

THE INTONATIONAL GRAMMAR OF PERSIAN

Nima Sadat-Tehrani

Linguistics Department, University of Manitoba

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 ([6], [4]). 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 with the pitch accent (L+)H* immediately dominated by the level Intonational Phrase, each level being marked by a low or high boundary tone. On the other hand, it scrutinizes the phonetic implementation of tones with regard to segments and shows that the L of an Accentual Phrase is aligned with the consonant preceding the stressed vowel, and the H with the consonant following this vowel in nuclear Accentual Phrases and with the next vowel in non-nuclear ones.

Keywords: Intonational phonology; Persian; Prosody; Alignment; Pitch track

1. INTRODUCTION

[3] 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 ([2]), which is the final syllable for nouns (*šuné* ‘comb’), adjectives (*kutáh* ‘short’), and most adverbs (*yaváš* ‘slowly’), and the final syllable of the main constituent for verbs (*xærid-æm* ‘I bought’), with verbal prefixes attracting the stress (*mí-xærid-æm* ‘I’d buy’). Previous studies ([5], [7]) 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 dissertation, there is a recurrent tonal/accental pattern for all utterances in Persian. The pattern, which following [5] 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 includes 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

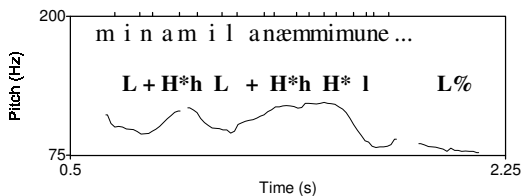
the l boundary tone, and other APs (which are not nuclear) take the h counterpart.

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.

miná milán-æm mí-mun-e čænd ruz. (1)
 Mína 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.’



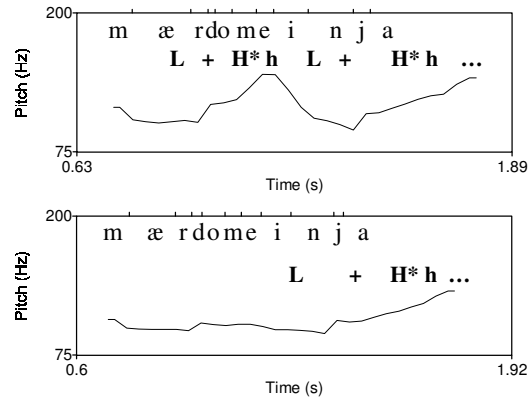
In (1), there are three APs corresponding to the noun Mina, the adverb Milan and its clitic *-æm*, and the verb. The first two carry the pitch accent L+H*, and the initially-stressed verb (*mi-mun-e*) bears the H* allophone. The first two APs are high-boundary-toned and the last is the nuclear AP and 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.

mærdom-e inja xeyli mehræbun-æn. (2)
 people-LINKER here very 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.

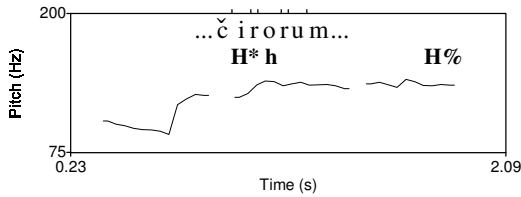
A contrastive focus AP behaves phonologically in the same way as an ordinary AP, i.e., it has the tonal pattern of (L+)H*. It is low-boundary-toned and causes deaccentuation up to the utterance end. Phonetically, a focused AP has more pitch excursion and longer duration than the non-focused counterpart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

The proposed prosodic system is less complicated than previous proposals which suggest an additional phrase accent for Persian ([5], [7]). 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).

arežú 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

This section reports the results 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 difference in pitch amounts 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.

30 utterances and 4 speakers (2 female, 2 male, age range 27-41) 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 type is given in (4), (5) and (6) respectively, in which the test AP is italicized and contrastive focus is shown with capitalization.

dirúz *namá*-mun umád. (4)
 yesterday letter-our come.PAST.3SG
 ‘Yesterday our letter arrived.’

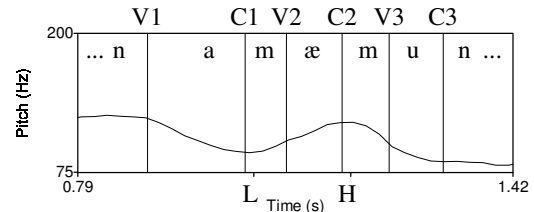
mal-e *namá*-mun bud. (5)
 property-LINKER letter-our be.PAST.3SG
 ‘It was for our letter.’

mal-e *NAMÁ*-MUN bud. (6)
 property-LINKER letter-our be.PAST.3SG
 ‘It was for OUR LETTER.’

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

Figure 4: Landmarks for *namá-mun* ‘our letter’ as a nuclear AP



A repeated-measures ANOVA on the means of speakers showed that the durations of the three AP types are different ($F(2,6) = 11.65, p = 0.009$). In order to detect where the difference was, a Tukey-Kramer multiple comparison test was performed and it was revealed that the difference 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 [8], 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.

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. CONCLUSION

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 a previous proposal 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 more 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 determined by the AP boundary tone: if it is high, the H is aligned in the vowel of the clitic following the stressed syllable, and if it is low, the H is aligned in the preceding consonant. 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.

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] Eslami, M., Bijankhan, M. 2002. Nezam-e ahæng-e zæban-e farsi [Persian intonation system]. *Iranian Journal of Linguistics* 34, 36-61.
- [3] 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.
- [4] Ladd, D.R. 1996. *Intonational phonology*. Cambridge: Cambridge University Press.
- [5] 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/archiv/e/dissertations/2002-3/behzad_mahjani.pdf
- [6] Pierrehumbert, J. 1980. The phonology and phonetics of English intonation. Ph.D. thesis, MIT, published 1988 by Indiana University Linguistics Club.
- [7] Scarborough, R. 2003. The intonation of focus in Farsi. UCLA working papers in phonetics, No. 105, retrieved from: http://www.stanford.edu/~rscar/index_files/Farsi%20Intonation.pdf
- [8] 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.