

ACOUSTIC ANALYSIS OF THE PERSIAN FRICATIVE-AFFRICATE CONTRAST

Zahra Mahmoodzade(1) and Mahmoud Bijankhan(2)

University of Tehran (1, 2)

zmahmodzdh@ut.ac.ir, mbjkhan@ut.ac.ir

ABSTRACT

Acoustic analysis of 18 words spoken in citation form by 10 male native speakers of Persian suggests that each of the following parameters can cue the post-alveolar fricative-affricate contrast in isolated word forms in Persian: silence duration, frication duration, rise time and amplitude rise slope. The acoustic values of each cue differ with the position of the test item in the word. Silence and frication duration of Persian affricates was longest in final position. The rise time of Persian voiceless affricates was also longest finally but the amplitude rise slope was longest initially. It is confirmed that there is a positive relation between frication duration and silence duration, frication duration and rise time and also silence duration and rise time. The trading relation between frication duration and amplitude rise slope was negative.

Keywords: Persian, Affricate, Fricative

1. INTRODUCTION

The purpose of this study is to find out the acoustic cues that distinguish between Persian post alveolar affricates and fricatives in canonical word form. The acoustic properties of sounds are expected to be enhanced in natural canonical form ([9]).

Several studies mention the combination of different acoustic cues in the fricative-affricate contrast. These acoustic variables include: Silence duration ([5, 6, 12]), frication duration ([13, 1, 11]), rise time ([6, 8, 12]), amplitude rise slope ([11, 15, 1]), rate of rise ([15]), F2 transition ([5, 6, 7]), release burst ([6, 8]), zero crossing rate ([15]), vowel duration ([3, 6]) and frication onset centroid frequency ([1]). In this paper, four acoustic characteristics for Persian affricates and fricatives have been studied: duration of silent interval, duration of fricative noise, rise time of the fricative noise and amplitude rise slope of the frication.

Silence duration followed by release burst is characteristic of stop consonants and affricates ([12, 13]) and constitutes the most observable acoustic differences in the spectrograms of

fricatives and affricates: affricates have a closure interval followed by release burst while fricatives have no silence or burst ([6, 13]). Dorman et al [5] concluded that silence can be a sufficient cue to the affricate-fricative distinction in medial and word final position. Howell and Rosen's [8] experiments showed that the perception of affricates increases as the duration of the silent gap increases. Several studies demonstrate that release burst and silence duration have perceptual salience in the affricate-fricative contrast ([6, 12]).

Duration of fricative noise is also a powerful cue for the fricative-affricate distinction. Experiments have demonstrated that the perception of fricatives increases with an increase in the frication noise while its decrease will change the perceived manner of articulation from fricative to stop and affricate ([12, 13, 3, 6]). There is a trading relation between silence duration and noise duration: as fricative noise increases, more silence is needed for the perception of an affricate ([12, 5]).

Rise time, which is defined as the time from the onset of fricative noise energy to the maximum amplitude ([6, 7]), has been also considered as an important cue for the fricative-affricate distinction. Several experiments demonstrate that rise time is perceptually relevant; sounds with short rise times are generally perceived as affricates while sounds with long rise times are heard as fricatives ([2, 8]). There is a relation between rise time and silence; in long noise rise times, more silence is needed to perceive an affricate but in rapid rise times, affricates can be heard by less silence duration ([5, 12, 7]). Some investigations demonstrate that rise time is a less reliable perceptual cue for the affricate-fricative contrast ([1, 12]). Diehl and Castleman [4] have pointed out that rise time is not a robust cue and it might be confounded by frication duration.

Amplitude rise slope, a variable related to rise time, is the rate of increase in the amplitude of the acoustic wave energy through time and expressed as dB/ms ([15]). Affricates have an abrupt onset but fricatives have a gradual rise in energy ([15,

2)). Mitani et al [11] demonstrated that the interaction between frication duration and onset amplitude rise slope influenced the perceptual contrast and discrimination of voiceless fricatives and affricates and they concluded rise slope is a function of frication duration.

2. MATERIALS AND METHODS

Ten male subjects aged between 20-40, all students at the university of Tehran, were recorded producing words with /d₃, tʃ/ and /ʒ, ʃ/ in vocalic contexts /'-aC²/, /Ca'-aC/ and /'Ca-/ in initial, medial and final word position. Four repetitions of the word list were produced. There were 480 tokens (4 phonemes X 3 positions X 10 subjects X 4 repetitions). 29 non-affricate realizations were excluded, leaving 451 tokens for analysis. The speech was sampled at the rate of 22050 HZ.

	d ₃		tʃ	
	mean	SD	mean	SD
M	45.40	11.66	47.08	11.66
F	80.62	23.81	79.32	24.63
Total	65.30	26.15	66.78	25.89
Combined total	Mean = 65.78		SD = 25.89	

Table 1: Silence duration means in ms.

Subjects were not aware of the aim of the experiment. In order to keep control on the parameters used, words were pronounced in isolation, meaning that we cannot make unqualified generalization to connected speech from our results.

The acoustic values of each cue were measured by considering both spectrograph and waveform. We did not measure closure duration in initial position following Castleman [1] since all test items are preceded by silence. In medial or final position, the beginning of the closure was marked at the offset of vowel F2 and after the last pitch period on signal. The end of closure was determined at the point of release burst on the waveform. The beginning of frication noise in initial position was marked after the release burst for affricates and for fricatives at the point where sustained frication noise began on the waveform. Medially and finally for fricatives it was marked at the offset of F2. The end of frication in non-final positions was marked at the onset of F2. In the final position, the end of frication was determined at the point where sustained frication ended on the waveform. Rise time was measured following Castleman [1] by determining the time of frication onset and the time of maximum amplitude rise using an energy

contour program. Amplitude rise slope was measured by using Excel Slope Function. SPSS package 13 was used for statistical analysis.

3. RESULTS AND DISCUSSION

3.1. Silence duration

The silence durations of the affricates in medial and final position are given in table I. It was confirmed by an analysis of variance ($F(1, 129) = 0.05$, $p > 0.05$) that the mean silence duration of voiced affricates is not significantly shorter than the unvoiced affricates. Furthermore, closure duration of affricates varies with the position of the test item. Both voiced and voiceless affricate closure durations are significantly longer in final than in medial position. The mean silence duration of /d₃/ medially is 45.4ms and finally 80.6ms ($F(1, 67) = 55.24$, $p < 0.05$). The mean silence duration of /tʃ/ medially is 47.1ms and finally 79.3ms ($F(1, 60) = 37.06$, $p < 0.05$).

The results demonstrate that the silent interval of both Persian affricates was longer finally. They also show that there is no significant difference in silence duration between voiced and voiceless affricates in this position ($F(1, 53) = 0.28$, $p > 0.05$), contrary to Howell & Rosen [8]. It is also confirmed that it can not be used as a cue for the voiced-voiceless affricate contrast in this language.

3.2. Frication duration

The results for frication duration are presented in table II. The mean frication duration of affricates is clearly shorter than for fricatives. This was confirmed statistically by an ANOVA ($F(1, 449) = 122.72$, $p < 0.05$). The mean frication duration of /d₃/ significantly varied with word position ($F(2, 106) = 74.66$, $p < 0.05$). An LSD Post Hoc test showed that there is a significant difference between initial and final ($p < 0.05$) and medial and final position ($p < 0.05$) for /d₃/ but not between initial and medial position ($p > 0.05$). The mean frication duration of /tʃ, ʃ, ʒ/ also differs with the position of each phoneme in the word. This was confirmed by the following one way ANOVAs ($F(2, 99) = 90.79$, $p < 0.05$), ($F(2, 117) = 102.98$, $p < 0.05$) and ($F(2, 117) = 172.72$, $p < 0.05$). LSD Post Hoc tests showed that there is a significant difference between initial and medial ($p < 0.05$), initial and final ($p < 0.05$) positions for /tʃ, ʃ, ʒ/, and between medial and final positions ($p < 0.05$). The mean frication duration of the voiced affricate was significantly shorter than for the voiceless affricate

(LSD Post Hoc Test, $p < 0.05$) and the voiced fricative (LSD Post Hoc Test, $p < 0.05$). The mean frication duration of voiceless affricates is also significantly shorter than voiceless fricatives (LSD Post Hoc Test, $p < 0.05$). The results show that frication duration is a potential acoustic cue to the fricative-affricate distinction in Persian. Voiced affricate frication duration is shorter than for voiced fricatives and voiceless affricates are shorter than voiceless fricatives. There was a significant effect of position on the mean of frication duration for both fricatives and affricates in Persian. For all phonemes in our data, frication duration is longest finally but shortest medially except for /dʒ/ which was shortest initially. There is also a significant difference in the mean of frication duration between voiced and voiceless affricates. There is a positive relation between frication duration and silence duration in Persian (Pearson correlation = 0.3, $p < 0.01$). Repp et al [12] and Dorman et al [5] have stated that as frication duration increases, more silence is needed for the perception of an affricate. Notably, the mean durations for /tʃ/ and /ʒ/ are not significantly different. It can be argued that there is no need for this difference because of other phonetic differences that can cue the distinction, namely the fricative-affricate difference and the voicing difference.

	dʒ		tʃ		ʒ		ʒ	
	mean	SD	mean	SD	mean	SD	mean	SD
I	46.68	10.94	112.50	25.83	179.43	33.92	111.73	24.37
M	47.63	10.07	87.04	13.75	127.68	18.52	81.60	11.69
F	95.56	29.98	170.49	30.69	236.83	44.48	168.63	25.01
total	64.43	30.53	127.29	42.54	181.31	56.06	120.65	41.92
total dʒ & tʃ	Mean = 94.81				SD = 48.38			
total ʒ & ʒ	Mean = 94.81				SD = 58.00			

Table II: Frication duration means in ms.

3.3. Rise time

Table III shows that the mean rise time for voiceless affricates is shorter than for voiceless fricatives ($F(1,220) = 78.82$, $p < 0.05$). The effect of position on affricate and fricative rise time was significant ($F(2, 99) = 37.45$, $p < 0.05$), ($F(2, 117) = 49.64$, $p < 0.05$).

LSD Post Hoc test shows that rise time difference for affricates is significant between initial and final ($p < 0.05$) and medial and final ($p, 0.05$) position but it is not significant between initial and medial position ($p > 0.05$). LSD Post Hoc test for fricatives showed that there is a significant difference between initial and medial ($p < 0.05$) and final

position ($p < 0.05$) but there is not such a difference between medial and final position ($p > 0.05$). Our findings indicate that voiceless affricate rise time was shorter than that of voiceless fricatives. For Persian voiceless affricate, rise time is shortest medially but longest finally, while for voiceless fricatives it is longest initially and shortest medially. Howell & Rosen [8] concluded that in voiceless affricates, rise time is longer finally. Our study on Persian voiceless affricates also shows this relation. A further result of this study was that Persian voiceless fricative rise time was longest initially.

	tʃ		ʒ	
	mean	SD	mean	SD
I	30.25	11.76	119.75	36.16
M	27.20	9.02	54.63	25.97
F	53.51	17.82	61.70	33.22
total	37.94	18.08	78.96	43.24

Table III: Rise time duration means in ms.

The mean rise time of Persian /tʃ/ of 37.94ms was much longer than that reported by Gerstman [7] of 5ms. Cutting and Rosner [2] reported the duration of rise time for affricates to be shorter than 40ms and for fricatives longer than 40ms. The mean rise time of Persian voiceless affricates is close to 40ms. Howell and Rosen [8] demonstrated that affricates were perceived as more natural with rise times of 30 to 50ms. They measured the mean rise time of voiceless affricates in isolated words as 49ms and voiceless fricatives 123ms. The mean rise time of Persian voiceless fricatives (78.69ms) is much shorter than this. Another interesting point is that there is a positive relation between frication duration and rise time and also silence duration and rise time in Persian (Pearson correlation = 0.05 and 0.4, $p < 0.01$). It has been reported that in long rise times, more silence is needed to perceive an affricate while in rapid rise times, less silence is required ([5, 12, 7]). Contrary to our results, Miller-Ockhuizen [10] has concluded that there is a negative relation between closure duration and rise time in Serbian palatal affricates. The Pearson correlation coefficient shows the silence duration-rise time trading relation is much stronger than frication duration-rise time relation in Persian. Miller-Ockhuizen [10] is also asserted that there is a positive relation between frication duration and rise time in Serbian affricates.

3.4. Amplitude rise slope

Results for amplitude rise slope in voiceless fricatives and affricates are presented in table IV.

Affricates have a steep rise slope but fricatives have a roughly linear slope ($F(1,220) = 190.89$, $p < 0.05$). Amplitude rise slope is also affected by position of the test item in the word. An LSD Post Hoc test revealed that there is a significant difference between final and initial ($p < 0.05$) or medial ($p < 0.05$) position for /tʃ/ but there is not such a difference between initial and medial position ($p > 0.05$) ($F(2, 99) = 22.49$, $p < 0.05$). ANOVA for /f/ shows that there is a significant difference between initial and medial ($p < 0.05$) and final position ($p < 0.05$) but no significant difference between medial or final position ($p > 0.05$).

	tʃ		ʃ	
	mean	SD	mean	SD
I	0.38	0.10	0.17	0.09
M	0.34	0.09	0.10	0.05
F	0.23	0.10	0.09	0.09
total	0.31	0.12	0.12	0.08

Table IV: Amplitude rise slope means in dB/ms.

Our findings show that voiceless affricates have a steep rise initially while voiceless fricatives have a more linear slope finally. But both sounds are steeper initially and less steep finally. Howell and Rosen [8] have claimed that fricatives have linear rise slope patterns. Mitani et al [11] found that there is a negative relation between frication duration and amplitude rise slope; as frication duration increases, slope steepness decreases. Pearson correlation coefficient (-0.4) shows that the trading relation between amplitude rise slope and fricative noise is negative in Persian ($p < 0.01$).

4. SUMMARY

The acoustic analysis of Persian fricatives and affricates demonstrated that each of the following variables can cue fricative-affricate contrast in isolated word forms: silence duration, frication duration, rise time and amplitude rise slope. Rise time and amplitude rise slope have been used as acoustic cues for voiceless fricative-affricate distinction. There was a significant effect of position on the acoustic values of each phoneme. Silence and frication duration of Persian affricates was longest in the final position. The rise time of Persian voiceless affricates was also longest finally but the amplitude rise slope was longest initially. It is confirmed that there is a positive relation between frication duration and silence duration, frication duration and rise time and also silence duration and rise time. The trading relation between frication duration and amplitude rise slope was negative.

5. REFERENCES

- [1] Castleman, W.A. 1997. Integrated perceptual properties of the [+/-continuant] distinction in fricatives and affricates. M.A dissertation. University of Texas at Austin.
- [2] Cutting, J., Rosner, B. 1974. Categories and boundaries in speech and music. *Perception and Psychophysics*. Vol. 16 (3), 564-570.
- [3] Cole, R.A. 1975. Perception of voicing in English affricates and fricatives. *J. Acoust. Soc. Am*, Vol. 58 (6), 1280-1287.
- [4] Diehl, R.L., Castleman, W.A. 1996. Integrated perceptual properties: The affricate/fricative distinction. In A.P. Simpson and Patzold(ed), *Sound patterns of connected speech: Proc. Symposium held at Kiel University*, 191-200.
- [5] Dorman, F.M., Raphael, J.L., Liberman, M.A. 1979. Some experiments on the sound of silence in phonetic perception. *J. Acoust. Soc. Am*, Vol. 65 (6), 1518-1532.
- [6] Dorman, F.M., Raphael, J.L., Isenberg, D. 1980. Acoustic cues for a fricative-affricate contrast in word final position. *J. Phonetics*, Vol. 8, 397-405.
- [7] Gerstman, L.J. 1956. Noise duration as a cue for distinguishing among fricative, affricate, and stop consonants. *J. Acoust. Soc. Am*, Vol. 28 (1), 160.
- [8] Howell, P., Rosen, S. 1983. Production and perception of rise time in the voiceless affricate/fricative distinction. *J. Acoust. Soc. Am*, Vol. 73 (3), 976-984.
- [9] Koreman, J. 2006. Perceived speech rate: The effects of articulation rate and speaking style in spontaneous speech. *J. Acoust. Soc. Am*, Vol.119 (1), 582-596.
- [10] Miller-Ockhuizen, A., Zec, D. 2002. Durational differences in Serbian palatal affricates. *Proc. 1st Pan-American / Iberian Meeting on Acoustics, Mexico*.
- [11] Mitani, S., Kitama, T., Sato, Y. 2006. Voiceless affricate/fricative distinction by frication duration and amplitude rise slope. *J. Acoust. Soc. Am*, Vol.120 (3), 1600-1607.
- [12] Repp, B.H, Liberman, A.M., Eccardt, T., Pesetsky, D. 1978. Perceptual integration of acoustic cues for stop, fricative, and affricate manner. *J. Experimental Psychology: Human Perception and Performance*, Vol. 4, 621-637.
- [13] Rosen, S. 1992. Temporal information in speech: acoustic, auditory and linguistic aspects. *Philosophical transaction: Biological Science*, Vol. 336 (1278), 367-373.
- [14] Rosen, S., Darling, A.M., Faulkner, A., Huckvale, M. 1993. Cue interaction in an intervocalic voiceless affricate/fricative contrast. *J. Acoust. Soc. Am*, Vol. 93 (4), 2392.
- [15] Weigelt, L.F, Sadoff, S.J., Miller, J.D. 1990. Plosive/fricative distinction: The voiceless case. *J. Acoust. Soc. Am*, Vol. 87 (6), 2729-2737.