View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by Radboud Repository

Radboud Repository

Radboud University Nijmegen

# PDF hosted at the Radboud Repository of the Radboud University Nijmegen

The following full text is an author's version which may differ from the publisher's version.

For additional information about this publication click this link.
http://hdl.handle.net/2066/86013

Please be advised that this information was generated on 2017-12-06 and may be subject to change.

# The position of clitics in Persian intonational structure

*Vahideh Abolhasani Zadeh*[1,2], *Carlos Gussenhoven*[2,3], *Mahmood Bijankhan*[1]

[1]Department of Linguistics, University of Tehran, Iran
[2]Department of Linguistics, Radboud University Nijmegen, Netherlands
[3]School of Languages, Linguistics and Film, Queen Mary University of London, UK

vahidehabolhasani@yahoo.com, c.gussenhoven@let.ru.nl, mbjkhan@ut.ac.ir

**Abstract**

Persian clitic groups differ from words. Most importantly, a pitch accent (L+)H* is associated with the word-final (i.e. base-final) syllable of clitic groups, but with the word-final syllable of words, meaning that clitics remain outside the domain of the word. The pitch accent marks the stress, but we found no independent durational or spectral differences between stressed and unstressed syllables. Interestingly, the intonational distinction between words and clitic groups remains intact in the stretch of speech after the focus. Unlike Germanic, Persian post-focal words are accented, though pronounced with reduced pitch range.

Index terms: **clitic group, phonological word, prosodic hierarchy, focus, pitch range**

## 1. Introduction

The aim of this paper is to introduce the clitic group (CG) as a level in the prosodic hierarchy of Persian. The hybrid nature of clitics, intermediate between affixes and words, is reflected in the phonology by assuming a distinct constituent between the phonological word, which groups affixes with stems, and the phonological phrase, which groups word with other words.

The motivation for assuming the CG is the assignment of a pitch accent to final syllables of phonological words. This rule skips right-edge clitics, but not suffixes. We illustrate this in (1a,b,c,d), where (1a) are two isolated words, (1b) two suffixed words, (1c) two words with a clitic, and (1d) a compound. As these data show, words and suffixed words have final accented syllables, compounds lose accent on their first constituent, while clitics are not assigned accent, causing the accent in cliticized words to be on the syllable before the clitic. The CG is required to account for (1c).

| | | | | |
|---|---|---|---|---|
| (1) | a. ke'tab | 'book' | xu'ne | 'house' |
| | b. ketab-'ha | 'books' | xune-'ha | 'houses' |
| | c. ke'tab-i | 'one book' | xu'ne-j-i | 'one house' |
| | d. ketabxu'ne | 'library' | | |

A second motivation for recognizing the CG is provided by SYNCOPE. This rule applies to word-final vowels before a vowel-initial clitic. While word-final vowels before vowel-initial suffixes are preserved, triggering the insertion of [j] to break the hiatus, word-final vowels before clitic-initial vowels are deleted, causing the loss of a syllable. This is illustrated in (2). This vowel deletion does not occur in the domain of phonological word and phonological phrase.

(2) a. [[dɒneʃ'ʤu] + [æm]]$_{CG}$ → [dɒneʃ'ʤum]$_{CG}$
student my 'my student'

b. [[dɒneʃ'ʤu]$_{PW}$ + [ɒn]$_{PW}$ → [dɒneʃʤu'jɒn]$_{PW}$
(*[dɒneš'ʤun])
student PL 'students'

c. [[dɒneʃ'ʤu] e]$_{CG}$ [ʔæ'vaz]]$_{PP}$ → [dɒneʃ'ʤu ʔæ'vaz]$_{PP}$
( *[dɒnešʤu'vaz])
student of æhvaz 'æhvaz student'

## 2. Stress in Persian

Persian stress assignment is quantity-insensitive. Final syllables of nouns, adjectives, most adverbs and unprefixed verbs have the stress [4], [7], and [12]. Prefixed verbs take stress on the prefix. [6] states that the uniformity in stress placement in nouns and its variability in verbs follows from a structural difference between these categories and the resulting difference in the way they map onto prosodic structures. Specifically, prefixes are separate phonological words in his analysis, and a phrase–level stress rule puts the stress on the initial phonological word in the phonological phrase.

The exclusion of right-edge clitics from stress assignment was noted by [7, p. 48], [13, p. 46]. Clitics don't affect in any way the location of stress on their host, but (other) affixes will, as illustrated in (1bc) [10, p.24], [13, p.46]. Many examples of minimal pairs can be given, like *gol* 'flower', which gives ['goli]$_{CG}$ 'one flower' with a clitic and [go'li] 'proper name' with an affix.

## 3. Intonation in Persian

The intonational structure of Persian has been interpreted as involving three levels of prosodic hierarchy, the accentual phrase (AP), the intermediate phrase and the intonational phrase [8], [11, p.36]. Stressed syllables are associated with pitch accents [3]. [2] posits four pitch accents, H*, L*, L*+H and L+H*, two phrase accents marking intermediate phrases, L- and H- as well as two boundary tones, L% and H%. According to Eslami, pitch accents carry meaning, as in (3).

| | | |
|---|---|---|
| (3) | H* | new information |
| | L* | given information |
| | L+H* | contrast |
| | L*+H | doubt |

As for the intermediate phrase tones, L- is taken to be used when the message is complete and H- when it is incomplete. Finally, L% is used for statements and H% for questions.

Sadat Tehrani [11] described the AP-contour as the pitch accent L+H*. There are two allophones for this pitch accent: L+H* and H*, whereby L+H* is used for polysyllabic and H* monosyllabic APs**.** Information structure is claimed to be expressed by deaccenting APs after the focus constituent. The AP is ended by a boundary tone, which [11] notates as ‘h’ in the case of non-final (‘non-nuclear’) APs and as ‘l’ in the case of the last accented (‘nuclear’) APs. In broad focus sentences, the nuclear pitch accent is on the last AP, followed by ‘l’, and pre-nuclear accents appear on non-final APs, followed by ‘h’. This predicts that if an AP boundary tone lands on an unaccented AP-final syllable, which must be a clitic, the clitic has equal pitch or higher pitch relative to the preceding accented syllable, which has H* [11, p. 47].

# 4. The experiment

We composed a corpus of sentences featuring two minimal pairs contrasting a noun and a noun plus clitic combination. This pair of minimal pairs itself contrasted only in the voicing of the obstruent in the onset of the last syllable, which in the CG was the last consonant of the stem. These data form part of a larger corpus, in which more segmental conditions are included. Since no obvious quadruplets were available in the segmental condition we report on here, one of the four words was a nonsense word: *tabeš* ‘light’ vs *tab-eš* ‘swing+his/her’ and *tapeš* ‘nonsense word’ vs *tap-eš* ‘tank-top+his/her’. The members of these minimal pairs were embedded in carrier sentences which varied across three focus conditions, referred to as neutral (5a), post-focal (5b) and focal (5c). For the neutral and post-focal carrier sentence we used *Un X-e* ‘That is X’, where *-e* is a clitic. Condition (5c) differs from (5a,b) in having *un* ‘that’ in final position, which allows *X* to be focused and *X-e* to be in first position in the sentence, the focus position. The sentences were presented in standard Persian orthography, which uses Arabic letters. Conditions (5a) and (5b) were distinguished by having bold print for the experimental word in (5a) and bold print for *un* in (5b), reproduced here in the romanized spelling. These twelve sentences were given twice, once with a question mark (؟) and once with a full stop (.) at the end, in order to elicit both declarative and interrogative intonation contours. Each sentence was read twice in a professional recording studio by five educated native speakers recruited from the Linguistics Department of University of Tehran. Speakers were freely allowed to repeat themselves if they thought they hadn’t read a sentence correctly. The two best versions were selected from the utterances of each sentence by each speaker. This precedure yielded 240 utterances. In the majority of cases, these were the only readings produced for the sentence. After inspecting these utterances, we decided to discard some of them because of disfluencies or technical problems. In such cases, all versions of the sentence were discarded. This occurred 11 out of 240 times, or 7.6% of the cases. This ensured that our data were free from artefacts that might have had a distorting effect on the results.

(5) a. un tɒ'beʃ-e     un 'tɒb-eʃ-e
that light-is     that swing-his/her-is
‘That is light’     ‘That is his/her swing’

b. **un** tɒ'beʃ-e     **un** 'tɒb-eʃ-e
‘THAT is light’     ‘THAT is his/her swing’

c. **tɒ'beʃ-e** un     **'tɒb-eʃ-e** un
‘That is LIGHT’     ‘That is his/her SWING’

## 4.1 Results

Textgrids were produced in Praat [10] in which all segment boundaries were determined. Instead of establishing only the start of the closure duration and the end of the stop burst of plosives, the boundary between closure and burst was included as a segmental boundary, for both voiced and voiceless plosives. In the case of voiced plosives, this meant that we had burst intervals of zero duration in a number of cases. Initial plosives were only measured for their bursts, as no reliable indication of the beginning of the closure is available. An example of a TextGrid with wave form is shown in Fig. 1. As can be seen, we also included separate tiers for segments, words and clitics.

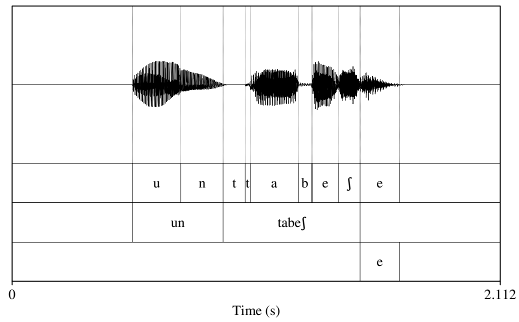


Figure 1. Praat textgrid for the noncliticized *Un tabeš-e*.

Subsequently, we averaged all values over the repetitions. Because of the way we supplied values for the 7.6% missing data, we have potentially reduced the variation in the data. For this reason we decided to adopt a 1% significance level for all results.

### 4.1.1 Duration

An analysis of variance (repeated measures) was performed on the durations with STRUCTURE (clitic vs nonclitic), MODE (declarative vs interrogative), FOCUS (neutral, post-focal, focal), VOICE (voiced vs unvoiced), and SEGMENT (t-burst, [a], plosive closure, plosive burst, [e], [ʃ] and [e]) as factors. None of the main factors or interactions was significant by Mauchly’s test for sphericity and we report uncorrected 1% significance levels. Our interest was in interactions between each of the four nonsegmental variables and SEGMENT. That is, we expected differential effects on the segment durations by the four nonsegmental factors. Three two-way interactions with SEGMENT were significant: MODE × SEGMENT ($F(1,5)=56.40, p<.001$), VOICE × SEGMENT ($F(1,5)=35.68, p<.001$) and FOCUS × SEGMENT ($F(2,5)=9.96, p<.001$). We observed a near-significant interaction STRUCTURE × SEGMENT ($F(1,5)=2.89$), $p=.02$). Other than an expected main effect for segment, there was a main effect for FOCUS ($F(2,5)=10.20, p<.01$) and MODE ($F(1,5)=44.66, p<.01$). The near-significant interaction for STRUCTURE is interpretable, since the main difference is in the duration of penultimate [e], which is accented in the nonclitic condition, where it is 17 ms longer than unaccented [e] in the clitic condition. The greater duration must be due to the presence of the pitch accent (accentual lengthening). This interpretation is reinforced by 19 ms longer plosive duration for [p/b] when the consonant opens an accented syllable, as it does in the clitic condition, taking burst and closure stages together. No durational difference is observed for [a], however. This suggests that accentual lengthening is confined to word-final syllables. In turn, this suggests that the difference is not due to inherent stress which might exist independently of the pitch accent, since lengthening due to stress typically occurs in all positions.

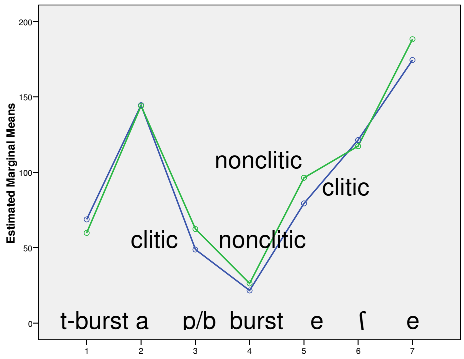


Figure 2. *Mean segment durations in clitic and nonclitic conditions.*

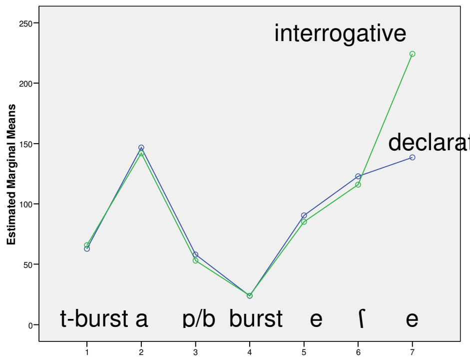


Figure 3. *Mean segment durations in questions and statements.*

The interaction with MODE is exclusive to the last vowel, [e], which is 224 ms in the interrogative and 138 ms in the declarative intonation. This relatively large difference suggests that question intonation is in part cued by increased final lengthening (see Fig. 3). The interaction with FOCUS is due the final two segments, which are shorter in the focus condition (not shown). This result rather shows that focus has no effect on the segment durations. The post-focal condition, in which the experimental items appear in utterance-final position after a focally accented *un* ('that') is virtually identical to the neutral condition, in which the focus is broad ('neutral'), but the word order in the same. The difference with the focal condition is to be attributed to phrase-final lengthening in the neutral and post-focal conditions, resulting in a final vowel which is 205 ms, compared to just 134 ms for the focal condition, where the vowel precedes *un*. An 11 ms concomitant difference for [ʃ] points to phrase-final lengthening of the last syllable (123 ms vs. 112 ms).

Lastly, the effect on segment durations by the voicing of the plosive is confined to the vowel [a] and the plosives themselves. As expected, the vowel is longer before [b] than before [p] (159 ms vs 129 ms; not shown), while conversely, [b] has a shorter duration than [p] (54 ms vs 105 ms for burst and closure together).

### 4.1.2 Centre of gravity

When the energy distribution in an area defined over some time window and frequency bandwidth is considered an object with mass, it will have a center of gravity (COG). The COG of a given segment is in a sense its mean frequency. For sonorants, the COG is related to spectral slope, whereby the steeper the slope, the lower is the COG. The measure is particularly useful for segments without well-defined formant structure, like those with voiceless friction[14]. COGs were calculated for [ʃ] and the burst of [p]. Analyses of variance with STRUCTURE (clitic vs nonclitic), MODE (declarative vs interrogative) and FOCUS (neutral, post-focal, focal) separately for [ʃ] and [p/b]-burst yielded no significant effects.

### 4.1.3 Fundamental frequency

We report mean f0 for the clitic and nonclitic forms for neutral, post-focal and focal conditions for declarative and interrogative intonation separately. Fig. 4, 5 and 6 show the declarative condition, while Fig. 7, 8 and 9 do the same for the interrogative condition.

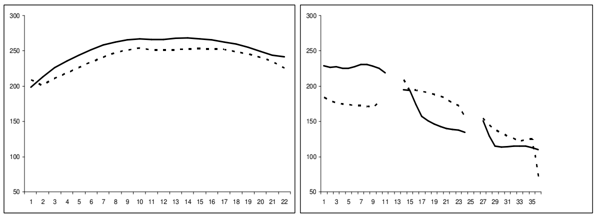

*Figure 4. Mean declarative F0 contours for* un *(left) and in neutral* tɒ'[b/p]eʃ-e/'tɒ[b/p]-eʃ-e *(right) on normalized time scales (clitic: —, nonclitic:----).*

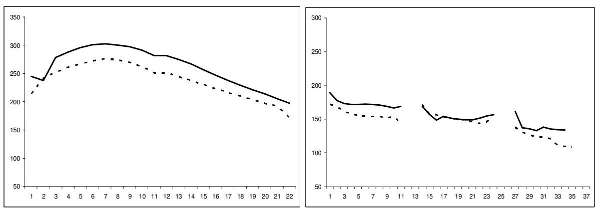

*Figure 5. Mean declarative F0 contours for* un *(left) and post-focal* tɒ'[b/p]eʃ-e/'tɒ[b/p]-eʃ-e *(right) on normalized time scales (clitic: —, nonclitic:----).*

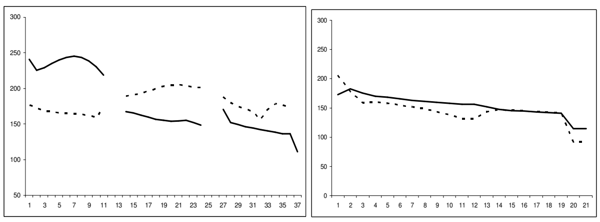

*Figure 6. Mean declarative F0 contours for focal* tɒ'[b/p]eʃ-e/'tɒ[b/p]-eʃ-e *(left) and* un *(right) on normalized time scales (clitic: —, nonclitic:----).*

Two things are suggested by the declarative contours. First, a comparison of the neutral (Fig. 4) with the post-focal (Fig 5) contrasts between cliticized and non-cliticized forms suggests that post-focal forms are not in fact deaccented. The pattern whereby the first syllable of the stem [tɒp]/[tɒb] has high pitch and the following clitic low pitch, evident in Fig. 4, is repeated in Fig. 5, but within a reduced pitch range. This suggests that the tonal structure is preserved after the focused *Un*, and that rather the phonetic realization is adjusted through pitch range reduction. Second, inspection of Fig. 6 confirms this in that the pitch of utterance-final, post-focal *un* is *higher* than the end of the preceding focal cliticized form. Since the latter ends low, and declaratives end low, the raised pitch between these two low targets must be due to a H-tone. This can only be the H* which is assigned to *un* on the grounds that it is an AP. Its realization, however, is affected by its post-focal status, as its pitch is some 60 Hz below that of the H*-bearing syllable [tɒ] in the cliticized form in Fig. 4, which is not post-focal.

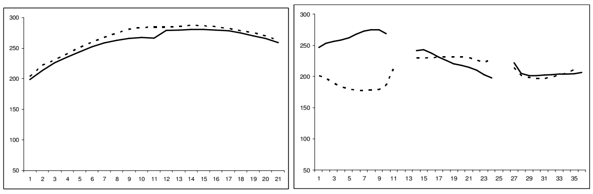

Figure 7. *Mean interrogative F0 contours for* 'Un' *(left) and in neutral* 'ta'[b/p]eš-e/'ta[b/p]-eš-e' *(right) on normalized time scales (clitic: —, nonclitic:----).*

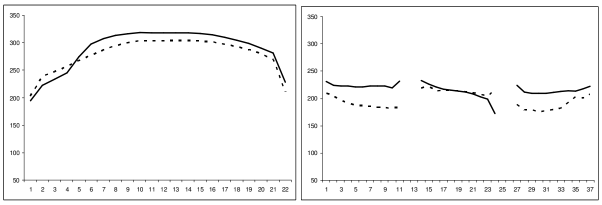

Figure 8. *Mean declarative F0 contours for* '**Un**' *(left) and post-focal* 'ta'[b/p]eš-e/'ta[b/p]-eš-e' *(right) on normalized time scales (clitic: —, nonclitic:----).*

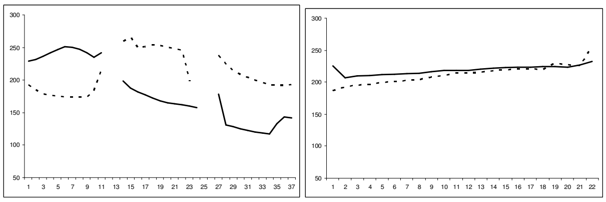

Fig.9. *Mean interrogative F0 contours for focal* **'ta'[b/p]eš-e/'ta[b/p]-eš-e'** *(left) and* 'Un' *(right) on normalized time scales (clitic: —, nonclitic:----).*

The interrogative contours confirm both conclusions, as shown in Fig. 7 and 8 for the comparison of the contrasts in neutral and post-focal positions, since the post-focal differences (Fig. 8) are again reduced versions of the contrast in neutral position (Fig. 7). Second, what this set of interrogative contours suggest beyond this conclusion is that there is no L%, as the pitch remains high, but does not contain a H% either, as the pitch does not rise, and if anything falls a little, though not as much as in the declaratives. A third conclusion is that there no evidence of a H- tone after the non-final AP. Rather than rising, the F0 falls, most clearly so in the situation in which unaccented syllables appear before the non-final AP-boundary, as in the clitic case shown in Fig. 9. Thus, non-final APs have L-, which is not clearly pronounced when H*-toned syllables define the AP-boundary.

### 5. Conclusions

Suffixed and unsuffixed Persian words have a pitch accent on their final syllable. CGs distinguish themselves from words by not shifting the pitch accent onto the right-attaching clitic or clitics, thus leaving the pitch accent on the syllable before the first clitic. Our investigation into the realization of segmentally identical structures that differ in the presence of a word-clitic boundary ('clitic' and 'non-clitic' conditions) revealed that post-focally tonal structures remain intact, but their realization is with reduced pitch range. The tonal structures of both the neutral and post-focal conditions thus is as in (6a) for the word condition and (6b) for the clitic condition. In the interrogative, the L% is not there, which we indicate by %, shown in (7a) fro the non-clitic and (7b) for the clitic condition. Representations of the focal cases is entirely parallel, but have *un* in final position.

(6) a. un tɒ'beʃ-e un 'tɒb-eʃ-e
H*L- L+H* L% H* L- H* L%

(7) a. un tɒ'beʃ-e un 'tɒb-eʃ-e
H*L- L+H* % H* L- H* %

## 6. Acknowledgements

This research is carried out with financial support from the Iran Telecommunication Research Center. We thank Joop Kerkhoff for technical assistance.

## 7. References


[1] Boersma, P. and Weenink, D., "Praat-doing phonetics by computer", 1992-2009.

[2] Eslami, M., "Šenaxt-e nævay-e goftar-e zæban-e farsi væ karbord-e an dar bazsazi væ bazšenas-ye rayan'i-ye goftar" [The prosody of the Persian language and its application in computer-aided speech recognition], Phd Thesis, Tehran university, 2000.

[3] Eslami, M. and Bjankhan, M., "Nezam-e ahæng-e zæban-e Farsi" [Persian intonation system], Iranian Journal of Linguistics, 34: 36-61, 2002.

[4] Ferguson, C., "Word stress in Persian", Language, 33:123-135, 1957.

[5] Gussenhoven, C., "The phonology of tone and intonation", Cambridge: Cambridge University Press, 2004.

[6] Kahnamuyipour, A., "Syntactic categories and Persian stress", Natural Language and Linguistic Theory, 21:333-379, 2003.

[7] Lazard, G., "Grammaire du Persan Contemporain", Paris: Librairie c, 1957.

[8] Mahjani, B.,"An instrumental study of prosodic features and intonation in modern Farsi(Persian)", MSc thesis, retrieved from www.ling.ed.ac.uk/teaching/postgrad/mscslp/archive/dissertations/2002-3/behzad_mahjani.pdf.

[9] Nespor, M. and Vogel, I., "Prosodic phonology", Foris: Cincinatti/Dordrecht, 1986.

[10] Noorbakhsh, M. and Bijankhan, M., "vajšenasi-ye nævayi-e zæban-e Farsi", Vijename zæbanšenasi, Pæjuhešgah Miras Færhangi, 14:18-33, 2007.

[11] Sadat Tehrani, N., "The intonational grammer of Persian", Phd Thesis, University of Manitoba, 2007.

[12] Same'i, H., "Tekye-ye fe'l dær zæban-e farsi : Yek bæresi-ye mojædæd" [Verb stress in Persian: A reexamination], Nameye Færhængestan, 1:6-21, 1996.

[13] Shaqaqi, V., "Clitics in Persian", Phd Thesis, University of Tehran, 1993.

[14] Van Son, R.J.J.H. and Pols, L., "An acoustic profile of consonant reduction", Institute of Phonetic Sciences and IFOTT, University of Amsterdam.