A PITCH ACCENT POSITION CONTRAST IN PERSIAN

Vahideh Abolhasani Zadeh[a,b], *Carlos Gussenhoven*[c,d] *& Mahmood Bijankhan*[b]

[a]Department of English Language & Literature, Shahid Bahonar University of Kerman, Iran;
[b]Department of Linguistics, University of Tehran, Iran;
[c]Department of Linguistics, Radboud University Nijmegen, the Netherlands;
[d]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

Small acoustic differences in duration, intensity and vowel formants were found between initial and final accented target words in Persian, by the side of substantial differences in f0. On the basis of these data and the results of a perception experiment in which an f0 continuum was superimposed on a single source utterance, we conclude that Persian has a Japanese-style pitch accent, not an English-style stress contrast.

Keywords: pitch accent, stress, clitic group, phonological word, Farsi, Persian

1. INTRODUCTION

Word prominence contrasts can broadly be divided into those that rely on f0 differences and those that rely on durational and spectral differences which are accompanied by f0 differences. Beckman [1] used the term *pitch accent* for the first of these prominence-lending phonological elements for which Japanese was her example, and the *stress accents* for the second exemplified by English. In later work, the notion *stress accent* was redefined as a stressed syllable which is accented by means of a tone melody like the pitch accent of Japanese. That is, the critical difference between the Japanese and English cases is the presence of a phonetically stressed syllable in the location of the tone and tones making up the pitch accent. In this contribution, the terms *pitch accent* is used in this more general sense [2].

In languages that do not obviously have a system of stressed and unstressed syllables independently of the presence of pitch accents, small phonetic differences may nevertheless be found in duration, intensity and spectral properties between accented and unaccented syllables [1, 7]. Instead of invalidating the distinction between English-type and Japanese-type languages, such small differences ought possibly to be interpreted as side-effects of the pitch accent placement. Even English, a language with a quite salient difference between stressed and unstressed syllables, presents durational increases of pitch accented stressed syllables as compared to unaccented stressed syllables [4]. Rather than analyzing English as having unstressed, stressed and superstressed syllables, researchers have assumed that the additional duration in pitch accented stressed syllables is a side effect of pitch accent placement, as indeed suggested by the term 'accentual lengthening'.

This contribution aims to show that Persian has a pitch accent location contrast between cliticized and non-cliticized words and, second, that the phonetic differences in duration and spectral properties between accented and unaccented syllables are small and attributable to the allocation of a pitch accent to the accented syllable and the absence of a pitch accent in the unaccented syllable. A production experiment with 12 speakers yielded large f0 differences between the cliticized and noncliticized conditions and significant, but small differences in duration, spectral properties and intensity. A perception experiment supported the conclusion that f0 alone is responsible for the perception of pitch accent location, leaving no residue for any durational, spectral or intensity features.

2. PROMINENCE IN PERSIAN

The contrast at issue arises in Persian through the existence of a rule which provides the final syllable of phonological words with a pitch accent, crucially skipping right-edge clitics [5]. 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 (1) shpws, simplex uninflected words, suffixed words and compounds have final accented syllables. Since compounds lose accent on their first constituent (1b), they are analyzed as single phonological words, along with inflected words.

However, since right-edge clitics are not assigned a pitch accent, they are not included in the phonological word, and instead form a Clitic Group (CG) with the phonological word on their left.

(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’

Prefixed verbs are accented on the prefix. According to Mahjani [8], the uniformity in stress placement in nouns and its variability in verbs follows from the way verbs map onto prosodic structures. Specifically, prefixes are separate phonological words; a phrase-level rule puts the stress, as this Persian prominence is called in the literature, on the initial phonological word in the phonological phrase in his analysis. The exclusion of right-edge clitics from stress assignment was noted by [6] p. 46, [10], p. 48. Clitics don’t affect the location of the prominence on their host in any way, regardless of the number of right-edge clitics [10], p. 46. There are a fair number of minimal pairs whose members either form single words or words-plus-clitc combinations, like *gol* ‘flower’, which gives [['gol]$_{PW}$i]$_{CG}$ ‘one flower’, which has a clitic, and [[go'li]$_{PW}$]$_{CG}$ ‘proper name’, which takes a hypocoristic suffix [i].

3. PRODUCTION

In order to investigate the difference between cliticized and uncliticized expressions, we composed a corpus of sentences featuring two minimal pairs contrasting a noun and a noun-plus-clitic combination. The two minimal pairs themselves contrasted only in the voicing of the obstruent corresponding to the final position in the pre-clitic noun. We did this in order to assess the effect of voiceless segments in a sensitive location on the perception of the prosodic contrast. These materials in fact form part of a larger corpus in which more segmental conditions are included. Since no obvious quadruplets were available in the segmental condition we wanted to include, one of the four words was a nonsense word: *tabeš* ‘light’ vs *tab-eš* ‘toy+his’ and *tapeš* ‘nonsense word’ vs *tap-eš* ‘tank-top+his’. The members of these minimal pairs were embedded in carrier sentences which varied across three focus conditions, illustrated for the voiced-obstruent pairs in (2). We refer to these as the *nuclear* (2a), *post-nuclear* (2b) and *focused* (2c) conditions. The nuclear and post-nuclear carrier sentences correspond to English ‘That X-is’ and ‘THAT X-is’, where X is the target item and ‘is’ corresponds to a clitic. The contrasting locations in these sentences are therefore the penultimate (non-cliticized word) and antepenultimate (cliticized word) syllables in the utterance, never the final syllable. This was done to avoid the interference of boundary tones in the experimental syllable. Condition (2c) differs from (2a,b) in having the target item in sentence-initial position: ‘X-is that’. Since this is the position for focused constituents in Persian the confounding of focus and position is inevitable in these materials.

The sentences were represented in standard Persian orthography, which uses Arabic letters. Conditions (2a) and (2b) were distinguished by having bold print for the experimental word in (2a) and bold print for *un* in (2b), 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. This was done in order to elicit both declarative and interrogative intonation contours. Since these are known to have different f0 characteristics [9], we needed to make sure that our results were generalizable across these two common intonation contours. This procedure yielded 24 sentences.

(2) a.	Un ta'beš-e That light-is That is light	Un 'tab-eš-e That swing-is That is his swing
b.	**Un** ta'beš-e That light-is That is light	**Un** 'tab-eš-e That swing-his-is That is his swing
c.	**Ta'beš-e** un Light-is it That is light	**'Tab-eš-e** un swing-his-is it That is his swing

3.1. Data collection

The materials were presented to 12 educated native speakers in random order, blocked by intonation contour. Speakers were recruited from the University of Tehran and were aged between 26 and 37. They read each sentence twice and were freely allowed to repeat themselves if they thought they hadn’t read a sentence correctly. The two best versions of each sentence by each speaker were selected; often these were the only two that were produced. In a few cases, we decided to discard utterances because of disfluencies, in which case all versions of the sentence were discarded. Of the 2(clitic)×2(voice)×3(focus)×2(mode)×12(speakers)

=586 utterances 572 were thus analysed with the help of Praat [3]. We determined all segment boundaries in the target words, including that between stop closure and stop burst, for both voiced and voiceless plosives. In the case of voiced plosives, this meant that we had intervals of zero duration in a number of cases. Initial [t]'s were only measured for their bursts, as no reliable indication of the beginning of the closure was available. We measured f0, segment durations, mean segment intensity levels, F1, F2 and F3 of the two vowels, and the Centre of Gravity of the stop bursts and [ʃ]. Subsequently, we averaged all values over the repetitions and performed analyses of variance (repeated measures) with STRUCTURE (clitic vs nonclitic), MODE (declarative vs interrogative), FOCUS (neutral, post-focal, focal), VOICE (voiced vs unvoiced), and SEGMENT. The levels for SEGMENT varied with the dependent variable. For duration, these included t-burst, [a], labial closure, stop burst, [e], [ʃ] and [e]; for formants and intensity these were the two vowels, and for Centre of Gravity the stop bursts and [ʃ]. We report results for STRUCTURE and FOCUS.

3.2. Results for duration

Duration of the burst of [t]. We found effects for STRUCTURE (F[1,11]=20.46, p<.001), due to a 9 ms longer duration of the [t]-burst in the accented syllable, FOCUS (F[2,10]=5.64, p<.01), due a 10 ms longer duration than in the postfocal condition.

Duration of [ɒ]. An interaction STRUCTURE×FOCUS (F[2,10]=4.13), p<.01) is due to a 5 ms overall longer duration of the vowel in the accented syllable than in the unaccented syllable, to which difference the focus condition contributed most.

Duration of labial closure. An interaction STRUCTURE×FOCUS (F[2,10]=8.27, p<.01) as well as main effects for STRUCTURE (F[1,11]=13.23, P<.01) and FOCUS (F[2,10]=4.47, p<.05) were found. The labial closure is 7 ms longer when closing the accented syllable and 4 ms longer in the focal condition than in the post-focal condition.

Duration of [e]. Main effects of STRUCTURE (F[1,11]=5.18, p<.01) and FOCUS (F[1,34]=6.69, p<.001) are due to a 5 ms longer duration in the accented syllable (the cliticized word condition), while it is 9 ms longer in the neutral condition than in the focus condition.

Duration of [ʃ]. There was a main effect for STRUCTURE (F[1,11]=7.73 p<.05) due to a 3 ms longer consonant in the unaccented syllable, and for FOCUS (F[2,10]=68.09, p<0.001), due to a 10 ms shorter duration in the focal condition than in the neutral and postfocal conditions.

3.3. Results for energy

We found a 2.06 dB increase in the accented [ɒ] relative to the unaccented vowel (F[1,11]=5.04, p<.05) and a 1.64 dB increase in the accented [e] as compared to the unaccented vowel (F[1,11]=5.89, p<.05). The [ɒ] is 3.03 dB more intense in the neutral than in the postfocal condition and 4.3 dB less intense than in the focal condition (F[2,10]=42.46, p<.001), while the energy of vowel [e] is 3.98 dB more intense in the neutral condition than in the postfocal condition and 1.36 dB less intense than in the focal condition (F[2,10]=42.47, p<.001).

3.4. Results for spectral properties

Formants of [ɒ]. In the accented syllable, [ɒ] has a marginally higher F1 (18 Hz) than in the accented condition (F[1,11]=7.07, p<.01) and a marginally higher F2 (34 Hz) in the neutral condition than in the postfocal condition (F[2,10]=3.65, p<.05).

Formants of [e]. F1 is marginally higher (19 Hz) in the accented condition than in the unaccented condition (F[1,11]=10,91, p<.01). In the neutral condition, F2 is 48 Hz and 24 Hz higher than in the focal and postfocal conditions, respectively (F[2,10]=3.80, p<.01). There were no significant effect for Centre of Gravity.

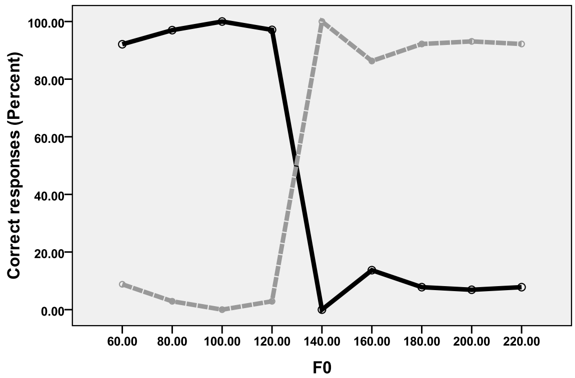
4. PERCEPTION

We found large differences in f0. The accented syllables were higher than the unaccented syllables is all cases. We here report the results of a perception test which more directly addressed the question of the perceptual effects of the acoustic differences we found. To establish more directly whether f0 is the sole cue differentiating pitch accent locations, we ran a perception experiment with manipulated versions of the same source utterance. We chose a sentence with neutral, declarative focus and a cliticized target word, *Un 'tab-eš-e* 'That is his swing', by a male speaker. Using Praat, we provided the first syllable of the target word, [tɒ], with a contour that was the average of the cliticized and non-cliticized words in the sentence pair spoken by that speaker. The shape of the contour was preserved by taking five equally spaced time points. Nine versions of this speech file were then produced in which the f0 contours in the second syllable of the target word, [beʃ], were 20 Hz apart, again taking five equally spaced points through the contour of [e], with the lowest step starting at 60 Hz and the highest at 220 Hz. These stimuli were presented in random order to a group of 30 female and 21 male listeners aged between 23 to 38 who were asked to indicate if they perceived the word for 'light'([tɒ'beʃ], non-cliticized) or the word for 'swing' (['tɒb], cliticized). These words were shown on a computer screen, on which subjects could click on the word they thought they heard. The results showed a complete perception shift from one condition to the next, with responses of 100% and 0% for at least one stimulus on each side of the switch point (see Fig. 1). A logistic regression analysis showed that this point lies at 155 Hz, which is approximately the mid point in our stimuli continuum. The effect of f0 was highly significant ($F(1,7)=4.54$, $p<0.01$). In this test, $R^2=.394$ and $\beta=.938$.

5. CONCLUSION

A highly detailed acoustic analysis of the productions by 12 speakers of 572 pairs of sentences with target words that contrasted in the location of a pitch accent revealed that in addition to large f0 differences, there were small differences in other acoustic variables. In particular, we found that accented vowels were somewhat longer, had marginally greater intensity and were marginally opener in the accented syllables than in the unaccented syllables. Less consistently, similar duration differences were found for the consonants. We interpret these differences as side effects of the occurrence of the pitch accent on the accented syllables, not as effects of an additional underlying difference in stress. Perceptually, f0 alone can be used to effect a 100% switch from one pitch accent location to the other, regardless of the original location of the pitch accent in the source utterance. This conclusion is supported by the fact that similar small acoustic differences in duration, intensity and vowel formants were found between different focus conditions. Further research will need to address the role of f0 in signaling focus.

Figure 1: The result of perception test.



6. ACKNOWLEDGEMENTS

This research is carried out with financial support from the Iran Telecommunication Research Center.

7. REFERENCES

[1] Beckman, M.E. 1986. *Stress and Non-stress Accent.* Dordrecht: Foris Publications

[2] Beckman, M.E., Pierrehumbert, J.B. 1986. Intonational structure in Japanese and English. *Phonology Yearbook* 3, 15-70.

[3] Boersma, P., Weenink, D. 2007. Praat - doing phonetics by computer. *http://www.praat.org/*

[4] Cambier-Langeveld, T.,Turk, A. 1999. A cross-linguistic study of accentual lengthening: Dutch vs. English. *Journal of Phonetics* 27, 255-280.

[5] Ferguson, C. 1957. Word stress in Persian. *Language* 33, 123-135.

[6] Lazard, G. 1957. *Grammaire du Persan Contemporain*. Paris: Librairie C. Klincksieck.

[7] Levi, S.V. 2005. Acoustic correlates of lexical accent in Turkish. *Journal of the International Phonetic Association* 35, 73-97.

[8] Mahjani, B. 2002. *An Instrumental Study of Prosodic Features and Intonation in Modern Farsi.* MSc thesis. *http://www.ling.ed.ac.uk/teaching/postgrad/mscslp/archive/dissertations/2002-3/behzad_mahjani.pdf*

[9] Sadat Tehrani, N. 2007. *The Intonational Grammar of Persian.* PhD diss., University of Manitoba.

[10] Shaqaqi, V. 1993. *Clitics in Persian.* PhD diss., University of Tehran.