Acoustics of contrastive palatal affricates predict phonological patterning

Amanda Miller-Ockhuizen† and Draga Zec†
† Cornell University, Ithaca, NY, USA
E-mail: am332@cornell.edu, dz17@cornell.edu

ABSTRACT

Serbian contains two classes of contrastive palatal affricates exemplified in *ćar* [tŝar] *'gain'* vs. *čar* [tʃar] *'magic'*. In the phonological process of iotization, [t] patterns with [tŝ], and [k] patterns with [tʃ]. Articulatorily, *ć* [tŝ] is laminal, more front with compressed lips, while *č* [tʃ] is apical, more back with protruded lips. Acoustically, two resonance frequency peaks in the frication noise distinguish the affricates. [tʃ] / [k] display lower frequency energy than [t] / [tŝ]. Different frequency ranges in the spectra of frication noise vs. stop bursts derive from constriction degree. We adopt a two-level model of the phonetics-phonology interface [1]. Phonetic differences in several articulator variables in the input contribute to cavity volume differences in the output, and define the natural classes observed in iotization. We propose a binary phonological feature, increased vs. decreased [front cavity volume], which allows a natural statement of iotization.

1. INTRODUCTION

Serbian contains two classes of contrastive palatal affricates, exemplified in the minimal pair *ćar* [tŝar] *'gain'* vs. *čar* [tʃar] *'magic'*. In a process of iotization triggered by certain affixes, [t] alternates with [tŝ] and [k] with [tʃ], as shown by the comparative and possessive forms in Table 1.

Base adj.	**Gloss**	**Masc. comp.**	**Fem. comp.**	**Gloss**
žu[t]	'yellow'	žu[tŝ]i	žu[tŝ]a	'more yellow'
ja[k]	'strong'	ja[tʃ]i	ja[tʃ]a	'stronger'
Base noun	**Gloss**	**Masc. poss.**	**Fem. poss.**	**Gloss**
pse[t]o	'dog'	pse[tŝ]i	pse[tŝ]a	'dog's'
vu[k]	'wolf'	vu[tʃ]ji	vu[tʃ]ja	'wolf's'

Table 1: Comparatives and Possessives involving iotization

Articulatorily, *ć* [tŝ] is laminal, more front (palato-alveolar), and produced with compressed lips, while *č* [tʃ] is apical, more back (alveo-palatal), produced with protruded lips [2,3]. While place of articulation differences could offer a natural explanation of the patterns, the differences are subtle, and even non-existent for some speakers. Our acoustic investigation shows that the two affricates are distinguished consistently by two prominent spectral frequency peaks, with [tʃ] having lower energy spectral peaks than are found in spectra of [tŝ]. Different frequency ranges in spectra of frication noise vs. stop burst spectra are related to constriction degree. We propose a binary physiologically grounded phonological feature, increased vs. decreased front cavity volume [3], which allows a phonetically accurate statement of the iotization pattern, by capturing the similarity of [t] / [tŝ] and [k] / [tʃ].

2. METHODS

Five repetitions of a wordlist were recorded by 3 female and 3 male subjects at 22050 Hz sampling rate directly onto a PC at IEFPG. Recordings were then digitally transferred to a SUN Sparc station in the Cornell University Phonetics Lab. Initial consonants appear in the onset of an accented syllable, while medial consonants appear in the onset of an unaccented syllable. All forms are bisyllabic, and both syllables contain short vowels. Each production was segmented and labeled using the labels in Table 2.

Label	**Acoustic Landmark**	**Acoustic Cue used to label landmark**
CB	Closure Begin	End F1
FB	Frication Begin	Begin sustained frication noise
VB	Vowel Begin	Begin F1
PFA	Peak Frication Amplitude	Peak in frication noise

Table 2: Labels used in the study

We measured the centroid frequency of the spectra [4], the frequency and amplitude values of two prominent spectral peaks found in the 2500-5000 Z (P1), and the 6000-8500 Z (P2) ranges, and the second formant frequency (F2) in the following vowel. A single peak in the stop-burst spectra was identified within the 0-5000 Z range.

FFT spectra were created using the ESPS program *fft* with 1024 points and 10 coefficients using a Hanning window. These parameters resulted in a 46.5 ms window, which was aligned to the beginning of sustained frication noise (FB label) for the affricates and fricatives. Stop burst spectra were created using a shorter 23.5 ms Hanning window, which was aligned with the beginning of the stop burst. The original FFT spectra were converted to bark-scaled spectra using the ESPS program *barkspec.* Speaker-specific and consonant-specific ranges for each of the two spectral peaks were identified through visual

 ISBN 1-876346-48-5

inspection, and a peak-picking algorithm was used to objectively identify the peak frequencies. F2 values were also identified using the ESPS program *formant* at the onset and 30 ms into the vowel following the target consonant.

We measured frication duration (FB-VB), closure duration (CB-FB, or CB-VB), rise time (RT) [7] and maximum rate of rise (MAX ROR) [8]. RT was computed as the duration between the FB and PFA labels, and MAX ROR as the largest amplitude change over a 5 ms interval within a 50 ms window centered at the FB label in fricatives and affricates, and at the VB label in stops.

3. RESULTS

Figure 1 plots the peak 1 frequency (P1) against peak 2 frequency (P2) for all six subjects. P1 categorically distinguishes the two affricates for 4/6 subjects. P1 values for the fricative [ʃ] are similar to those for the affricate [tʃ]. One subject (sp4) actually displays a slight difference in the opposite direction for the two affricates in P1, but with complete overlap of the two categories. P2 categorically distinguishes the affricates for all subjects in this context (though Subject sp4's P2 range is small).

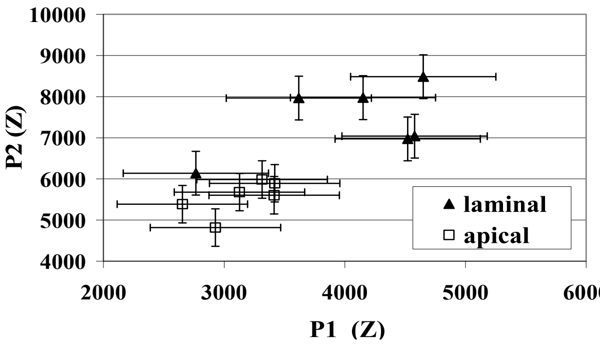

Figure 1: P1 against P2 in word-initial accented syllable (each point represents the mean value for a single subject)

Figure 2 displays P1 against P2 in medial position, in the onset of an unaccented syllable. The spectral peaks are not as distinct in this position. One subject (sp4) displays a similar lack of distinction in P1 to that found in initial position, but P2 distinguishes the affricates for all subjects.

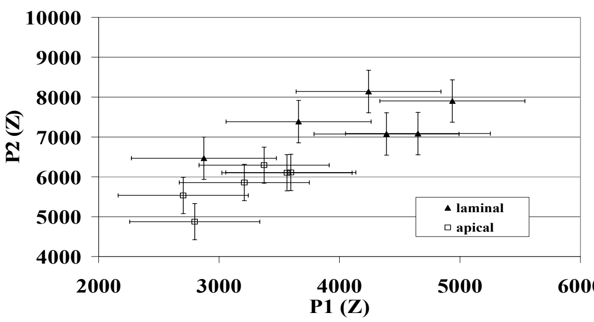

Figure 2: P1 against P2 in word-medial unaccented syllable
(each point represents the mean value for a single subject)

Mean and standard deviation values of the single spectral peaks in [t] and [k] stop bursts are provided in Table 3.

	sp1	sp2	sp3	sp4	sp5	sp6
[t]	3827 (1669)	3765 (1345)	2618 (1104)	3446 (1295)	4140 (1006)	3112 (1258)
[k]	2612 (1563)	2924 (1872)	2301 (1375)	2574 (1341)	2479 (1472)	2426 (1425)

Table 3: Mean burst peaks for [t] / [k] in the i#_(i,e,a,o,u) context (word-initial position)

Centroid values capture the overall frequency range differences for 4/6 subjects, but not for Subjects sp4 and sp6, as shown in Figure 3. The centroid values of the affricates differ by about 300-500 Z in both word positions.

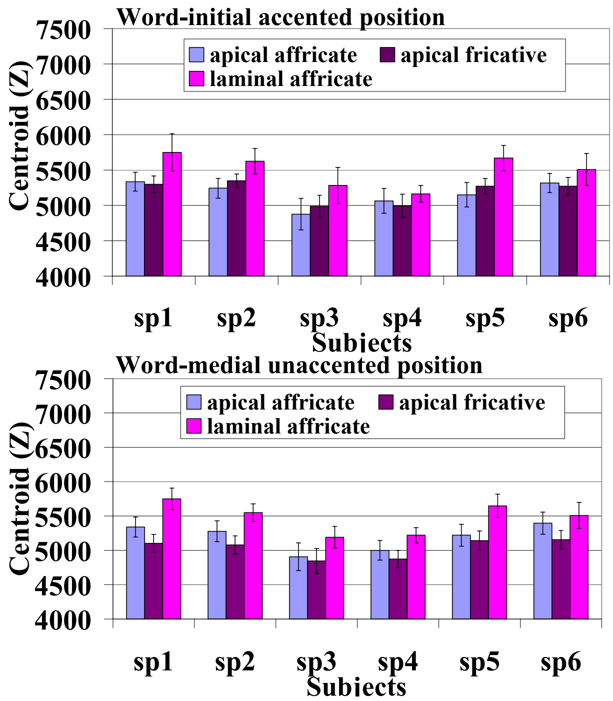

Figure 3: Centroid frequencies in word-initial and word-medial positions (i# _i,e,a,o,u context)

Figure 4 plots frication duration against closure duration in word-initial accented syllable position for Subject sp2. The duration of the two phases is inversely proportional for all six subjects [3].

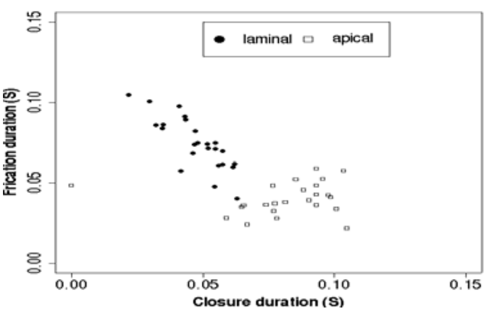

Figure 4: Closure duration vs. frication duration (word-initial accented syllable position, Subject sp2)

ISBN 1-876346-48-5 © 2003 UAB

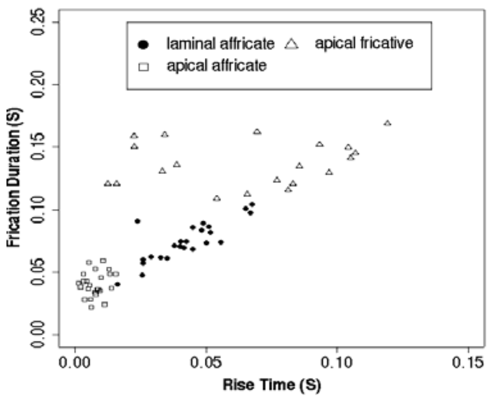

Figure 5: Rise time against frication duration (word-initial position, Subject sp2)

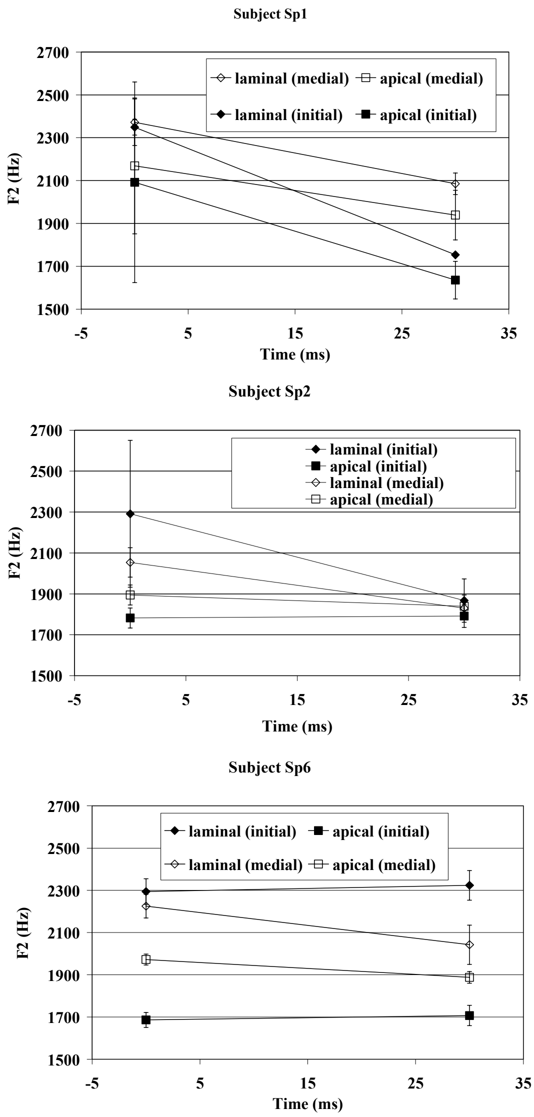

Figure 6: F2 in initial and medial position (0 and 30 ms)

Figure 5 plots RT [7] against frication duration for the same subject (sp2). The two measures are highly related for the affricates. The variability in the RT of [ʃ] is considerable because it is unconstrained by pressure build-up during a preceding stop closure. The lack of variability for some subjects' productions shown in [3] indicates that some subjects have associated RT with a cue to apicality. They have carried the cue over to fricatives, even though there is little physiological basis for it, suggesting that the goal is perceptual.

F2 values in the following vowel (i_e vocalic context) also distinguish the two affricates. As seen in Figure 6, F2 of the following vowel is higher after the laminal affricate than after the apical affricate, showing a coarticulatory fronting effect that is often found with palatals. As can be seen, there are variant neutralization patterns in the unaccented syllable. Subject sp1 displays higher F2 values for each of the affricates in word-medial unaccented position compared with word-initial accented position. Subjects sp2 and sp6, on the other hand, show neutralization of F2 values (which is most extreme for sp2).

The two affricates also display a strong relationship between ROR and closure duration, as seen in Figure 7 (initial position) and Figure 8 (medial position). However, stops that have longer closure durations than each of the affricates do not necessarily have a larger ROR in either position. The affricates have less distinctive ROR values in word-medial unaccented syllable onset position.

4. DISCUSSION

The frequency differences found in the salient spectral peaks of the two palatal affricates are attributed to a host of subtle articulatory differences. [tŝ] is laminal and more front, while [tʃ] is apical and more back with lip protrusion [2,3]. As is typical with coronals, there is a lot of inter-speaker variation [9]. In a study of 35 speakers' productions [2], about half of the subjects don't make a consistent or sizable distinction in place. While the apical vs. laminal distinction is consistent, that articulatory attribute is not applicable to non-coronals. Therefore, it does not predict the connection between [k] and [tʃ] found in the process of iotization. Crucial for phonological patterning are the acoustic attributes: [tʃ] and [k] are characterized by lower frequency peaks, and [tŝ] and [t], by higher frequency peaks in the spectra. [k] and [t] have overall lower frequency range stop bursts, given that the resonances are only seen in the formant transitions that occur during the opening phase. In the open phase, the resonances of the cavity behind the constriction are relevant as well as the resonance in front of the constriction. In the affricates, only the resonance in front of the constriction is acoustically realized since the peaks are realized in the relatively closed frication interval. While in [tŝ] and [tʃ] the higher frequency peaks reflect differences in cavity volume achieved through several articulatory gestures [10,12], the difference in front cavity volume and the resultant frequencies are attributable purely to place of articulation that subsumes cavity volume in [t] and [k]. It is not possible to analyze the frequency difference among the two affricates as one of stridency, since higher frequency peaks for both [tŝ] and [tʃ] exceed peaks found in typical strident

ISBN 1-876346-48-5 © 2003 UAB

fricatives or affricates [3]. Nor do the affricates differ along the compact vs. diffuse dimension [11], as P1 / P2 values are similar in frequency and amplitude for the affricates [3].

Temporal cues reflect the apical vs. laminal distinction. Cross-linguistically, apical stops tend to have shorter, and laminal stops tend to have longer closure duration [5]. The results for the two affricates are opposite, with [tʃ] having greater closure duration than [tŝ], but the expected relationship is found in the frication interval, which is longer for the laminal [tŝ]. This indicates that the physiological differences are manifested in the frication interval of affricates, which corresponds to the opening phase. Closure durations are controlled in order to realize MAX ROR differences and similar segment durations [6].

In a two-level model [1], multiple articulatory attributes of each affricate, represented as articulatory gestures [10, 12], map to a physiologically-based feature [front cavity volume], which accounts for both the phonetic differences and the phonological patterning of the two affricates [3].

ACKNOWLEDGEMENTS

We thank Mirjana Sovilj, Director of IEFPG, for providing facilities and technical support in Belgrade.

REFERENCES

[1] Zsiga, E. (1997) Features, gestures and Igbo vowels: An approach to the phonology-phonetics interface, *Language 73 (2),* 227-274.

[2] Miletić, B. (1933) *Izgovor srpskohrvatskih glasova*, Beograd: Izdanje zadužbine Milana Kujundžića.

[3] Miller-Ockhuizen, A. and Zec, D. (2003) The phonetics and phonology of Serbian palatal affricates*, WPCPL 15.*

[4] Forrest, K., Weismer, G., Milenkovic, P. and Dougall, R.N. (1989) Statistical analysis of word initial voiceless obstruents: preliminary data. *JASA 84,* 115-23.

[5] Maddieson, I. (1997) Phonetic universals In W.J. Hardcastle and J. Laver, Eds. *The Handbook of Phonetic Sciences*, Oxford: Blackwell, 640-673.

[6] Miller-Ockhuizen, A. and Zec, D. (2002). Durational differences in Serbian palatal affricates, *Proceedings of the First Pan-American / Iberian Meeting on Acoustics.*

[7] Howell, P. and Rosen, S. (1983) Production and perception of rise time in the voiceless affricate / fricative distinction, *JASA. 73(3)*, March 1983.

[8] Weigelt, L., Sadoff, S. and Miller, J. (1990) Plosive / fricative distinction: the voiceless case, *JASA 8(6)*, 2729-2737.

[9] Dart, S. (1991) *Articulatory and Acoustic properties of Apical and Laminal articulations, UCLAWPP 79.*

[10] Browman, C. and Goldstein, L. (1989) Articulatory gestures as phonological units. *Phonology 6,* 201-251.

[11] Jakobson, R., Fant, G. and Halle, M. (1968) *Preliminaries to Speech Analysis*: *The Distinctive Features and their correlates*, Cambridge, MA., The MIT Press.

[12] Fujimura, O. (1994) C/D Model: A computational model of phonetic implementation, In E. Ristad, Ed., *Language Computations: 1-20*. DIMACS Series in Discrete Mathematics and Theoretical Computer Science, Vol. 17, Providence, RI: American Mathematical Society.

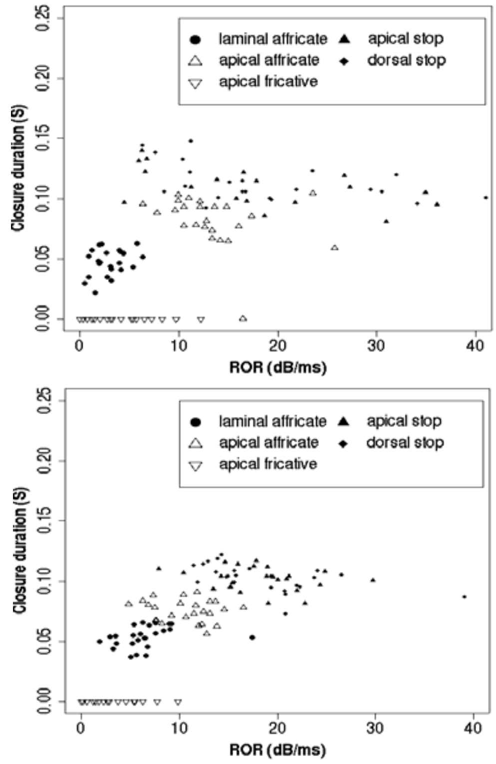

Figure 7: ROR against closure duration for Subjects sp2 and sp5 (word-initial accented syllable position)

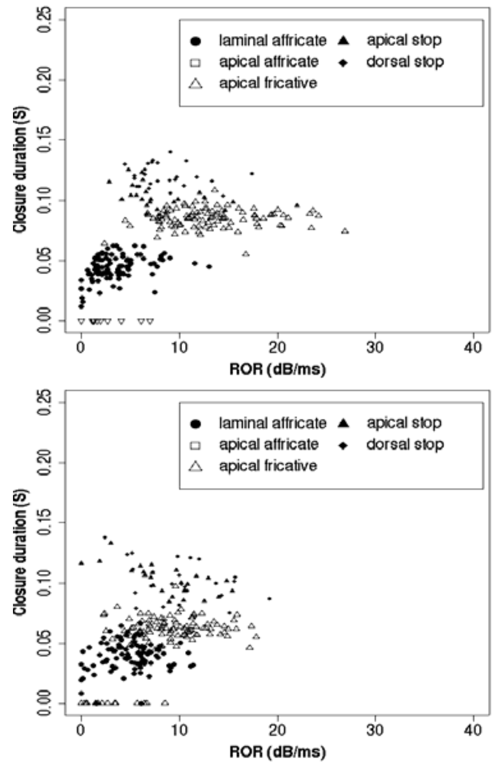

Figure 8: ROR against closure duration for Subjects sp2 and sp5 (word-medial unaccented syllable position)