Compensatory lengthening in Persian: the timing of non-modal phonation

Vahid Sadeghi

Department of English Language, Imam Khomeini International University, Iran

<https://doi.org/10.36505/ExLing-2008/02/0049/000108>

Abstract

Persian has CVGC (or CVCG) sequences (G, a glottal consonant; /h/ or /ʔ/) which become reduced in certain occurrences, with the perceptual effect of the loss of the glottal consonant. The purpose of this study is to provide an acoustic description of the sequences in reduced forms. A production study examined three acoustic measurements of phonation types: H1-H2, H1-F1, and f0. The measurements were made at 15 ms time intervals throughout the vowel to determine the time course of phonation effect. Results indicate that G is realized as breathiness in CVhC (or CVCh) and laryngealization in CVʔC (or CVCʔ), and that such non-modal voice qualities spread onto half of the portion of the preceding vowels.

Key words: Spectral tilt, compensatory lengthening, voice quality variation

Introduction

It has long been suggested that Persian exhibits compensatory lengthening (CL), which is triggered by the deletion of glottal consonants in coda position in informal speech (Shademan 2003). This implies that Persian could possibly contain surface minimal pairs, such as /bæʔd/ → [bæ:d] “later” and /bæd/ → [bad] “bad” that would contrast only with respect to vowel length. Results of recent studies, however, indicate that vowel offset f0 and spectral tilt measurements for CVGC tokens changed significantly relative to CVC, suggesting CL might better be interpreted as a quantitative variation of glottal gesture (Bijankhan 2000).

The present paper attempts to explore the time course of voice quality variation in CL data. Measurements of the time course of variation in phonation in the data indicate that non-modal phonation in CVGC data is not localized to vowel offset, as was already assumed; rather, glottal consonants extend their non-modal quality over half of the portion of the preceding vowel. Thus, it is suggested that CL involves a sequencing of modal and non-modal phonation. The portion characterized by modal quality corresponds to the entire length of the vowel preceding glottal consonants, and the second portion which is associated with non-modal voice quality and has a shorter duration relative to the modal part corresponds to the entire length of the coda glottals, realized with magnitude reduction of glottal gesture. Thus, the additional length associated with CL data could result

from the laryngeal variation of coda glottals that adds an interval of non-modal vocalic gesture to the preceding vowel.

Figure 1. The spectrograms and f0 contours for the words [shar?] (a), [sharh] (b) and [shar] (c), by the same speaker.

Figure 2. FFT spectra for the vowel [a] in [shar?] (a), [sharh] (b) and [shar] (c) sampled at 45 ms to the vowel offset.

Phonetics of phonation types

Phonation types can be quantified through a number of phonetic parameters. These include fundamental frequency and spectral tilt, among others. Non-modal-phonation types are commonly associated with lowering of fundamental frequency (Gerfen and Baker 2005), (Gordon and Ladefoged 2001). Spectral tilt is also known to differentiate phonation types in a number of languages including Jalapa Mazatec (Blankenship 2002), Tagalog (Blankenship 2002), and !Xoo (Gordon and Ladefoged 2001). It can be quantified by comparing the amplitude of the first harmonic to that of the second and higher frequency harmonics. It is assumed that spectral differences yield positive values for breathy vowels which have the greatest drop off in energy as frequency increases, and negative values for creaky vowels which display the largest boost in energy at higher frequencies.

Displayed in figure 1 are the spectrograms, f0 contours for the words [ʃarʔ] ‘religion’ (a), [ʃarh] ‘explanation’ (b) and [ʃar] ‘evil’ (c), produced by the same speaker. In [ʃarʔ], though there is little obvious evidence in the spectrogram signaling the presence of laryngealization, f0 contour shows clear f0 drop over the second portion of the vowel. In [ʃarh], on the other hand, the spectrogram shows a clear case of breathiness in the second portion of the vowel, as reflected in the decreased overall acoustic intensity and aperiodic energy at higher frequencies. The same portion is also characterized by the lowering of fundamental frequency in the f0 contour. The modal token [ʃar] shows no evidence of aperiodicity, intensity drop or vowel excursions throughout the vowel. It is also apparent from the spectrogram that the duration of the vowel [æ] in [ʃarʔ] and [ʃarh] is considerably greater than their modal counterpart in [ʃar]. Spectral slope differences are illustrated in Fig. 2 which shows FFT spectra for the vowel [a] in [ʃarʔ] (a), [ʃarh] (b) and [ʃar] (c) sampled at about 45 ms to the vowel offset. The FFT spectrum for [æ] in [ʃarʔ] displays a steeply positive spectral tilt due to a fall off in energy at H1, while the spectrum for [æ] in [ʃarh] shows a most deeply negative spectral tilt due to a considerable loss of energy at H2 and higher frequencies. The modal [æ] in [ʃar] occupies the intermediate ground with its higher frequencies having slightly less amplitude than the fundamental.

Experiment

Subjects and tokens

Two sets of monosyllabic words were selected. Each set contained ten words. The words in the first set were of the CVC type, and the words in the second set were of the form CVGC. Five words of the CVGC data contained the glottal fricative /h/, and the other five contained the glottal stop /ʔ/. The words were embedded in ten sentence frames with the main stress falling on the embedded words. 15 native speakers with no known knowledge of linguistics utter colloquially the randomized sentences.

The last 80 ms portion of vowels in each related CVC / CVGC tokens was designated for comparison. The portion of each vowel was tagged at 15 ms time intervals. A Fast Fourier transform (FFT) was calculated over a 25.6 ms window centred at each tag. Values of f0, spectral parameters H1-H2, H1-F2, and vowel duration were computed for vowels in each related CVC / CVGC tokens.

Table 1: means of overall vowel length as well as means of f0 and spectral tilt values at 50 and 20 time intervals for /V/ in CVC and CVGC tokens.

		Mean		Std. dv		t		p	
		50	20	50	20	50	20	50	20
D(H1-H2)	VCV	2.4	2.5	0.7	0.7				
	V?V	1.3	1.1	0.7	0.5	9.7	12.4	.00	.00
D(H1-F1)	VCV	-4.2	-4.1	1.0	0.9				
	V?V	-7.4	-7.1	1.1	1.1	16.5	15.6	.00	.00
F0	VCV	133	132	10.2	10.5				
	V?V	126	123	10.6	11.8	4.4	4	.00	.00
L	VCV	96		6.6					
	V?V	141		5.8		-49		.00	

		Mean		Std. dv		t		p	
		50	20	50	20	50	20	50	20
D(H1-H2)	VCV	2.4	2.1	0.7	0.8				
	VhV	6.2	6.3	0.8	0.8	-29	-28	.00	.00
D(H1-F1)	VCV	-3.7	-4.1	1.2	1.0				
	VhV	9.6	9.8	1.1	1.0	-75	-75	.00	.00
F0	VCV	129	129	11.5	9.8				
	VhV	121	121	10.3	8.9	5.7	7.1	.00	.00
L	VCV	92		8.3					
	VhV	144		5.8		-49		.00	

Recordings

Digitization was performed at 20 KHz. Separate recordings were made for each of the glottal consonants /ʔ/ and /h/. The data recorded for each included 150 tokens (15 speakers × 2 syllable templates × 5 words).

Results and discussion

For each of the parameter H1-H2, H1-F2 and f0, a paired-samples T-test was conducted on the mean values calculated at each time interval across the 80 ms vowel portion. Results indicate that the means of H1-H2, H1-F2, and f0 for the vowels in CVhC and CV?C are significantly different from the vowels in CVC throughout the vowel duration except for the first two time intervals. In addition, T-test shows significant duration differences for modal versus breathy or laryngealized vowel contrast.

A significant finding for CL data acoustic structure was substantial changes in spectral tilt and f0 values during the second half portion of the vowel, as opposed to CVC data. These values are phonetically significant acoustic cues that can be used to recognize words involving CVGC or CVCG sequences. Hence, the result of this study can be used to improve ASR.

References

- Bijankhan, M. 2000. Farsi vowel compensatory lengthening: an experimental approach. Proc. 5th ICSLP Beijing.
- Blankenship, B. 2002. The timing of non-modal phonation in vowels. *Journal of Phonetics* 30, 163-191.
- Gerfen, C., Baker, K. 2005. The production and perception of laryngealized vowels in Coatzacoaspan Mixtec. *Journal of Phonetics* 33, 311-334.
- Gordon, M., Ladefoged, P. 2001. Phonation types: a cross-linguistic overview. *Journal of Phonetics* 29, 383-406.
- Shademan, Sh. 2003. Glottal-deletion and compensatory lengthening in Farsi. *UCLA Working Papers in Phonetics*, No, 104, 61-81.