase. Both levels are marked at the right edge by a low or high boundary tone. The system suggested in this paper is less complicated than previous proposals which considered the level Intermediate Phrase between the two. The present work also looked at the phonetic properties of the Persian AP. It was shown that the duration of the focus AP is longer than that of the non-focused counterparts. This difference is caused only by vowels. The L is always aligned in the consonant preceding the stressed vowel but is significantly earlier in the Focus type. The alignment of H is in the vowel of the clitic following the stressed syllable in Non-nuclear APs (which have a high boundary tone) and in the preceding consonant in the other AP types (which have a low boundary tone). The former is significantly later than the latter. The pitch excursion of the nuclear AP is significantly less than that of the non-nuclear and the focus AP. This difference is caused only by the different highs and not by lows.

5. References

[1] Atterer, M., Ladd, D.R. 2004. On the phonetics and phonology of "segmental anchoring" of F0: evidence from German. *Journal of Phonetics* 32, 177-197.

[2] Boersma, P., Weenink, D. 2007. Praat: Doing phonetics by computer (Version 4.5.24) [Computer program], retrieved from: http://www.praat.org/

[3] Eslami, M., Bijankhan, M. 2002. Nezam-e ahæng-e zæban-e farsi [Persian intonation system]. *Iranian Journal of Linguistics* 34, 36-61.

[4] Jun, S.-A. 2005. Prosodic typology. In: Jun, S.-A. (ed.), *Prosodic typology: The phonology of intonation and phrasing*. Oxford: Oxford University Press, 430-458.

[5] Ladd, D.R. 1996. *Intonational phonology*. Cambridge: Cambridge University Press.

[6] Mahjani, B. 2003. 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

[7] Pierrehumbert, J. 1980. The phonology and phonetics of English intonation. Ph.D. thesis, MIT, published 1988 by Indiana University Linguistics Club.

[8] Scarborough, R. 2007. The intonation of focus in Farsi. UCLA working papers in phonetics, No. 105, pp. 19-34, retrieved from: http://www.linguistics.ucla.edu/faciliti/workpapph/105/2-Farsi%20Focus%20-%20for%20WPP.pdf

[9] Schepman, A., Lickley, R., Ladd, D.R. 2006. Effects of vowel length and "right context" on the alignment of Dutch nuclear accents. *Journal of Phonetics* 34, 1-28.