ACOUSTIC CORRELATES OF LEXICAL STRESS IN PERSIAN

Vahid Sadeghi

Imam Khomeini International University, Iran
Vsadeghi5603@gmail.com

ABSTRACT

This paper examines the effects of lexical stress on intensity and duration in Persian both in the presence of the intonational prominence contrast and in the abstraction from the compounding accent condition. A production study was conducted in which 10 speakers produced Persian lexical and reiterant disyllabic minimal stress pairs spoken with and without an accent in a fixed carrier sentence.

Duration, overall intensity, and spectral levels in four contiguous frequency bands were measured. Results showed a significant difference in overall intensity and spectral tilt between accented and unaccented syllables. However, no significant differences were found for these measures between stressed and unstressed syllables in the absence of the accent contrast. In contrast, duration proved the most reliable correlate of stress, as Persian speakers produced stressed syllables consistently longer than their unstressed counterparts independent of the accompanying intonational prominence contrast.

Keywords: accent, spectral levels, overall intensity

1. INTRODUCTION

The primary purpose of this paper is the acoustical characterization of linguistic stress in Persian. Stress is a linguistic property of a word that specifies which syllable in the word is stronger than any of the others. However, understanding how prominence at the word level is expressed requires disentangling the effects of stress from those of accent which signals prominence at the utterance level. At this latter level, a speaker presents a word communicatively important by realizing a pitch accent on the prosodic head of that word; that is, the stressed syllable. For this reason, pitch movement has always been advanced as the most important phonetic correlate of linguistic stress ([2, 5, 7]). The earliest extensive instrumental study on this topic is that of Fry [5, 6] for English. He suggested that the traditional account of stress as a local increase in loudness could not be tenable. The experiments showed that overall intensity can differentiate stressed from unstressed syllables, but this measure is not a reliable indicator of the structural difference in the absence of rigid control for other things that affect intensity including F0. Huss [7] examined the effect of duration on the perception of stress in English stress minimal pairs ("im**port**" verb versus "**im**port" noun) in post-focal deaccented contexts, and found that in the absence of F0 and vowel reduction, listeners could not perceive stress.

Sluijter and van Heuven [11], however, pointed out that all previous studies had covaried accent and stress. They suggested that the traditional claim of loudness as a perceptual cue for stress seems justified if a more accurate measure of loudness is chosen than intensity. They examined the balance of energy in higher frequency bands relative to lower frequencies in a Dutch corpus with varying focus, and found an effect of primary stress even outside of focal accent position. They also found a significant effect for duration in non-focal deaccented position.

Campbell and Beckman [3] extended Sluijter and van Heuven's work to examine spectral correlates of stress and accent in a corpus of English sentences with varying focus, but found no difference in spectral tilt between vowels in stressed versus unstressed syllables in the absence of the intonational prominence contrast. They suggested that the difference in findings is in keeping with the difference between the two languages in the perceptibility of stress in the absence of accent or vowel reduction contrast. However, Ortega-Liberia and Prieto [9] found new evidence on the issue from Spanish that supported Sluijter and van Heuven's findings. They examined the patterns of production of stress in Catalan and Castilian Spanish which differed maximally in vowel reduction. Their results showed that in both languages, stress was cued by duration and intensity in the absence of pitch accents.

Earlier studies of Phonetic correlates of stress in Persian have shown the salience of F0 contour in cueing stress in minimal stress word pairs [1]. No specific experimental study, however, has been directed towards the acoustical realization of stress in the absence of the confounding effect of accent. In this paper, duration and intensity (overall intensity and spectral tilt) measures are examined as the acoustic correlates of stress in a corpus of Persian materials that varies lexical stress independently of accentual prominence. The questions asked in the present production experiment are: Can these measures reliably differentiate accented from unaccented syllables? Are they reliable acoustic correlates of linguistic stress when possible confounding with high F0 due to accent is undone.

2. PERSIAN STRESS PATTERN

The majority of lexical words in Persian are stressed on the final syllables [4]. Word-final stress pattern applies to nouns, adjectives, most adverbs, and simple verbs. However, prefixes in inflected verbs attract stress, resulting in recessive stress pattern. In addition, right-edge clitics, unlike suffixes, do not attract stress, leaving the stress pattern of the stem unaffected [1]. The intonational structure of Persian involves two levels of prosodic hierarchy, the accentual phrase (AP) and the intonational phrase (IP) [10]. Aps are associated with 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* for monosyllabic APs**.**

3. METHODOLOGY

3.1. Material

The Persian minimal stress pair [sɑ**zeʃ**] 'compromise' versus [**sɑ**zeʃ] 'his instrument' differing only in stress position were selected as the target tokens. The reiterant version of this word pair was also used, where each syllable was replaced by the syllable /nɑ/: /**nɑ**nɑ/ versus /nɑ**nɑ**/. Reiterant speech allows us to study prosodic phenomena while abstracting from segmental influence [8]. The vowel /ɑ/ was chosen because it is the most open, longest vowel in Persian, resulting in the largest distance between F0 and F1. The target words were embedded in pre-final position in a carrier sentence: *ʔanha be donbale [target] hastand* "they in search of [target] are". Targets were spoken with and without a pitch movement on the stressed syllable.

3.2. Subjects and procedure

The resulting four stimulus types (2 stress positions * 2 accent conditions) with their reiterant versions were read 4 times each by 5 male and 5 female speakers in two sessions (two times in each reading session). The Shure microphone (SM58 Cardiod Dynamic) was used for recordings. Subjects produced lexical and reiterant versions of each stimulus in immediate succession. The condition with the target outside focus ([-accent]) was realized by placing a single contrastive accent on the first word of the sentence: *ʔanha*. In the other condition ([+accent]), a single nuclear pitch accent was placed on the stressed syllable of the target. The target words were accompanied by synonyms within parenthesis to yield the appropriate reading. The word bearing contrastive accent, *ʔanha* was realized in bold face on the monitor, and the speakers were told that the word was to be interpreted as expressing a narrow contrast.

3.3. Data analysis

The 320 utterances (2 stress positions * 2 accent conditions * 2 versions [lexical versus reiterant] * 10 speakers * 4 repetitions) were digitized (10 KHz sampling rate, 5 KHz low-pass filtering) and analyzed by Praat (version 5.2.12). Syllable durations were measured in a straightforward fashion using the visual criteria for determining segmentation boundaries described by van Zanten, et al. [12]. The overall intensity of the stressed and unstressed vowels was defined as the average amplitude of the Fourier spectrum between 0 and 5 KHz over the stable portion of the vowel. Intensity was also measured in four contiguous frequency bands B1-B4, 1-0.5, 0.5-1.0, 1.0-2.0, and 2.0-4.0 KHz. The spectral level of a frequency band was defined as average amplitude of the Fourier spectrum over the stable portion of the vowel in that frequency band. The lowest band was chosen so that it included the fundamental frequency. The second, third, and fourth bands were chosen so that they included F1, F2, and F3 respectively.

4. RESULTS

Three-way analyses of variance were run for both lexical and reiterant speech data with stress, accent condition and syllable position as fixed effects and

with repetition and speaker as repeated measures. Three-way analyses of variance were also run for both lexical and reiterant speech data to examine the effects of stress, accent condition on overall intensity and spectral levels.

4.1. Duration

Table 1 shows mean values for the syllable duration measure broken down by speech type, accent condition, and stress position. As can be seen, Stressed syllables are longer than unstressed syllables [lexical: $F(1,312)= 181.30$, $p< 0.001$; reiterant: $F(1,312)= 75.426$, $p< 0.001$]. Accented words, also have longer syllables than unaccented words [lexical: $F(1,312)= 109.22$, $p< 0.00$; reiterant: $F(1,312)= 191.94$, $p< 0.001$]. There were no interactions between stress and accent [lexical $F(1,312)= 1.21$, $p= 0.272$; reiterant: $F(1,312)= 0.043$, $p= 0.836$]. Furthermore, the differences between stressed and unstressed syllables in the final-stressed words are smaller than the differences in the initial-stressed words. This contradicts earlier findings of Sluijter and van Heuven [11], who examined data with similar syllable structure, and found longer durations for final syllables due to pre-boundary lengthening. However, there was a significant interaction between stress and syllable position for lexical speech [lexical: $F(1,312)= 5.05$, $p= 0.025$; reiterant: $F(1,312)= 1.06$, $p= 0.302$], indicating that the combined effects of stress and syllable position are not completely additive. There were no significant interactions between accent and syllable position [both data: $F <1$].

Table 1: Mean syllable duration (ms) and standard deviations of the first (σ1) and second (σ2) syllables of sɑzeʃ (lexical) and nɑnɑ (reiterant). The data are presented for [+Acc] and [-Acc] conditions.

		Lexical		Reiterant	
Accent	Stress	σ1	σ2	σ1	σ2
[+Acc]	[+str]	284 (41)	235 (44)	263 (44)	224 (41)
	[-str]	251 (39)	212 (37)	238 (40)	208 (33)
[-Acc]	[+str]	253 (41)	216 (39)	229 (43)	192 (28)
	[-str]	225 (44)	188 (28)	207 (31)	171 (26)

4.2. Overall intensity

Table 2 shows means and standard deviations of the overall intensity measure. In the [+acc] condition, there is a difference of about 5 dB between the stressed and unstressed vowels. In the [-acc] condition, the difference is only about 1 dB. Stress proved significant for both lexical and reiterant speech [lexical: *syll. 1* $F (1,156) = 9.907$, $p= 0.002$, *syll. 2* $F (1,156)= 6.772$, $p= 0.010$, reiterant: *syll. 1* $F (1,156) = 6.459$, $p= 0.012$, *syll. 2* $F (1,156)= 5.103$, $p= 0.025$].

Table 2: Mean overall intensity (dB) and standard deviations of the first (σ1) and second (σ2) syllables of sɑzeʃ (lexical) and nɑnɑ (reiterant). The data are presented for [+Acc] and [-Acc] conditions.

		Lexical		Reiterant	
Accent	Stress	σ1	σ2	σ1	σ2
[+Acc]	[+str]	71.80 (6.8)	71.77 (7.2)	71.84 (6.9)	70.39 (7.3)
	[-str]	66.42 (6.2)	65.80 (6.4)	66.35 (6.5)	67.21 (6.7)
[-Acc]	[+str]	66.72 (7.3)	65.14 (7.1)	66.17 (7.1)	67.31 (6.2)
	[-str]	65.85 (6.8)	64.52 (6.6)	65.49 (6.3)	65.79 (6.8)

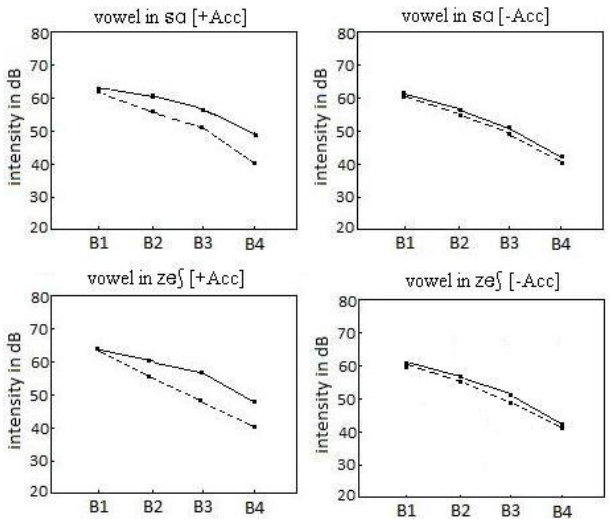
Accent also produced a significant effect for both types of speech [lexical *syll. 1* $F (1,156)= 8.104$, $p= 0.005$, *syll. 2* $F (1,156)= 9.719$, $p= 0.002$, reiterant: *syll. 1* $F (1,156)= 7.259$, $p= 0.008$, *syll. 2* $F (1,156)= 4.678$, $p= 0.032$]. Importantly, the interaction between stress and accent condition was significant in all cases except for the second syllable of the reiterant token. [lexical *syll. 1* $F (1,156)= 5.161$, $p= 0.024$, *syll. 2* $F (1,156)= 4.458$, $p= 0.036$, reiterant: *syll. 1* $F (1,156)= 3.947$, $p= 0.049$, *syll. 2* $F (1,156)= 0.641$, $p= 0.42$]. Stress only differentiated stressed from unstressed syllables in the [+acc] condition. This strongly suggests that overall intensity in Persian like English ([3]) and Dutch ([11]), is more likely an acoustic correlate of accent than stress.

4.3. Spectral levels

Figure 1 shows mean intensity of stressed (solid line) and unstressed (dashed line) vowels in *sɑ* and *zeʃ* in [+acc] condition (left) and [-acc] condition (right). As can be seen, in the [+acc] condition, the negative spectral tilt of unstressed vowels is considerably steeper than that of stressed vowels. The intensity in the lower filter band is hardly affected by stress, while there are considerable differences in higher bands. In [-acc] condition, however, stressed vowels have only a slightly gentler spectral slope than unstressed vowels. The intensity differences in all filter bands are hardly noticeable. Stress did not cause a significant effect on the intensity of the lowest band of both speech data [all cases $F<1$] , but did exert a significant effect on the intensity on all other bands [B2 lexical: *syll. 1* $F(1,156)= 8.925$ $p=$

0.003, *syll. 2* F(1,156)= 7.362, *p*= 0.007, reiterant: *syll. 1* F(1,156)= 7.609, *p*= 0.007, *syll. 2* F(1,156)= 10/01, *p*= 0.002; B3 lexical: *syll. 1* F(1,156)= 15.526 *p*= <0001, *syll. 2* F(1,156)= 36.314, *p*< 0.001; reiterant: *syll. 1* F(1,156)= 12.11, *p*< 0.001, *syll. 2* F(1,156)= 18.832, *p*< 0.001; B4 all cases, *p*< 0.001]. In addition, accent produced a significant for all spectral measures except B1 in both lexical and reiterant speech data and B2 in the second syllable of the lexical speech [B1lexical: *syll. 1* F(1,156)= 1.45, *p*= 0.23, *syll. 2* F(1,156)= 2.88, *p*= 0.092; reiterant: *syll. 1* F(1,156)= 1.83, *p*= 0.17, *syll. 2* F(1,156)= <1; B2 lexical: *syll. 1* F(1,156)= 7.912, *p*= 0.006, *syll. 2* F(1,156)= 0.467; reiterant: *syll. 1* F(1,156)= 7.231, *p*= 0.008, *syll. 2* F(1,156)= 8.13, *p*= 0.005; B3 lexical: *syll. 1* F(1,156)= 7.264, *p*= 0.008, *syll. 2* F(1,156)= 4.953 *p*= 0.02; reiterant: both syllables *p*< 0.001, B4 all cases *p*< 0.001]. Crucially, the interaction of stress and accent was (almost) significant for all the filter levels B2 lexical: *syll. 1* F(1,156)= 3.962, *p*= 0.04, *syll. 2* F(1,156)= 4.863 *p*= 0.02; reiterant: *syll. 1* F(1,156)= 4.061, *p*= 0.04, *syll. 2* F(1,156)= 4.597, *p*= 0.03; B3 lexical: *syll. 1* F(1,156)= 6.722, *p*= 0.01, *syll. 2* F(1,156)= 4.748 *p*= 0.03; reiterant: *syll. 1* F(1,156)= 8.65, *p*= 0.004, *syll. 2* F(1,156)= 9.865 *p*= 0.002;, B4 all cases *p*< 0.001], indicating that the effects of stress on the intensity in the higher spectral levels are considerably stronger in [+acc] tokens than in [-acc] tokens. The results contrast earlier findings of Sluijter and van Heuven [11], who found a significant effect for stress both in the focal and non-focal accent condition.

Figure 1: Mean intensity (dB) of unstressed (dashed) and stressed (solid) vowels in for [+Acc] (left) and [-Acc] (right) conditions.



5. CONCLUSIONS

This study examined the acoustical correlates of stress and accent in Persian. The measurements were made in both [+acc] and [-acc] conditions. Results showed that duration is the most reliable acoustic correlate of stress, irrespective of accent. Overall intensity and spectral differences, however, are poor indicators of stress position as they could only distinguish stressed from unstressed tokens in the [+acc] condition. As such, overall intensity and spectral intensity distribution can be regarded as the reliable acoustic correlates of accent rather than stress, and that it is conceivable to suggest that the greater intensity and energy distribution in higher bands, associated with stressed vowels, are mainly caused by the larger amplitude of the voice source. These results contradict traditional accounts of stress as a dynamic phenomenon which regard loudness as the primary phonetic correlate of stress resulting from greater physiological effort.

6. REFERENCES

[1] Ablhasani Zadeh, V., Gussenhoven, C., Bijankhan, M. 2010. The Position of clitics in Persian intonational structure. *Speech Prosody* Chicago.

[2] Beckman, M., Edwards, J. 1994. Articulatory evidence for differentiating stress categories. In Keating, P.A. (ed.), *Papers in Laboratory Phonology* Cambridge: Cambridge University Press, 3, 7-33.

[3] Campbell, N., Beckman, M. 1997. Stress, prominence and spectral tilt. In Botinis, A, Kouroupetroglou, G, Carayiannis, G. (eds.), *Intonation: Theory, Models and Applications.* ESCA, 67-70.

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

[5] Fry, D.B. 1955. Duration and intensity as physical correlates of linguistic stress. *J. Acoust. Soc. Am.* 27, 765-768.

[6] Fry, D.B. 1958. Experiments in the perception of stress. *Language and Speech* 1, 126-152.

[7] Huss, V. 1977. English word stress in post-nuclear position. *Phonetica* 35, 86-105.

[8] Liberman, M.Y., Steeter, L.A. 1978. Use nonsense-syllable mimicry in the study of prosodic phenomena. *J. Acoust. Soc. Am.* 63, 231-233.

[9] Ortega-Liberia, M., Prieto, P. 2007. The phonetics of stress in Catalan and Spanish, *Proc. 16th ICPhS* Saarbrücken, Germany.

[10] Sadat Tehrani, N., 2007. *The Intonational Grammer of Persian*. Ph.D. Thesis, University of Manitoba.

[11] Sluijter, A., van Heuven, V. 1996. Spectral balance as an acoustic correlate of linguistic stress. *J. Acoust. Soc.Am.* 100(4), 2471-2485.

[12] Zanten, E., van, Damen, L., van Houten, E. 1991. The ASSP Speech Database. *SPIN/ASSP-report*. Speech Technology Foundation, Utrecht, 41.