Working Papers of the Cornell Phonetics Laboratory
2003, v.15, pp.130-193

Phonetics and Phonology of Contrastive Palatal Affricates*

Amanda Miller-Ockhuizen and Draga Zec

Serbian contains two classes of contrastive palatal affricates exemplified in *tŝar* 'gain' vs. *tʃar* 'magic'. In the phonological process of iotization, [t] patterns with [tŝ], and [k] with [tʃ]. Articulatorily, [tŝ] is laminal, more front with compressed lips, while [tʃ] is apical, more back with protruded lips. Acoustically, the affricates are distinguished by two prominent spectral peaks in the frication noise interval. [tʃ]/[k] display lower frequency spectral peaks than [t]/[tŝ]. Different frequency ranges in the spectra of frication noise vs. stop bursts derive from constriction degree. Temporal acoustic attributes show that they behave as a class of affricates, different from both stops and fricatives. Phonetic differences in several articulatory attributes in the input contribute to cavity volume differences in the output, which result in higher frequency spectral peaks for the laminal affricate than the apical affricate. These differences define the natural classes observed in iotization and allow a phonetically accurate statement of the [t]/[tŝ] and [k]/[tʃ] patterning.

1. Introduction

Standard Serbian (formerly known as the eastern variant of Serbo-Croatian) possesses two classes of contrastive palatal affricates, each containing a voiced/voiceless pair. The difference between these classes can be broadly captured in terms of apical vs. laminal articulation. The apical class includes [tʃ], which is voiceless, and [dʒ], which is voiced. The members of the laminal class are [tŝ] and [dẑ], which are voiceless and voiced respectively.[1] Articulatorily, the sounds in the laminal class are more front and produced with compressed lips, while the sounds in the apical class are more back and produced with protruded lips.

But overall, the phonetic differences between the two classes of palatals are small, as noted in Miletić's (1933) phonetic study of Serbian sounds. Their places of

---

* We thank audiences at the Acoustical Society of America/First Pan-American/Iberian Meeting on Acoustics in Cancun, Mexico (Miller-Ockhuizen and Zec 2002), at CNRS in Paris, particularly Jacqueline Vassiere, and at the International Congress of Phonetic Sciences in Barcelona, Spain (Miller-Ockhuizen and Zec 2003). We also thank Ian Maddieson for helpful discussions, and Abby Cohn for helpful comments on a draft version of this paper. We would like to thank Mariana Sovilj, Director of IEFPG, for generously providing facilities and technical support in Belgrade. We also thank Johanna Brugman, Emily Hanna, Cliff Crawford, and Marisol del Teso Craveotto for help with labeling the data.

[1] The apical class is represented by the standard IPA complex symbols [tʃ] and [dʒ]. But, because no IPA phonetic symbols are readily available for the laminal class, we introduce the phonetic symbols [tŝ] and [dẑ], modeled in part on the [ŝ] symbol used for a sibilant laminal fricative found in Ubykh and Abkhaz (Catford MS, Ladefoged and Maddieson 1996: 161). The orthographic symbols used for the two pairs of affricates are *č* and *dž* for the apical class, and *ć* and *đ* for the laminal class.

articulation are fairly close and, with some speakers, even overlapping. Moreover, the auditory difference among them, with members of the laminal class sounding “higher” and “softer” and members of the apical class sounding “lower” and “harder” (Miletić 1933), is subtle, yet sufficient for sustaining a phonological contrast. The contrast among these classes occurs both initially and medially. In word-initial position, the voiceless contrast is illustrated by the minimal pair *ćar* [tŝar]‘gain’ vs. *čar* [tʃar]‘magic’, and in word-medial position, the voiceless contrast is illustrated by the minimal pair ve*će* [vetŝe] ‘bigger’ vs. *veče* [vetʃe] ‘evening’.

The place of the two pairs of palatal affricates in the consonantal system of Serbian is presented in Table I. Note that, among palatal sounds, only affricates include both the apical and laminal classes. Palatal fricatives belong to a single, apical class, while all palatal sonorants are laminal (Miletić 1933, 1960).[2]

		Labial	Dental	Palatal (laminal)	Palatal (apical)	Velar
Obstruents	Stops	p b	t d			k g
	Affricates		ts	tŝ dẑ	tʃ dʒ	
	Fricatives	f	s z		ʃ ʒ	χ
Sonorants	Nasals	m	n	ɲ		
	Liquids		l r	ʎ		
	Approximants	v		j		

Table I: The inventory of Serbian consonants

The goal of this paper is to identify the phonetic basis for the phonological contrast between the laminal and apical classes of palatal affricates. We first argue, in Section 2, that due to a high degree of inter-speaker variation, no single articulatory attribute is sufficient for the phonetic characterization of these classes. In Sections 3 and 4, we present results of two acoustic experiments which investigate temporal and spectral

[2]Note that, in the class of dental obstruents, [s], [z] and [ts] have a laminal tongue constriction shape, while [t] and [d] are apical.

attributes of the affricates. In the experiments, we focus on the voiceless members of the two classes, [t̂ŝ] and [tʃ], and also include the stops [t] and [k], and the fricative [ʃ], as points of comparison. In Experiment 1, we test the hypothesis that the two classes of affricates differ in manner of articulation, while Experiment 2 is driven by the hypothesis that the two affricate classes differ in several spectral attributes. In Section 5, we show that the acoustic properties that we establish for the two classes of palatals are a good predictor of their phonological patterning. In the phonological process of iotization, the coronal stops [t]/[d] alternate with the laminal class, while the dorsal stops [k]/[g] alternate with the apical class. We account for this patterning by establishing acoustic similarity between [t] and [t̂ŝ] on the one hand, and [k] and [tʃ], on the other. We propose a new phonological feature that accounts for the acoustic similarity, as well as the [t̂ŝ]/[tʃ] contrast and the phonological patterning that these sounds participate in.

2. Articulatory properties of Serbian affricates

In this section, we show that the articulatory attributes of the two classes of palatals are not sufficiently distinct to serve as the sole phonetic basis for sustaining a phonological contrast. The articulatory evidence comes from the extensive study of the Serbian sounds in Miletić (1933), which uses palatography as the principal method. With 34 subjects participating in this study, palatograms were made for each subject's production of all Serbian sounds, including the two series of palatal affricates under investigation. In addition to the palatographic evidence for the two sounds, Miletić also includes lip photographs for a single subject (Subject 35), and presents a more qualitative description of the degree of lip protrusion for all subjects. He also provides linguagrams for all Serbian sounds, including the palatal affricates, for a single subject (Subject 20).

The articulatory data in Miletić's study allow us to critically evaluate how well the two palatal affricate classes are differentiated in articulatory terms. Miletić's descriptions will serve as basis for evaluations along three articulatory dimensions: (i) place of articulation, (ii) constriction location on the tongue (tip vs. blade, or laminal vs. apical), and (iii) presence or absence of labial compression (compressed vs. protruded lips). While the contrasting palatal affricates exhibit slight differences on all of these dimensions, the most salient being the laminal/apical distinction, we show that no single articulatory difference is either consistently present across speakers or robust enough to support a linguistic contrast. In Table II below, we present an overall summary of our findings. We will be referring back to it throughout this section.

Subjects	Place of Articulation		Tongue Constriction Location		Lip Protrusion	
	[tʃ]	**[tŝ]**	**[tʃ]**	**[tŝ]**	**[tʃ]**	**[tŝ]**
5,7	Denti-alveolar	Denti-alveolar	Apical	Laminal	Slightly protruded	Slightly compressed
14	Denti-alveolar	Alveolar	Apical	Laminal	Slightly protruded	Slightly compressed
33	Denti-alveolar	Denti-alveolar	Apical	Laminal	Slightly protruded	Slightly compressed
13	Denti-alveolar	Denti-alveolar	Apical	Laminal	Protruded	Compressed
31	Denti-alveolar	Denti-alveolar	Apical	Apical	Protruded	Compressed
21	Alveolar	Alveolar	Apical	Laminal	Slightly protruded	Slightly compressed
1,3,6,16, 19,23, 26,30	Alveolar	Denti-alveolar	Apical	Laminal	Protruded	Compressed
2,8,9,10, 22,27, 28,29	Alveolar	Alveolar	Apical	Laminal	Protruded	Compressed
18	Alveolar	Post-alveolar	Apical	Laminal	Protruded	Slightly Compressed
17	Post-alveolar	Denti-alveolar	Apical	Laminal	Slightly protruded	Slightly compressed
20	Post-alveolar	Alveolar	Apical	Laminal	Slightly protruded	Slightly compressed
25	Post-alveolar	Alveolar	Apical	Laminal	Slightly protruded	Slightly compressed
4	Post-alveolar	Alveolar	Apical	Laminal	Protruded	Compressed
11,12,15, 34	Post-alveolar	Alveolar	Apical	Laminal	Protruded	Compressed
24,32	Post-alveolar	Denti-alveolar	Apical	Laminal	Protruded	Compressed

Table II: Attested combinations of three articulatory attributes relevant to contrastive Serbian affricates based on our interpretations of articulatory data presented in Miletić (1933)

We begin with place of articulation. The type of place difference that will need to be characterized is illustrated in Figure 1, which presents palatograms for the laminal and apical palatal affricates produced by one of Miletić's subjects (Subject 20). In these palatograms, which are fairly typical, the coronal places of articulation overlap, and the difference between them is small.

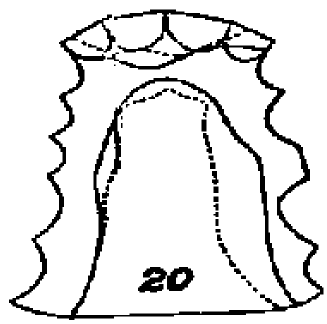

Figure 1: Palatograms of (a) the laminal palatal affricate [tŝ] in the word bȁća [bâtsa] 'bloke (nom.sg.)' (dotted line) and (b) the apical palatal affricate [tʃ] in the word màča [matʃa] ʻsword (1st gen. sg.)' (full line, Miletić 1933: Figure 83, Subject 20)

Our place of articulation descriptors are based on Dart's (1998: 76-77) system of place of articulation labels which "take into account only the most forward part of the contact area" (Dart 1998: 76).[3] Dart selects six, numerically labeled place of articulation points, in the coronal continuum. For our purposes, we single out four places of articulation: dental, denti-alveolar, alveolar and post-alveolar. The interpretations assigned to these place descriptors are consistent with articulatory target regions in Ladefoged and Maddieson (1996:15). Table III provides the region of contact for each place of articulation, as well as the correspondence of our place of articulation descriptors to Ladefoged and Maddieson's articulatory target regions, and Dart's numerical labels.

[3] Dart (1998) doesn't name her place categories, and uses instead a six point continuum.

Our categories	Region of contact	Dart (1998)	Ladefoged & Maddieson (1996)
Dental	Exhibit complete contact with the back of the upper central incisors	1	Dental
Denti-alveolar	Exhibit some contact with the teeth, from the base up to the midpoint of the teeth.	2 and 3	Dental and alveolar
Alveolar	Exhibit contact just behind the teeth, with no contact on the teeth	4	Alveolar
Post-alveolar	Exhibit a space between the edge of the teeth and the beginning of contact, with the size of the space being disregarded	5 and 6	Post-alveolar

Table III: Place of articulation category correspondences

Our findings, based on palatograms for all 34 speakers in Miletić's study, are summarized in the first column of Table II. As can be seen, there is a high degree of inter-speaker variation in place of articulation for each of the affricates, ranging from denti-alveolar to post-alveolar. It is striking, however, that both affricates are produced within this narrow articulatory region by all 34 subjects. Miletić (1933) himself noted that Subjects 4, 7, 13, 21 and 25 do not produce any difference in the place of articulation. Our own more fine-grained place of articulation classifications of Miletić's data detailed in Table II, point out subtle place of articulation differences between the two affricates for most subjects.

Consistent with the small differences in place of articulation seen in Figure 1, our place of articulation classifications do not indicate differences between the two affricates for Subjects 2, 5, 7, 8, 9, 10, 13, 21, 22, 27, 28, 29, 31 and 33. That is, 14 out of 34 subjects (41%) for which place of articulation data are available do not distinguish between the place of articulation of the two affricates based on our fairly fine-grained classifications. Further, some of the subjects that do produce the affricates with a place of articulation distinction produce a denti-alveolar/alveolar contrast, while others produce an alveolar/post-alveolar contrast, making a distinction in terms of place of articulation inconsistent. Only Subjects 17, 24 and 32 produce a comparatively larger distinction between denti-alveolar and post-alveolar places of articulation, which could be characterized using the place categories of [dental] vs. [alveolar]. Yet all of these places of constriction fall within Recasens' (1990) prepalatal zone. We question whether subtle place of articulation differences such as denti-alveolar vs. alveolar can be the sole basis of a contrast within a single language. Miletić notes that only two subjects (Subjects 7 and 25) completely neutralize the contrast. Thus, as Serbian speakers are able to distinguish minimal pairs containing the two affricates on a regular basis, there must be an additional aspect of the contrast not involving place of articulation that listeners rely upon.

Next, we turn to constriction location on the tongue, either tip or blade, responsible for the laminal/apical distinction. The linguagrams in Figure 2 (Miletić 1933: Figure 84) show the tongue constriction locations for the two affricates. Our understanding of *laminal* is in line with Ladefoged and Maddieson's (1996: 11) description of tongue blade as being centered just below the alveolar ridge when the tongue is at rest. With this understanding of the tongue blade, [tŝ] is laminal, and [tʃ] is apical.[4]

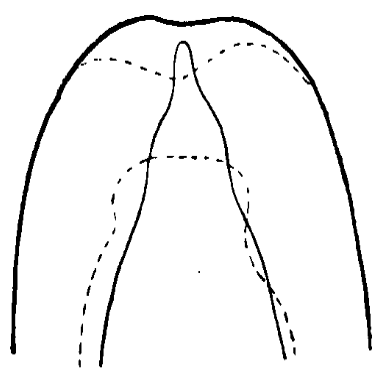

Figure 2: Linguagrams of (a) the laminal palatal affricate [tŝ] in the word bȁća [bâtsa] 'bloke (nom.sg.)'(dotted line) and (b) the apical palatal affricate [tʃ] in the word màča [matʃa] 'sword (1st gen. sg.)' (solid line) (Miletić 1933: Figure 84, Subject 20)

[4] Miletić (1933:34) provides a chart of Serbian consonants, in which both series of palatals are classified as alveolar; [tʃ] is further classified as produced with the tip of the tongue, and [tŝ], as produced with the tongue blade, or *praedorsum*, defined as "part of the tongue below the back part of the alveolar ridge" (Miletić 1933:20).

These are the only linguagrams provided in Miletić's study. It is on the basis of these linguagrams (provided for Subject 20), as well as the palatograms, and interpretations of Miletić's statements, that we analyze the difference between the two affricates as a difference between a more laminal and a more apical articulation. Consistent with this is Miletić's (1933) observation that all subjects have the tongue tip on the lower teeth in the articulation of [tŝ], and the tongue tip raised in the articulation of [tʃ], except for Subject 31, who produces both sounds with the tongue tip raised.

Tongue constriction location appears a consistent articulatory difference between the two affricates. The apical vs. laminal distinction is in fact the most consistent articulatory difference found for this group of subjects. Our own preliminary ultrasound investigation for one single Serbian speaker (one of the authors) confirms that tongue constriction location is laminal for [tŝ] and apical for [tʃ], as seen in Figure 3.

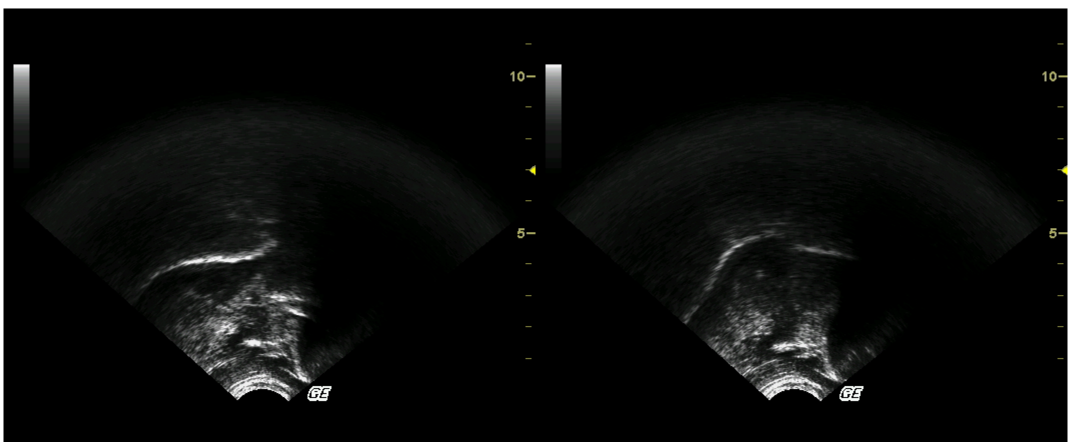

Figure 3: Mid-sagital images showing an apical tongue constriction location for [tʃ] in the word *čar* [tʃar]'magic'(left) and laminal tongue constriction location for [tŝ] in the word *ćar* [tŝar]'gain' (right)

Finally, lip protrusion differentiates the affricates quite well. According to Miletić's descriptions, all speakers produce [tʃ] with some degree of lip protrusion, although there is a considerable degree of inter-speaker variability in the degree of protrusion.

To summarize, we have identified articulatory differences between the two affricate classes along three dimensions, none of which can be singled out as the bearer of the phonological contrast. We hypothesize that the phonetic differences among the affricate classes point at the difference in the front cavity volume. Studies relevant for framing this hypothesis are Perkell et al (1979) and Keating (1988).

Perkell et al (1979) describe the difference between English [s] and [ʃ] in the following way: while [s] is characterized by contact between the tongue blade and the lower incisors, [ʃ] is characterized by the lack of such contact, which creates a sublingual cavity in the latter, but not in the former articulation. The overall effect is greater front cavity volume in the production of [ʃ] than in the production of [s]. Likewise, Miletić (1933) notes that the tongue blade touches the lower incisors in [tŝ] but not in [tʃ], suggesting the existence of a sublingual cavity in the production of [tʃ]. This further suggests that the front cavity for [tʃ] is greater for [tŝ].

Keating's (1988) study of palatal articulations contains several sounds that could serve as points of comparison for the two classes of affricates under study. The Serbian apical affricate is most similar to the Czech palato-alveolar affricate (Keating's Figure 8, taken from Hála 1962). Both sounds have relatively short constrictions and relatively large sublingual cavities. The Serbian laminal affricate bears resemblance to the Mandarin alveolo-palatal affricate (Keating's Figure 9, taken from Ohnesorg and Svarny 1955) in the relatively longer constriction, and to the Czech palatal nasal (Keating's Figures 1a and 1b, from Hála 1962), in the positioning of the tongue tip behind the lower incisors. While Keating claims that the Czech nasal has no sublingual cavity, we suspect that the situation with the Serbian laminal affricate is more variable. We already mentioned that one of Miletić's subjects produces both affricates as apicals, which for that subject suggests the existence of a sublingual cavity in the production of both affricates.

We hypothesize that Serbian speakers consistently differentiate the two affricates in terms of front cavity volume, using different articulatory strategies to achieve this, as is documented in the different articulatory strategies used by the subjects in Miletić's study. This hypothesis is tested explicitly in Experiment 2, by investigating the spectral properties of the frication noise.

In Experiment 1 we ask the larger question whether the two affricates differ in terms of the manner of articulation. An obvious hypothesis is that the two palatals are not both affricates, one possibility being that one of the sounds is a stop with an affricated burst, while the other is an affricate. We explicitly test this by investigating rise time, which has been shown to differentiate English fricatives from affricates, and maximum rate of rise, which has been shown to differentiate English affricates from stops.

We report on Experiment 1 in Section 3, and then turn to Experiment 2 in Section 4. In Section 5 we show that our experimental findings shed light on the phonological patterning of the apical affricates with dorsals stops, and of the laminal affricates with coronal stops, a pattern that cannot be explained in terms of any of the relevant articulatory attributes discussed in this section.

3. Experiment 1: Temporal properties of Serbian obstruents

3.1 Introduction

In this experiment, temporal properties of the two affricates are investigated. Closure durations of the two affricates are compared with closure durations in the stops [t] and [k], and frication durations of the affricates are compared with frication durations of the fricative [ʃ]. The hypothesis that the two affricates differ in terms of manner of articulation is explicitly tested. Specifically, rise time (Howell and Rosen 1983) of the two affricates are compared to the rise time of the apical palatal fricative [ʃ], in order to test the hypothesis that one of the two affricates is more fricative-like than the other. Similarly, maximum rate of rise (Weigelt et. al 1990) of the two affricates is compared to that of the stops in order to test the hypothesis that one of the affricates is more stop-like.

3.2 Methods

Six subjects between the ages of 23-27 were recorded producing words with the two types of contrastive palatal affricates [tʃ] and [tŝ], the single palatal fricative [ʃ], and the closest stops in place of articulation, [t] and [k], in the onset position of the initial accented syllable, or in the onset of the (second) final unaccented syllable in bisyllabic words. Vowel length was controlled, and all words contain phonemically short vowels in both syllables. There were 79 tokens total, with 5 tokens of each consonant type in word-initial position (one for each vowel context), and 20 for each consonant type in word-medial position (in five different vowel contexts). Five repetitions of the wordlist in the Appendix were produced. Words were all produced in the phrasal context *Reci* ____ *deset puta.* 'Say ___ ten times'. Thus in word-initial position, the preceding vowel is always [i].

Productions were recorded on a PC computer at IEFPG (Institute of Experimental Phonetics and Speech Pathology) in Belgrade, Yugoslovia. Speech was sampled at the rate of 22,050 Hz. Recordings were then transferred to SUN Sparc stations in the Cornell University Phonetics Laboratory for acoustic analysis.

Each production was segmented and labeled using the labels provided in (1):

(1) Labels used in this study

Label	Acoustic Landmark	Acoustic attribute used to label landmarks
CB	Closure Begin	Point where first formant dies out
FB	Frication Begin	Point where sustained frication noise begins
VB	Vowel Begin	begin of first formant
PFA	Peak Frication Amplitude	Highest amplitude peak identified in the frication interval (ignoring spit spikes)

Rise time (RT) was calculated as the time between the FB and PFA labels. Closure duration was calculated as CB-FB in the case of affricates, or CB-VB in the case of stops. Frication duration was calculated as the difference between the beginning of frication noise (FB label) and the beginning of the following vowel (VB label). MAX ROR was calculated, by identifying a 20 ms window in which to search for the greatest change in amplitude. This window was centered over the labeled vowel beginning (VB) for stops (e.g. 10 ms before the onset of F1), and centered over the labeled beginning of sustained frication noise in affricates and fricatives (FB label). MAX ROR was calculated as the largest difference in amplitude between any two successive points within this 20 ms window. A representative waveform and spectrogram of a medial laminal affricate token is provided in Figure 4 to illustrate how the labels were placed.

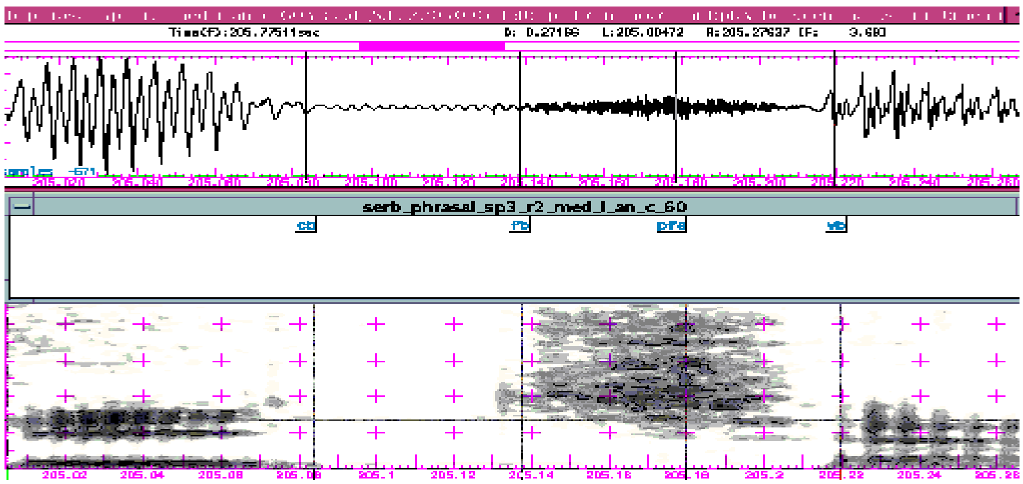

Figure 4: Labeled waveform and spectrogram for the word *sićan* [sitŝan] produced by Subject M1

Figure 5 provides a waveform and a spectrogram of a representative token of a medial laminal affricate, with the labels exemplified.

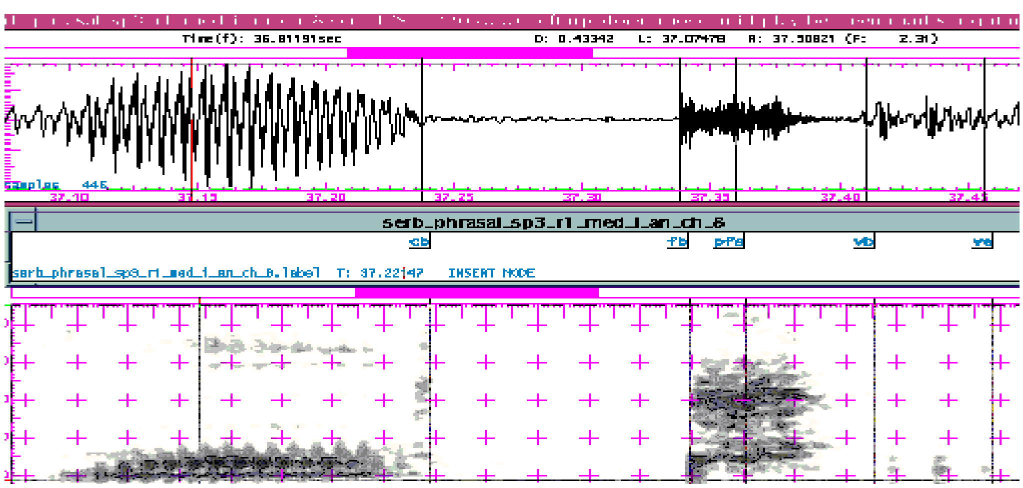

Figure 5: Labeled waveform and spectrogram for the word *vičan* [vitʃan] produced by Subject M1

Temporal differences associated with the two affricates are seen clearly by comparing the spectrograms and waveforms in Figures 4 and 5, respectively. As can be seen, the medial laminal affricate in Figure 4 displays a short closure duration and a long frication duration relative to the apical affricate in Figure 5, which displays a relatively longer closure duration and shorter frication duration. Further, the laminal affricate displays a longer rise time than the apical affricate. Quantitative results are provided in Section 3.3.

3.3 Results

Results for word-initial accented syllable onset position are presented first in Section 3.3.1, followed by results for word-medial unaccented syllable onset position presented in Section 3.3.2. In Section 3.3.3, we discuss the results of the two contexts in parallel.

3.3.1 Word-initial accented position

Figure 6 plots frication duration against closure duration. As can be seen, all of the subjects consistently differentiate the two affricates in terms of the inversely proportional duration of each of the phases. There is a negative correlation between the two durational phases of the affricates, which may well be due to like targets for overall segment durations, which are consistently maintained. Subject W2 differentiates the two classes of affricates better on this dimension than subject M1, although there is no overlap between the two categories for either subject. Subject W3's productions display a different pattern. She seems to be distinguishing frication duration to a larger extent than closure duration, as seen by the clearer separation on the y-axis than on the x-axis in Figure 6. Subject M3 displays a larger separation on the x-axis, showing a larger distinction in terms of closure duration, than is found in frication duration.

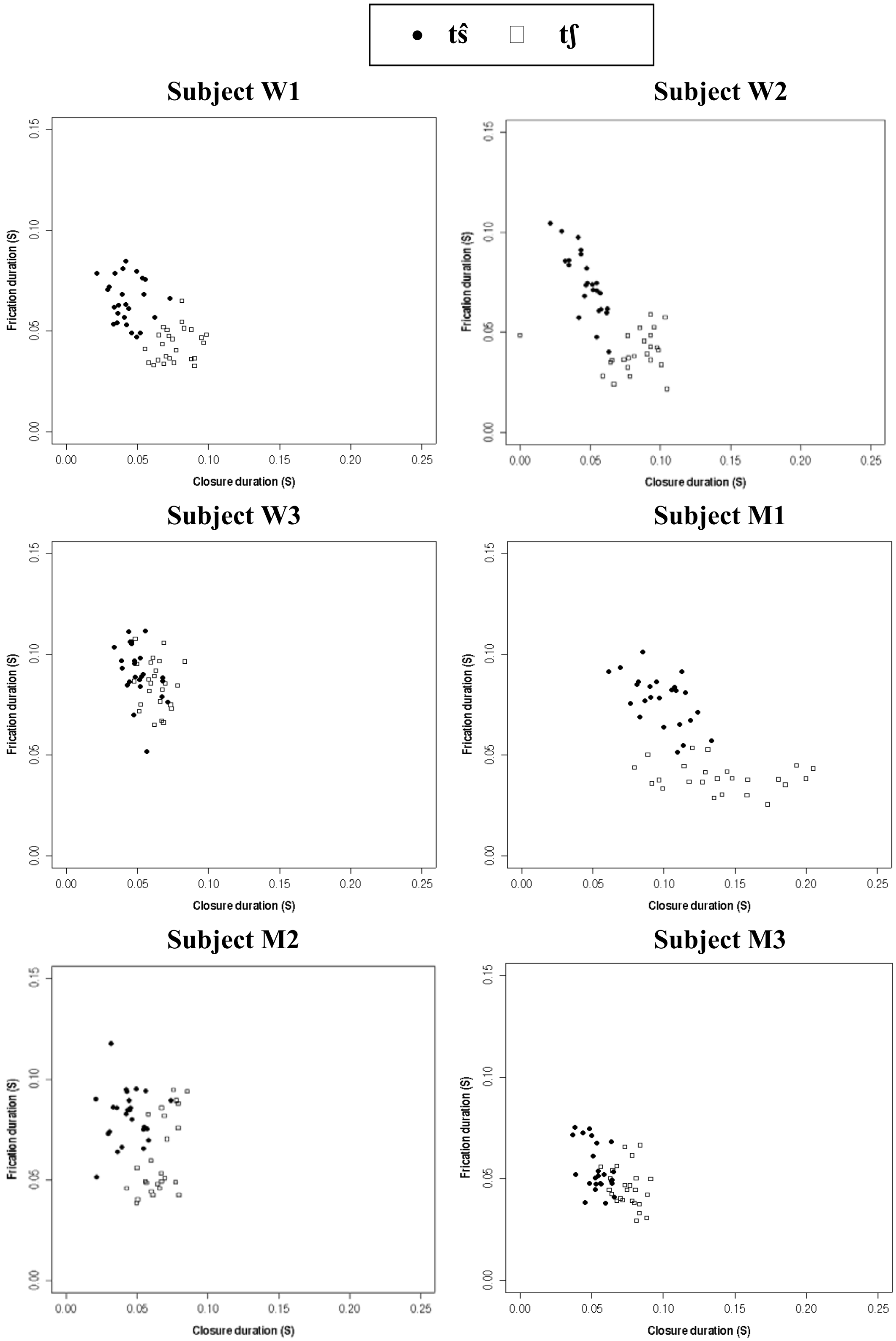

Figure 6: Frication duration by closure duration in word-initial accented syllable onset position

Table IV provides the adjusted R^2 values associated with a regression, with closure duration and manner of articulation against frication duration.

Subject	Adjusted R^2	p-value closure duration	p-value manner of articulation
W1	0.94	$p<.001$	$p<.001$
W2	0.91	$p<.001$	$p<.001$
W3	0.90	$p<.001$	$p<.001$
M1	0.94	$p<.001$	$p<.001$
M2	0.93	$p<.001$	$p<.001$
M3	0.97	$p<.001$	$p<.001$

Table IV: Adjusted R^2 values associated with a regression of closure duration and manner of articulation against frication duration

Figure 7 compares closure duration for the two affricates and coronal and dorsal stops in initial position. As can be seen, both the affricates and stops display similar closure duration differences. The laminal affricate has a shorter closure duration similar to the coronal [t], while the apical affricate displays slightly longer closure durations similar to the dorsal stop [k]. Most of the speakers display much shorter closure durations overall for the stops than the affricates.

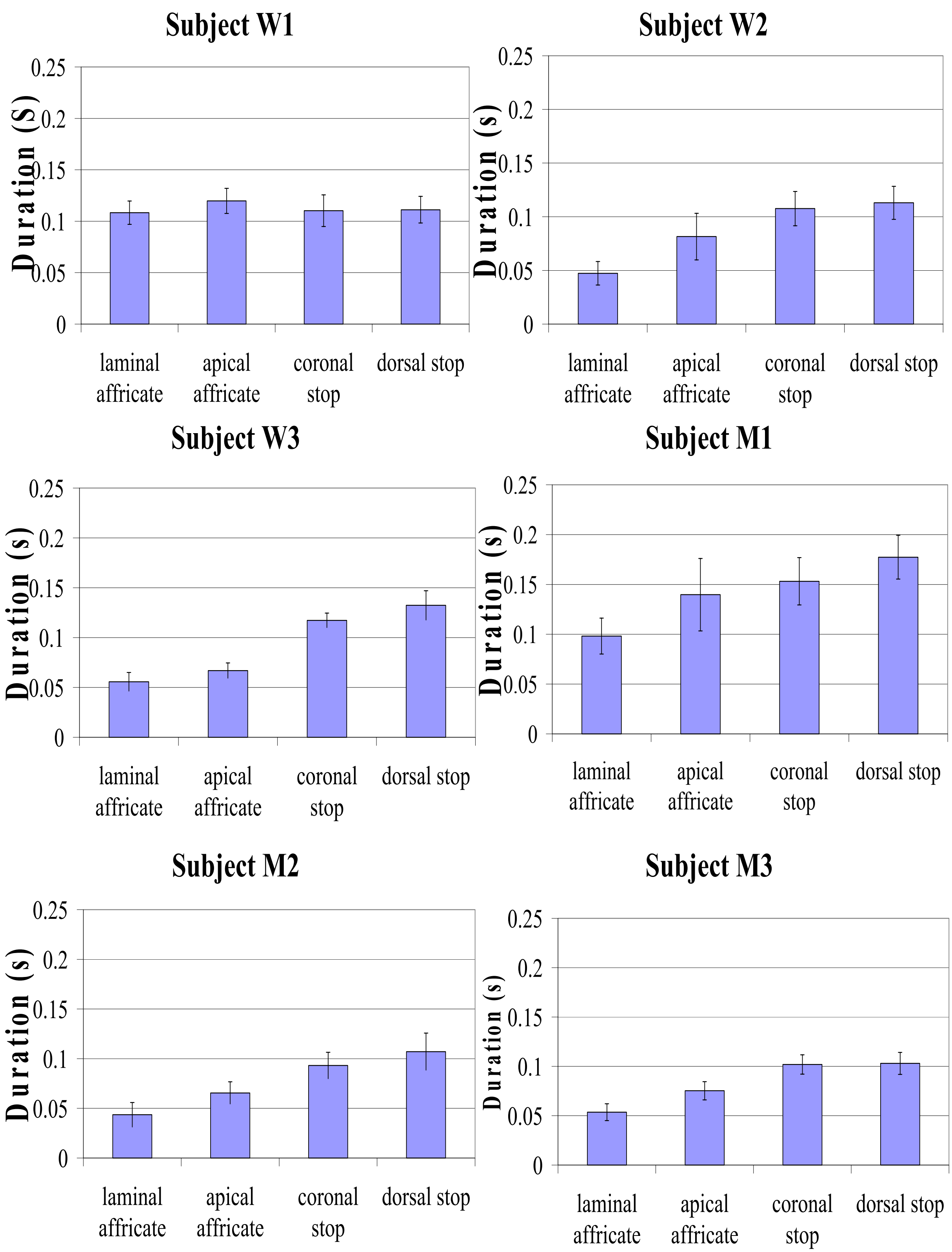

Figure 7: Closure duration in word-initial accented syllable onset position

Figure 8 displays rise time plotted against frication duration. Notice that in addition to the two affricates, these graphs also include the fricative [ʃ]. As can be seen in the graph, there is a positive correlation between frication duration and rise time. That is, as frication duration increases, rise time also increases. Given the inverse correlation between closure duration and frication duration for the two affricates shown in Figure 6 above, this means that there is also an inverse correlation between closure duration and rise time. That is, as closure duration increases, rise time decreases.

While there is a linear correlation between frication duration and rise time for the two affricates, the rise time associated with the fricative [ʃ] displays more variability. That is, the frication duration is consistently longer for [ʃ] than for the two affricates, in keeping with the goal of having all segments under study with similar segment durations, as shown above in Figure 6. But, since there is no closure preceding the frication interval in the fricative, there is no pressure buildup that constrains the rise time in the frication noise, and the distribution of the rise time values is rather large. However, Subjects W2 and M2 do maintain a linear relationship between rise time and frication duration even for the fricative [ʃ]. For these two subjects, the rise time in the fricatives is much longer than the rise time in the affricates, similar to the results found for [tʃ] vs. [ʃ] in English by Howell and Rosen (1983).

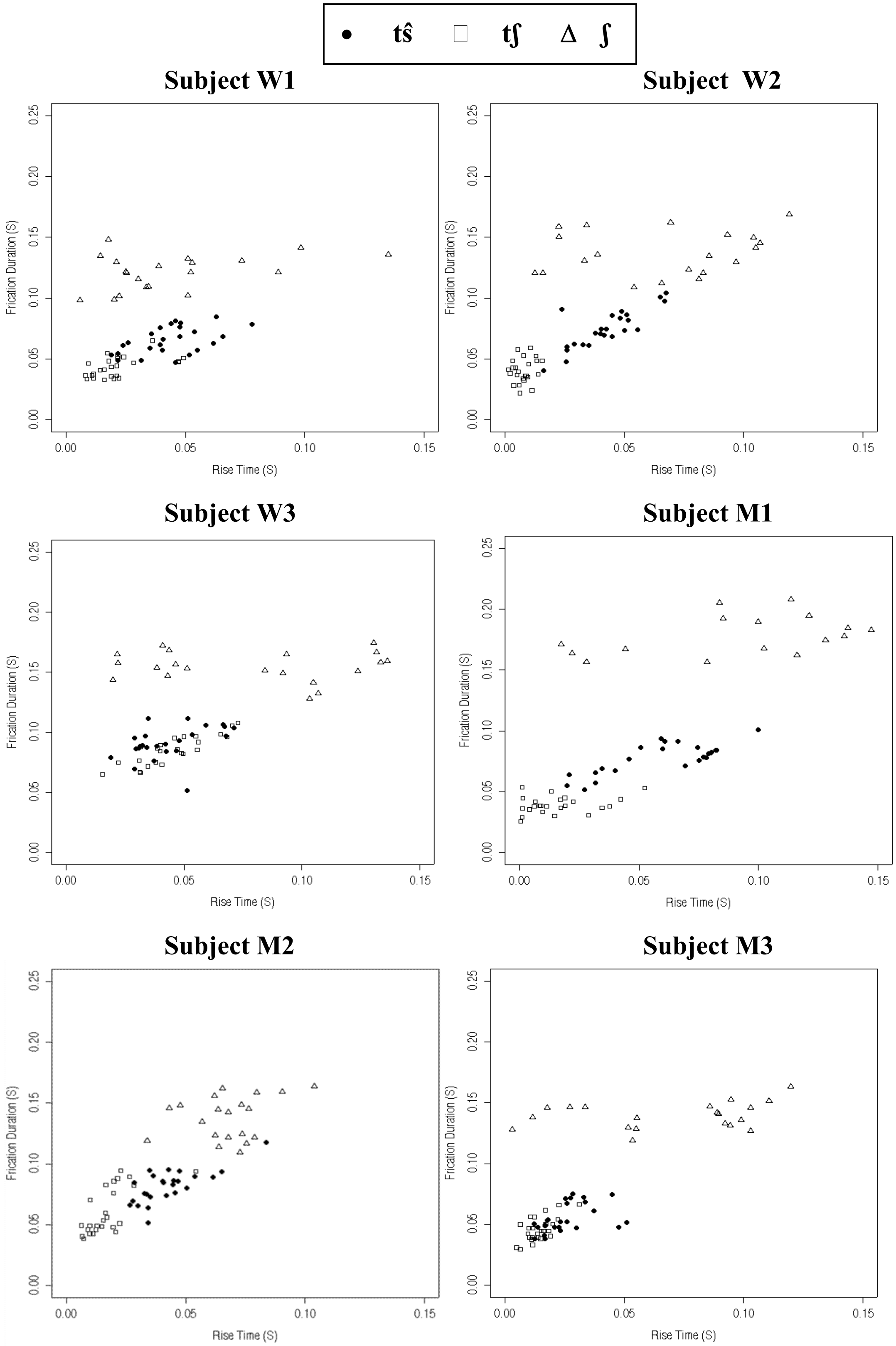

Figure 8: Rise time by frication duration in word-initial accented syllable onset position

The R-squared values for the correlations between frication duration against rise time are provided in Table V. The initial column provides the R^2 values with only frication duration as a predictor, and the third column provides R^2 values with both manner of articulation and frication duration as predictors. As can be seen, the R^2 values are much higher when manner of articulation is included. This is because of the lower correlation between rise time and frication duration for the fricatives. An ANOVA showed that the two R^2 values differ significantly, confirming that manner of articulation adds significantly to the prediction of rise time, showing that the affricates behave as a class phonetically as opposed to the fricative [ʃ].

Subject	Adjusted R^2 with frication duration	p-value frication duration	Adjusted R^2 with both manner and frication duration	p-value manner of articulation	p-value frication duration
W1	R^2 =0.01	NS	R^2 = 0.95	$p<.001$	$p<.001$
W2	R^2 =0.03	$p<.05$	R^2 = 0.93	$p<.001$	$p<.001$
W3	R^2 =0.26	$p<.001$	R^2 = 0.93	$p<.001$	$p<.001$
M1	R^2 =0.51	$p<.001$	R^2 = 0.96	$p<.001$	$p<.001$
M2	R^2 =0.18	$p<.001$	R^2 = 0.94	$p<.001$	$p<.001$
M3	R^2 =0.27	$p<.001$	R^2 = 0.97	$p<.001$	$p<.001$

Table V: Adjusted R^2 values for a regression of both frication duration and manner of articulation against rise time (120 df)

We now turn to the relationship between maximum rate of rise (MAX ROR) and closure duration. Weigelt et. al (1990) showed that MAX ROR differentiates stops from affricates in English. In Figure 9, MAX ROR is plotted against closure duration. This figure contains both of the affricates, the fricative [ʃ], as well as the stops [t] and [k]. As can be seen, MAX ROR is more adept at capturing the relationship between abruptness and closure duration than rise time is. Fricatives, which have the smallest closure durations (zero), exhibit the smallest MAX ROR values. However, the ROR values for fricatives are also widely distributed for some speakers (Subjects W1, W2 and W3). The affricate [tŝ] exhibits a smaller closure duration than the affricate [tʃ], which corresponds to smaller MAX ROR values. The stops exhibit the largest MAX ROR values, having the longest closure durations. The two stops are not differentiated among themselves. The patterns seem to be displaying a universal relationship between closure duration and MAX ROR, which has its basis in the relationship between closure duration and pressure buildup. This relationship has only been demonstrated for affricates *vs.* stops in English (Weigelt et. al 1990), but has never been shown to be maintained in contrastive affricates. This supposed universal results in Serbian consonants displaying a four-way contrast in abruptness, which is nicely captured in a continuum by the acoustic measure of MAX ROR here.

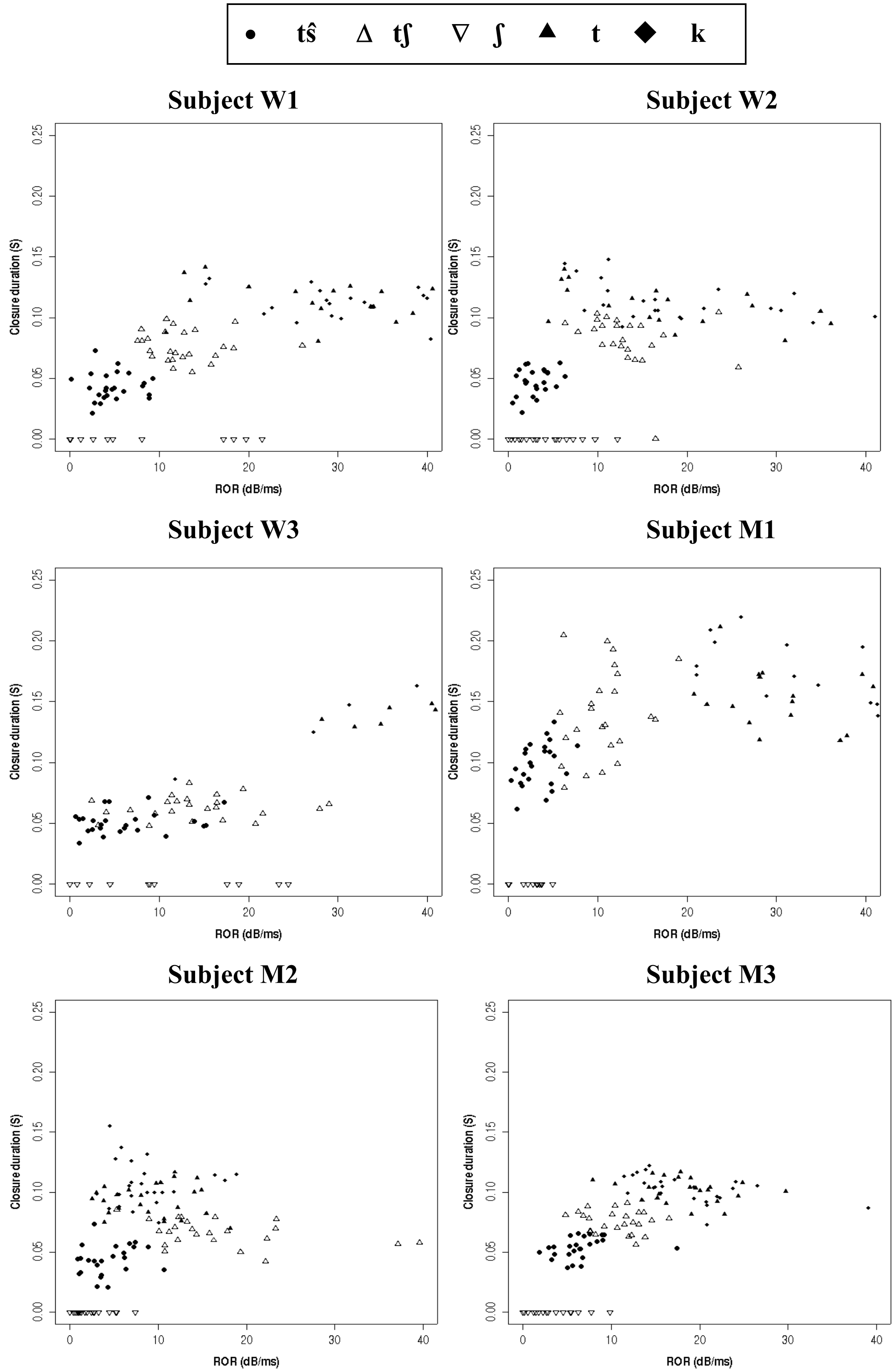

Figure 9: MAX ROR by closure duration in word-initial accented syllable position

Table VI provides R^2 values associated with regressions of MAX ROR against both closure duration and manner of articulation for each subject. Notice that there is a high degree of inter-speaker variability for the degree of correlation in this context.

Subject	Adjusted R^2 (affricates only)	p-value manner	p-value closure duration
W1	$R^2 = 0.55$	p<.001	NS
W2	$R^2 = 0.32$	p<.001	NS
W3	$R^2 = 0.76$	p<.001	NS
M1	$R^2 = 0.68$	p<.001	p<.01
M2	$R^2 = 0.15$	p<.001	NS
M3	$R^2 = 0.64$	p<.001	NS

Table VI: Adjusted R^2 values for manner of articulation and closure duration against MAX ROR (120 df)

3.3.2 Word-medial unaccented syllable onset position

Figure 10 plots closure duration against frication duration for all speakers. As can be seen, the two measures provide a categorical distinction. Subject W2 maintains the clearest separation on these measures in this position. The other five subjects' productions display a considerable shrinkage of the acoustic space in this prosodic context.

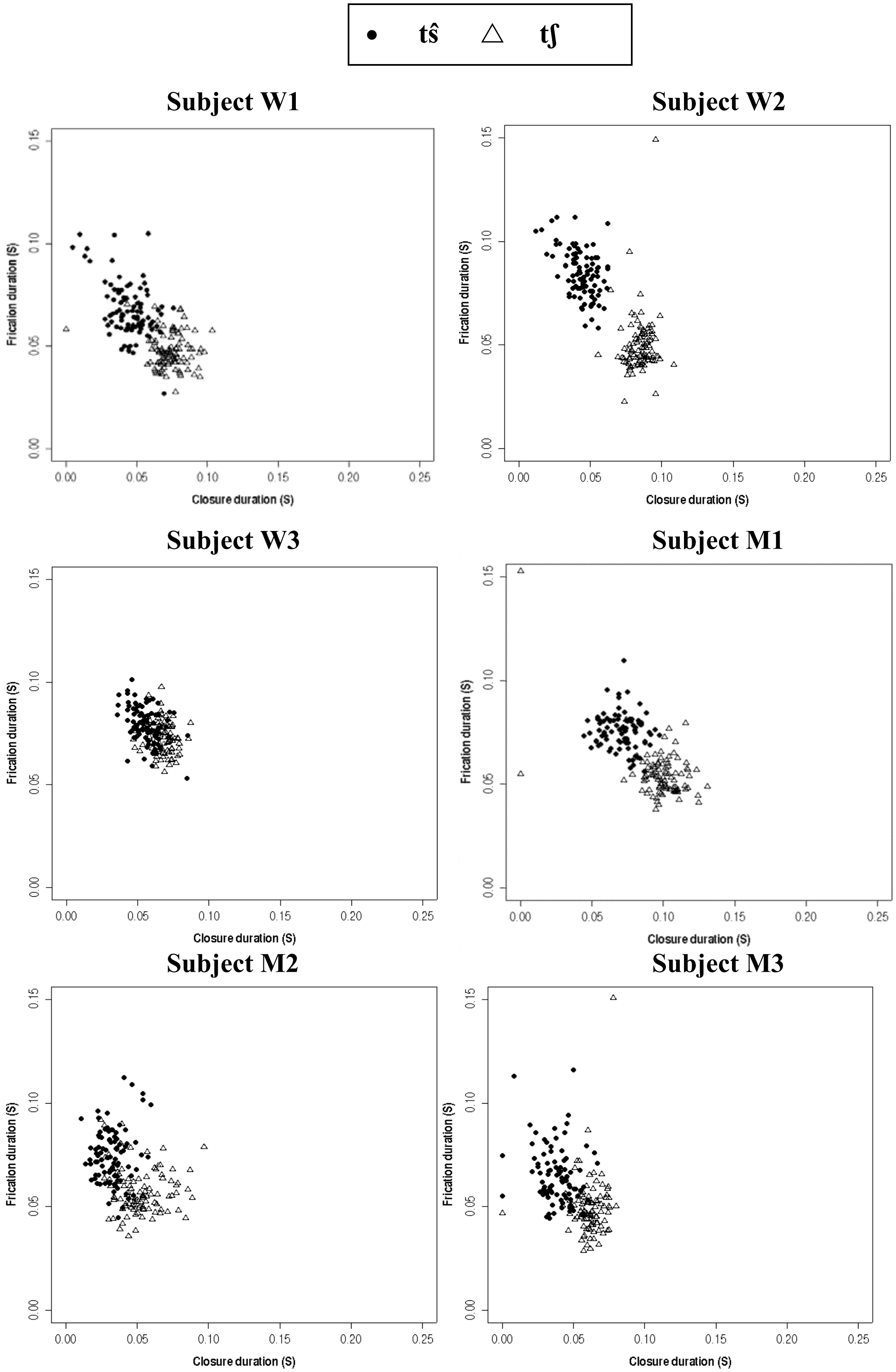

Figure 10: Frication duration by closure duration in word-medial unaccented syllable onset position

Table VII provides adjusted R^2 values associated with regressions of closure duration against frication duration for each subject.

Subject	Adjusted R^2	p-values closure duration	p-values manner
W1	$R^2 = 0.90$	p<.001	p<.001
W2	$R^2 = 0.75$	NS	p<.001
W3	$R^2 = 0.95$	p<.001	p<.001
M1	$R^2 = 0.92$	p<.001	p<.001
M2	$R^2 = -0.009$	NS	NS
M3	$R^2 = -0.0003$	NS	NS

Table VII: Adjusted R^2 values associated with a regression of closure duration and manner of articulation against frication duration (265 df)

Figure 11 provides closure durations for the two affricates and the coronal and dorsal stops in word-medial unaccented syllable onset position. In this prosodic context, durational differences associated with place of articulation and tongue constriction location become more apparent than they were in initial accented syllable onset position. Again, we see relatively shorter durations for the coronal stop and the laminal affricate, and relatively longer durations for the dorsal stop and the apical affricate. Interestingly, in medial position, all subjects display these closure duration differences, even though they likely do not all display place of articulation differences, as was shown by the other 34 subjects investigated by Miletić (1933).

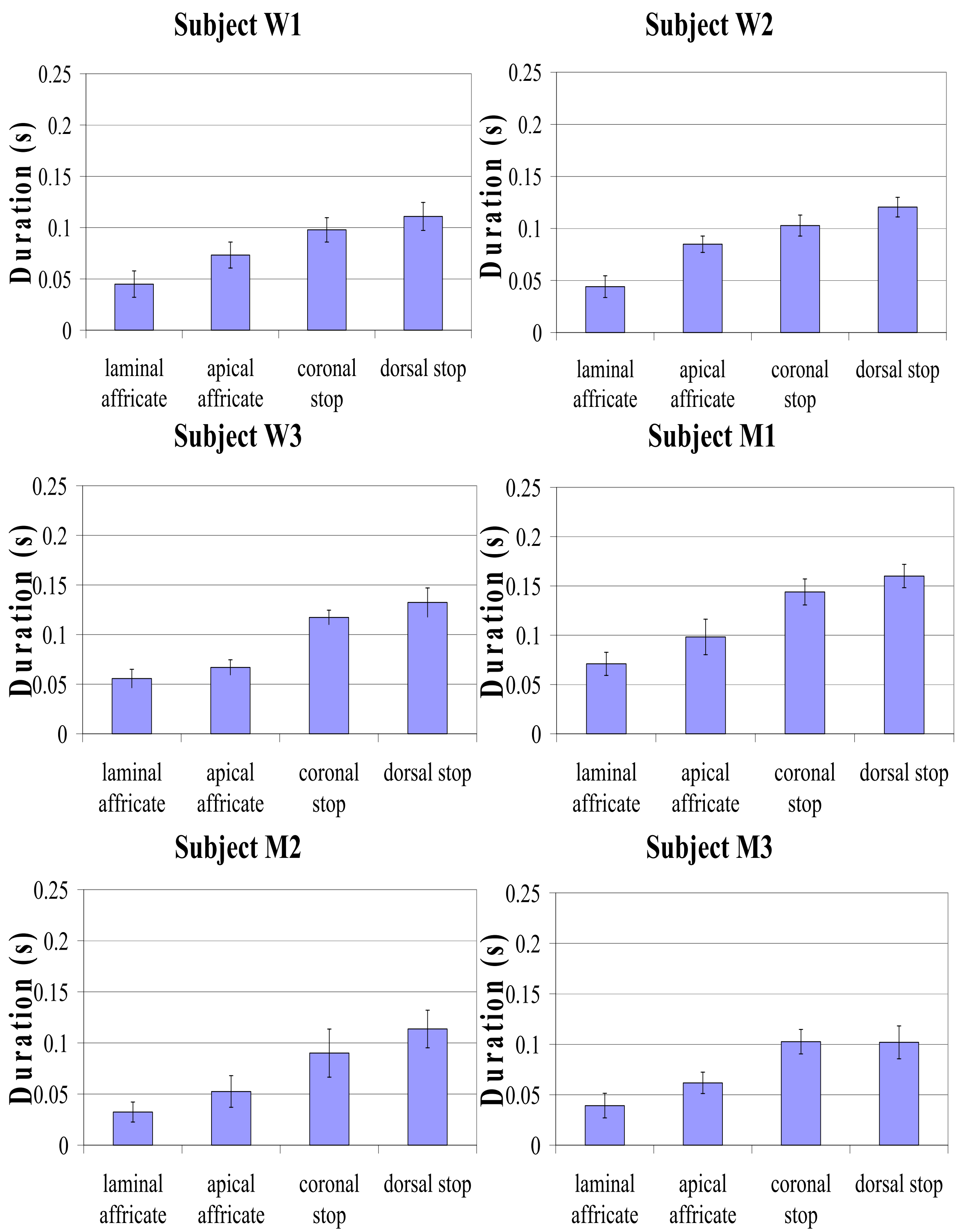

Figure 11: Closure duration in word-medial unaccented syllable onset position

Figure 12 displays rise time plotted against frication duration. For most subjects, there is a clear linear relationship between the two measures. However, for half of the subjects (subjects W1, W2 and M1), the frication duration results are much more widely distributed for the fricative than for the two affricates.

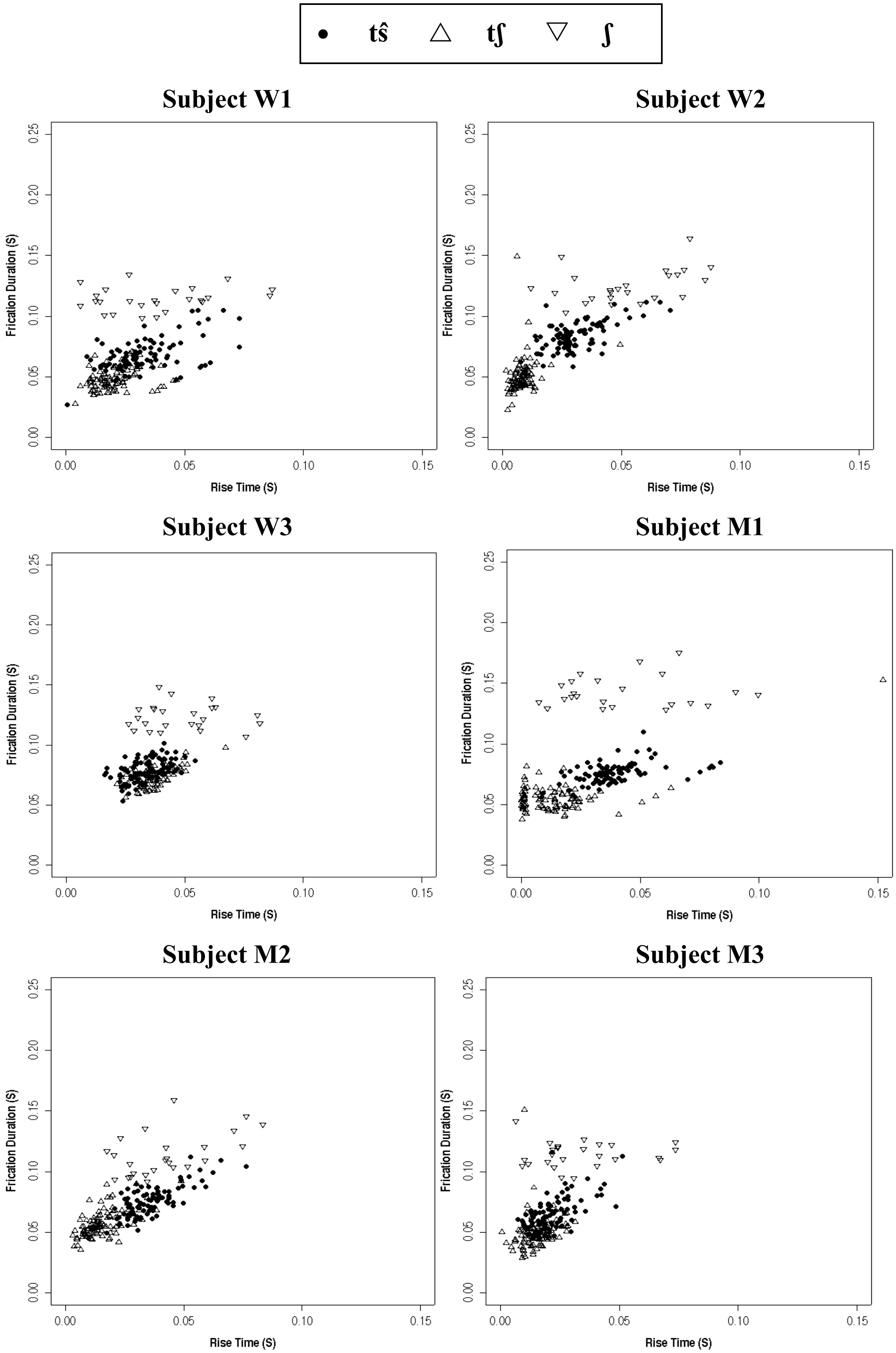

Figure 12: Rise time by frication duration in word-medial unaccented syllable onset position

Table VIII provides adjusted R^2 values associated with regressions of frication duration and manner of articulation against rise time for each subject. As can be seen, manner of articulation contributes a great deal above and beyond frication duration to the prediction of rise time.

Subject	Adjusted R^2	p-value frication duration	Adjusted R^2	p-value manner	p-value frication duration
W1	$R^2 = 0.65$	p<.001	$R^2 = 0.86$	p<.001	p<.001
W2	$R^2 = 0.21$	p<.001	$R^2 = 0.86$	p<.001	p<.001
W3	$R^2 = 0.44$	p<.001	$R^2 = 0.95$	p<.001	p<.001
M1	$R^2 = 0.29$	p<.001	$R^2 = 0.93$	p<.001	p<.01
M2	$R^2 = .002$	NS	$R^2 = -0.001$	NS	NS
M3	$R^2 = 0.61$	p<.001	$R^2 = 0.65$	p<.001	p<.001

Table VIII: Adjusted R^2 values associated with a regression of manner of articulation and frication duration against rise time (261 df)

Figure 13 displays closure duration plotted against MAX ROR. Subjects W1 and W3 display a greater degree of variability in the MAX ROR values for the fricatives, for which there is no pressure buildup during a preceding closure. The other four subjects' (W2, M1, M2 and M3) productions display an almost linear relationship for all of the segment types.

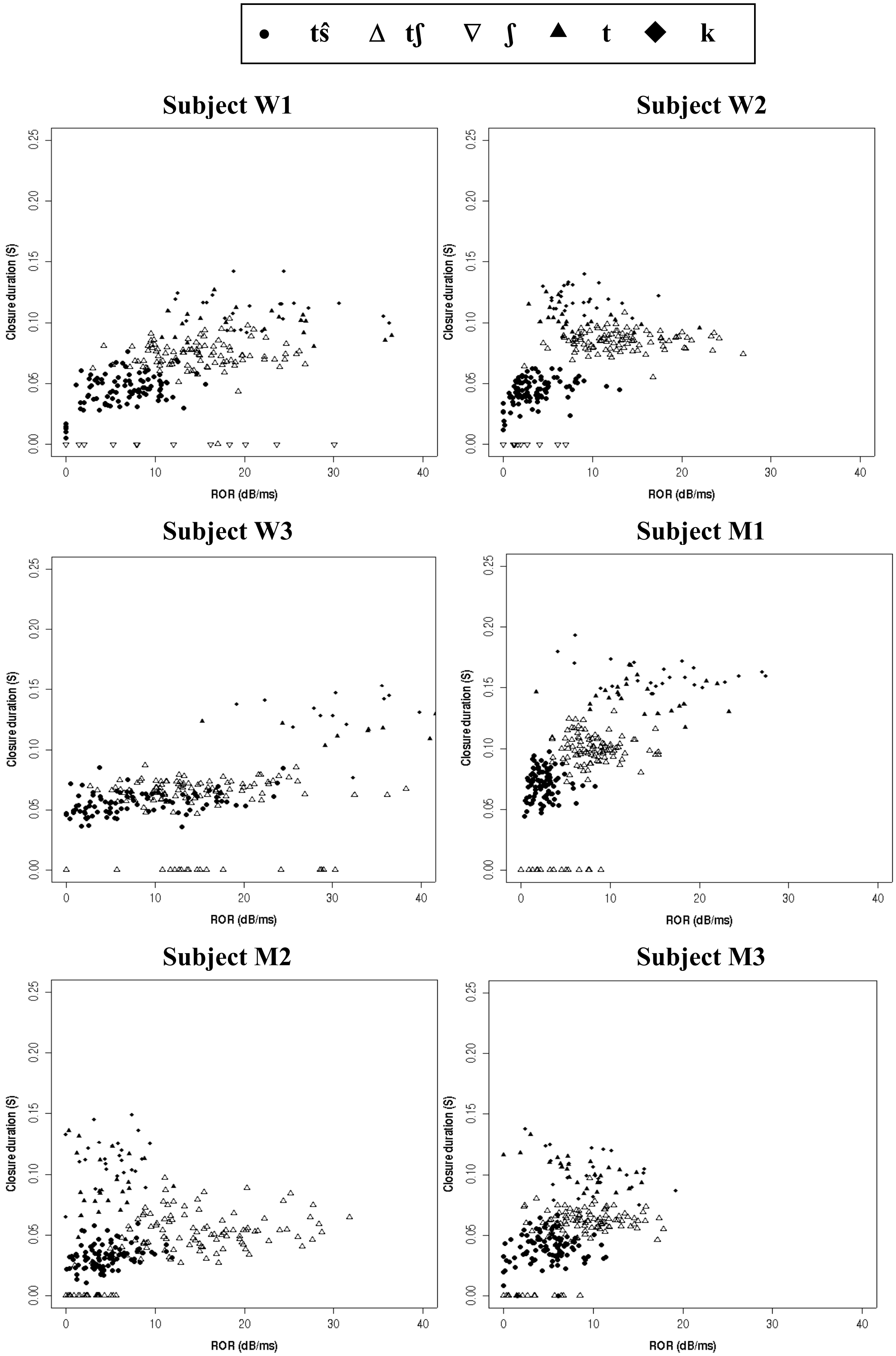

Figure 13: MAX ROR by closure duration in word-medial unaccented syllable onset position

Table IX provides adjusted R^2 values associated with regressions of MAX ROR against closure duration and manner of articulation for each subject.

Subject	Adjusted R^2	p-value manner	p-value closure duration
W1	$R^2 = 0.42$	p<.001	p<.001
W2	$R^2 = 0.11$	p<.001	NS
W3	$R^2 = 0.67$	p<.001	p<.001
M1	$R^2 = 0.55$	p<.01	p<.001
M2	$R^2 = 0.33$	p<.001	p<.001
M3	$R^2 = 0.26$	p<.05	p<.05

Table IX: R^2 values associated with a regression of manner of articulation and closure duration against MAX ROR (261 df)

3.4 Discussion

The results of Experiment 1 show that the two affricates investigated here consistently differ in closure duration, frication duration, rise time and MAX ROR. Only one of the subjects, Subject W3, displays a weak distinction between the two affricates on the measures of frication duration and closure duration. The two phases of the affricates are proportional, and add up to similar segment durations for the two affricates, as well as for the fricative [ʃ], and the stops [t] and [k]. Similarly, all subjects display closure duration differences that may be linked to place of articulation (see Maddieson 1997 and references therein) for both the stops and the affricates. The temporal differences linked to place are larger in word-medial unaccented syllable onset position than in word-initial accented syllable onset position. The shrinkage of the acoustic space in medial position is due to the weak prosodic context.

Rise time and MAX ROR values are categorically distinct for the two affricates in all subjects' productions collected in this study. Rise time is correlated to frication duration, and MAX ROR is correlated to closure duration, suggesting physiological linkages between the duration of affricate phases, Rise time and MAX ROR. However, the phonological class of manner of articulation also contributes a great deal to the prediction of both rise time and MAX ROR. MAX ROR and closure duration are better correlated in word-medial unaccented syllable onset position than they are in word-initial accented syllable onset position, suggesting that MAX ROR may be weighted more highly in the perception of the two affricates and fricatives in word-medial position. The variability seen in this segmental and prosodic context is due to an increased degree of frication noise present in the closure interval of the stops and the affricates, with the laminal affricate that has the shortest closure duration exhibiting the highest degree of frication noise. The degree of frication of stops and affricates in word-medial unaccented syllable onset position is speaker specific, and exhibits a high degree of variability. Therefore, it interrupts the straight-forward relationship between closure duration and pressure build-up, which determines the MAX ROR directly, since some airflow escapes during fricated closures.

However, for some speakers, fricatives display a wide range of variability in rise time, and stops display a wide range of variability in MAX ROR values, due to the lack of such physiological constraints. Still, some subjects maintain the linear relationship between closure duration and rise time found with affricates even in the production of fricatives, suggesting that these subjects may be manipulating the relevant acoustic cues for the aid of the listener. These subjects' productions are suggestive that Serbian listeners would behave similar to English listeners in using both rise time (Howell and Rosen 1983) and Maximum ROR (Weigelt et. al 1990) to distinguish fricatives from affricates, and stops from affricates, respectively. We are in the process of undertaking perceptual experiments to explicitly test the weighting of the different temporal and spectral cues in different prosodic contexts.

A similar physiologically-based relationship is found between rise time and frication duration for the two affricates. This separates the affricates as a class from the fricatives, where the rise time is not determined by the degree of pressure buildup, given the complete lack of closure. Similarly, the two affricates display a strong correlation between MAX ROR and closure duration, which is not found for the stops. These two phonetic patterns, along with the fact that [tʃ] and [tŝ] both display rather long closure durations and frication durations, provide evidence that the affricates behave phonetically as a separate class from both fricatives and stops.

4. Experiment 2: Spectral attributes of Serbian obstruents

4.1 Introduction

In Experiment 2, we investigate the spectral properties found in the frication noise of the two voiceless palatal affricates and the apical fricative [ʃ] as well as in the stop noise-bursts of [t] and [k], in order to test the hypothesis that the individual articulatory differences found between the two affricates cumulatively result in a consistent and relatively larger difference in the cavity volume in front of the constriction. Further, we hypothesize that spectral differences found in the frication noise of the affricates parallel spectral differences between the spectra of stop noise-bursts of [t] and [k], that are known to be higher for [t] than for [k]. Specifically, we investigate the spectral center of gravity, the frequencies and amplitudes of two prominent spectral peaks, and the differences in the frequency and amplitude values of these two identified peaks. We also investigate the second formant frequency differences found in the vowel following the affricates. In Section 4.2, we provide the methodology used in the experiment. In Section 4.3.1, we provide the results in word-initial accented syllable onset position and in Section 4.3.2, we provide the results in word-medial unaccented syllable onset position. We discuss the two sets of results in parallel in Section 4.3.3.

4.2 Methods

The same data used for Experiment 1 was used for Experiment 2. Spectral attributes measured are the centroid frequency of the spectrum, frequency and amplitude values of two prominent spectral peaks found in the 2500-5000 Z range, and the 6000-8500 Z ranges, as well as the average of these two spectral peaks, the difference between the frequencies of the two peaks, and the difference between the amplitude of the two peaks.

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. For both stops and affricates, the original FFT spectra were converted to bark-scaled spectra using the ESPS program *barkspec*, which uses the algorithm described in Hermensky (1990), Wang et al. (1992) and Sekey and Hanson (1994). Bark spectra are scaled to reflect frequency processing by the human auditory apparatus. Speaker-specific and consonant-specific ranges for each of the two spectral peaks were identified through visual inspection of all spectra of each type simultaneously, and then a peak-picking algorithm was used to objectively identify the peak frequencies within each of the identified ranges for each of the tokens. For all but one of the speakers, speaker-specific ranges were identified which captured the peak frequencies for each consonant across all vowel contexts. A vowel-specific range needed to be identified for the [u] context for [tʃ] for Subject M2. Speaker-specific ranges used are provided in Table X. The ranges used for [ʃ] were identical to the ranges used for [tʃ] for all subjects.

	[tʃ]		[tŝ]	
Subject	P1 frequency range	P2 frequency range	P1 frequency range	P2 frequency range
W1	2000-4500	5000-7500	3000-5500	7000-9500
W2	1500-4000	5000-7500	3000-5500	6000-8500
W3	1500-3900	5000-7500	3000-5500	6000-8800
M1	1500-4000	5000-7500	3000-5500	7000-9500
M2	900-3400 [u] 900-2700	4100-7000	2000-5000	5100-9000
M3	1500-4000	5000-7500	3000-5500	7000-9500

Table X: Speaker-Specific ranges for P1 and P2 frequencies

Figure 14 provides bark spectra for the two affricates in initial position in the i# _i context with the search ranges provided in Table XI delimited on the graphs with lines.

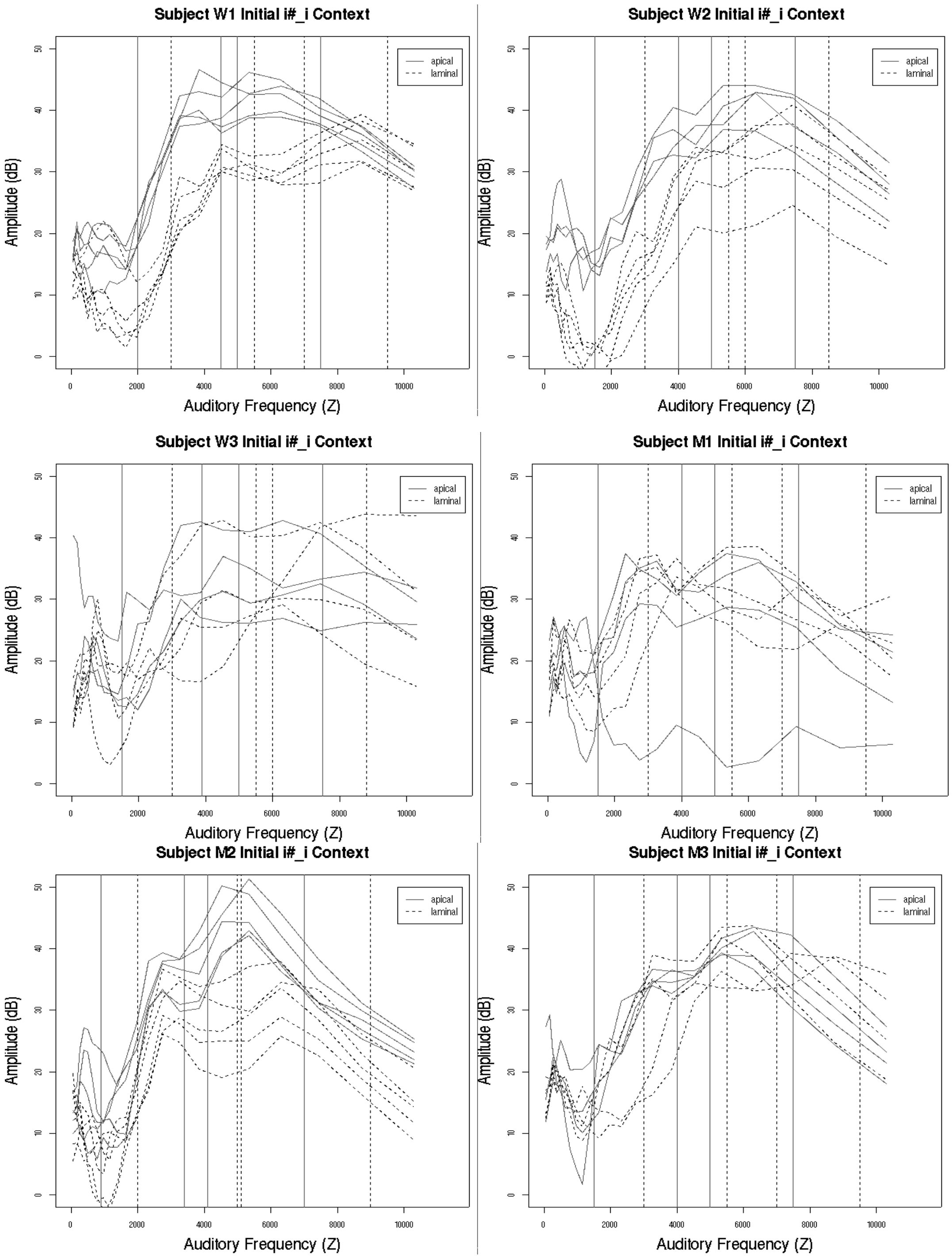

Figure 14: Bark-scaled spectra for the two palatal affricates in initial accented syllable onset position in the i# _i context

Second formant frequency (F2) values were identified using LPC spectra created over a 46.5 ms. window created with 1024 points and 18 LPC coefficients appropriate for files sampled at 22,050 Hz. F2 values were calculated at two different time points: (1) at the onset of the vowel, with the window aligned to the beginning of the vowel, and (2) with the window beginning 30 ms into the vowel.

4.3 Results

Results for word-initial accented syllable onset position are provided in Section 4.3.1, and the results for the word-medial unaccented syllable onset position are provided in Section 4.3.2. Results are discussed in parallel in Section 4.3.3.

4.3.1 Word-initial accented syllable onset position

Figure 15 plots Peak 1 frequency (P1) against Peak 2 frequency (P2) for all six subjects. P1 frequency categorically distinguishes the two affricates for subjects W1, W2, M1, and W3 (but not for subjects M2 and M3). P1 values for the fricative [ʃ] are similar to those found for the affricate [tʃ]. Subject M2 actually displays a slight difference in the opposite direction for P1 for the two affricates, but with complete overlap of the two categories on this measure. Subject M3 only displays slight overlap in the distribution of P1 values for the two affricates.

P2 frequency categorically distinguishes the two affricates for all speakers, with subjects W1, M1 and M3 displaying the greatest differences, as shown in Figure 15. Both P1 and P2 values are higher in the frication interval of [t̂ŝ] than that seen in Polish [ç̦], which has a peak associated with F2 at around 3000-4000 Hz (Halle and Stevens 1971), and higher than the frication noise associated with Abkhaz [ŝ], which has energy concentrated at about 3000 Hz in the spectrogram in Ladefoged and Maddieson (1996: 162). It is difficult to know if the slightly lower frequency energy seen in the Abkhaz fricative reported in Ladefoged and Maddieson is an individual difference rather than a linguistic difference, as one of our Serbian subjects (subject M2) also displays a similarly lower frequency P1, and the Polish results are for just two subjects, and there is no quantitative study of the Abkhaz fricatives.

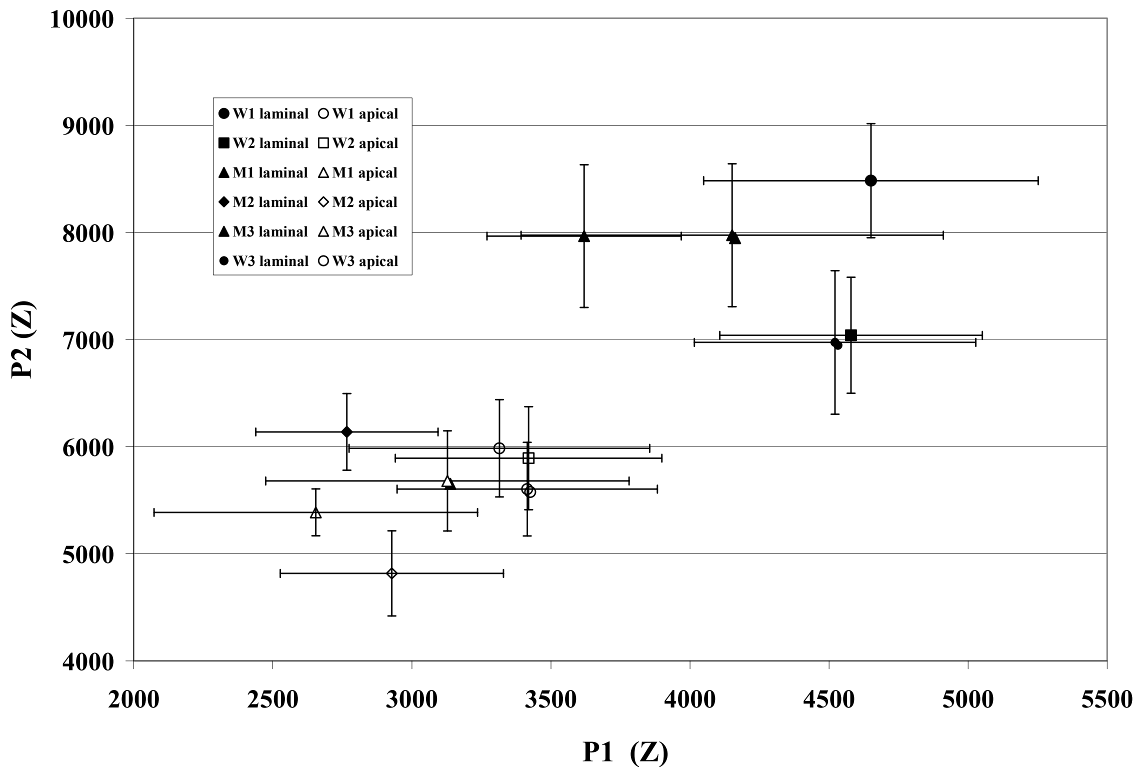

Figure 15: P1 plotted against P2 for six speakers of Standard Serbian in word-initial accented syllable onset position

A two-factor ANOVA was calculated with SPEAKER and CONSONANT TYPE as factors for both P1 and P2 frequencies. Both CONSONANT TYPE and SPEAKER were significant at the level of p<.001 for all subjects for each factor. P-values associated with post-hoc pair-wise comparisons using Holm's (1979) adjustment method are provided in Table XI.

	P1 frequency	P2 frequency
W1	p<.001	p<.001
W2	p<.001	p<.001
W3	p<.001	p<.001
M1	p<.001	p<.001
M2	NS	p<.001
M3	p<.001	p<.001

Table XI: Post-hoc pair-wise comparison values for the two affricate types using Holm's adjustment method

Figure 16 provides the values of a single spectral peak found during the stop noise-burst between 2000-5000 Hz for all subjects for [t] and [k].

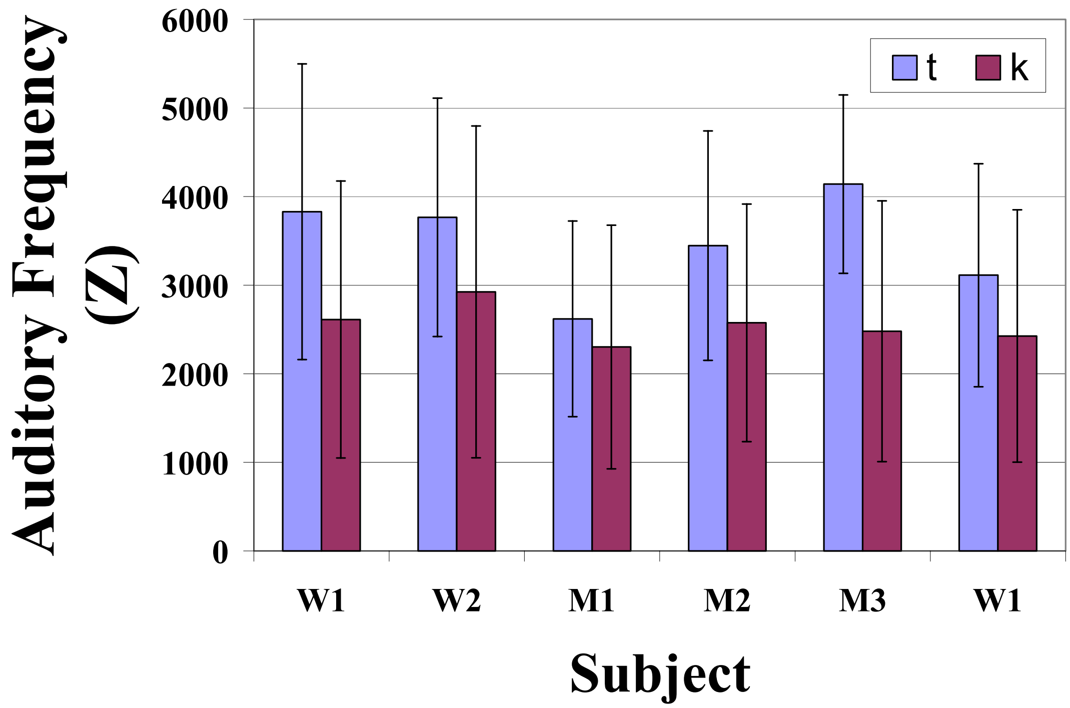

Figure 16: Frequency of a prominent spectral peak found in the noise-bursts of the stops [t] and [k] in word-initial accented syllable onset position in i#_[a,e,i,o,u] context

As can be seen, [t] has a higher peak similar to [tŝ], while [k] has a lower frequency peak similar to [tʃ].

P1-P2 values are plotted in Figure 17. These provide a way of assessing the compactness of the spectra. As can be seen in Figure 17, there is a tendency for there to be a larger degree of separation between the two peaks for [tŝ] than for [tʃ], although Subject W2 does not produce the distinction. Subject W3's productions display a smaller but consistent difference, since she displays less variability. This suggests that in initial position, [tʃ] is more compact than [tŝ], thus providing another acoustic difference between the two affricates.

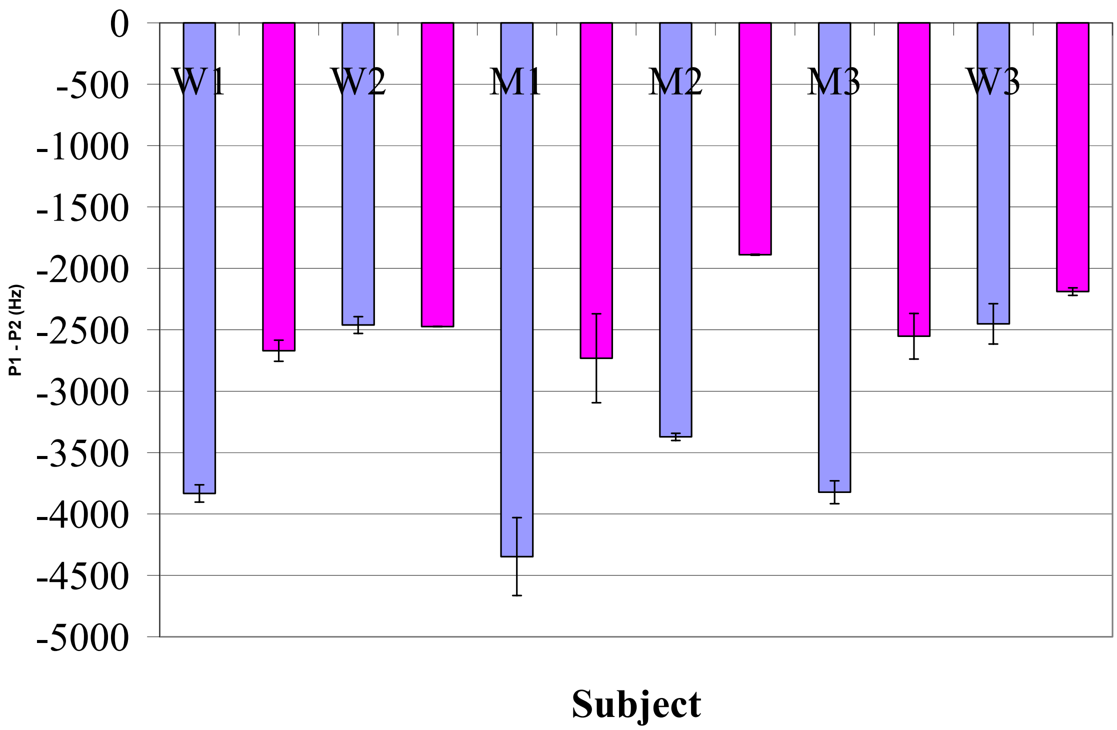

Figure 17. P1-P2 frequency values for the two palatal affricates in word-initial accented syllable onset position

An ANOVA with CONSONANT TYPE and SUBJECT as factors revealed significance for both factors at the level of p<.001. P-values associated with post-hoc comparisons using the adjustment method of Holm (1979) for a cross between SUBJECT and CONSONANT TYPE are provided in Table XII.

Subject	p-value
W1	p<.001
W2	NS
W3	NS
M1	p<.001
M2	p<.001
M3	p<.001

Table XII: P-values associated with post-hoc pair-wise comparisons for CONSONANT TYPE crossed with SUBJECT

Figure 18 provides centroid frequency results for the two affricates. As can be seen, centroid frequency values distinguish the two affricates for 3 out of 6 of the subjects (subjects W1, W2 and M3). Interestingly, this measure does not distinguish subjects M1, M3 and W3's productions, despite the fact that W3 categorically distinguishes the affricates on each of the two peaks individually. Given the small differences for each of the peaks, and the identity of the frequency differences between the two peaks, the centroid frequency values are quite similar.

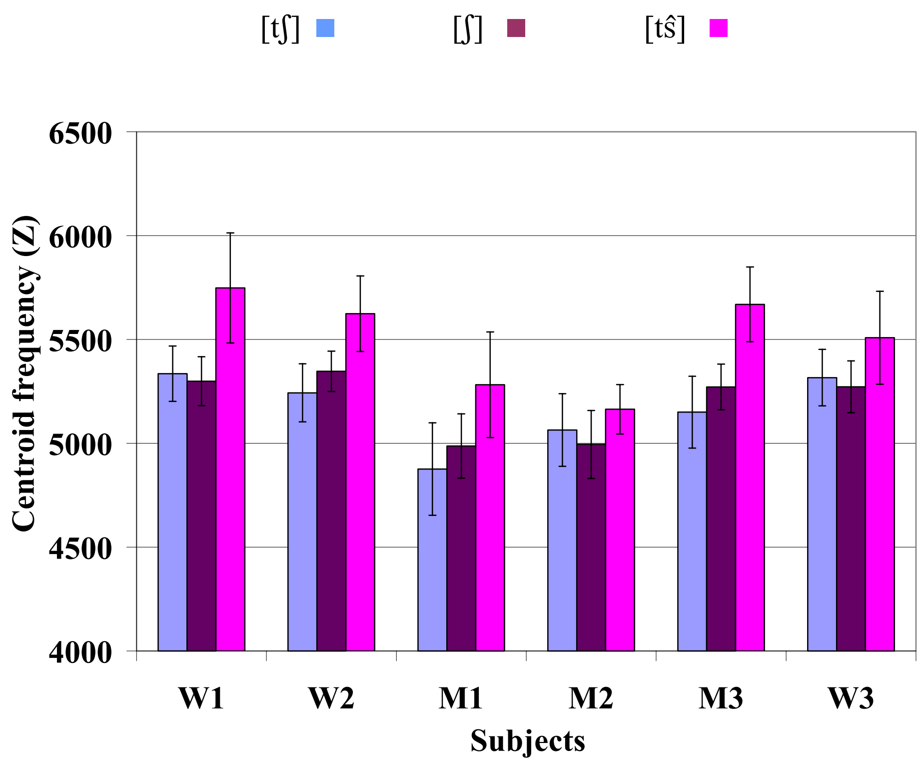

Figure 18: Centroid frequencies for six Serbian subjects in word-initial accented syllable onset position in the i#_[a,e,i,o,u] context

Centroid values for the two affricates were compared using pair-wise t-tests with AFFRICATE TYPE and SUBJECT as factors. Results revealed that AFFRICATE TYPE was significant for 5 out of 6 of the subjects, as shown in Table XIII. Comparisons between [tʃ] and [ʃ] were not significant for all subjects.

Subject	p-value affricate type
W1	p<.001
W2	p<.001
W3	NS
M1	p<.001
M2	p<.001
M3	p<.001

Table XIII: P-values associated with post-hoc pair-wise comparisons for the factor AFFRICATE TYPE

Figure 19 provides second formant frequency (F2) values in the vowel following the two types of affricates in word-initial accented syllable onset position, and 30 ms into the vowel. The vowel context shown here is the i#_e context. At the beginning of the vowel, four out of six subjects in this study display a consistent categorical difference in the second formant frequency. Two of the subjects display a high degree of variability in the F2 value following the [tʃ] affricates. Thirty ms into the vowel, five out of six subjects distinguish the two affricates based on second formant frequency values in the following vowel. Even the sixth subject (W1) displays a small difference, with extremely small standard deviations for each affricate type.

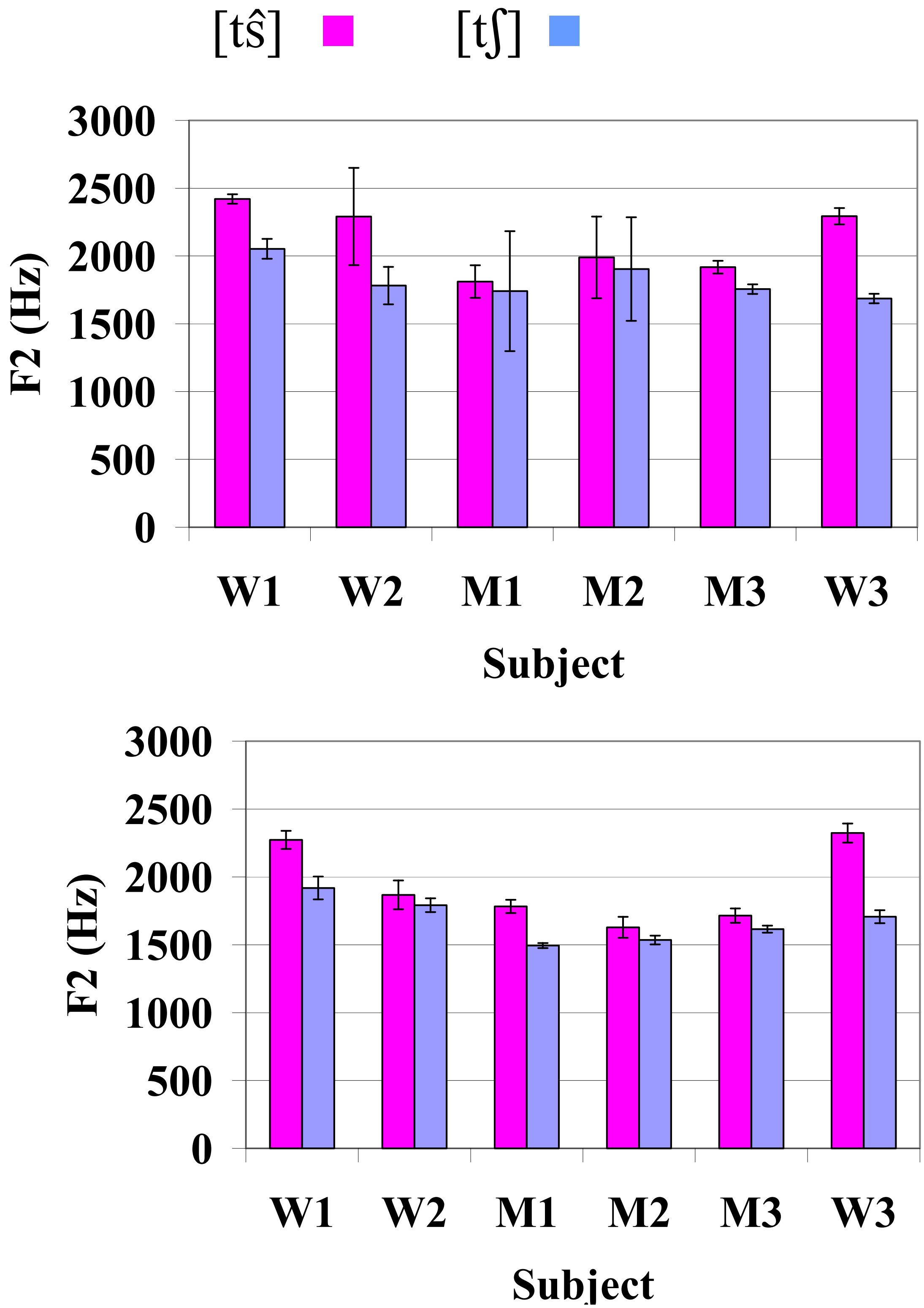

Figure 19: F2 values associated with Serbian affricates in word-initial accented syllable onset position (0 ms, top and 30 ms into the vowel, bottom)

The numbers of tokens are small here across all vowel conditions, so no statistical tests were undertaken in this position.

Figure 20 provides Peak 1 amplitude (A1) results plotted against Peak 2 amplitude (A2) results. As can be seen, [tʃ] has greater amplitude within the lower frequency range than [tŝ], although all speakers exhibit some overlap, and for some subjects (M2 and M3), the amplitude range employed in P1 is almost completely overlapping for the two affricates. The A1 values are very similar for [tʃ] and [ʃ] for subjects W2, M1 and M3. Results show that the amplitude of P2 is higher for [tʃ] than for [tŝ] in the upper frequency range, similar to the results found in the lower range of the spectrum. Results thus suggest that the fricated interval of [tʃ] and the fricative [ʃ] are louder than [tŝ] overall.

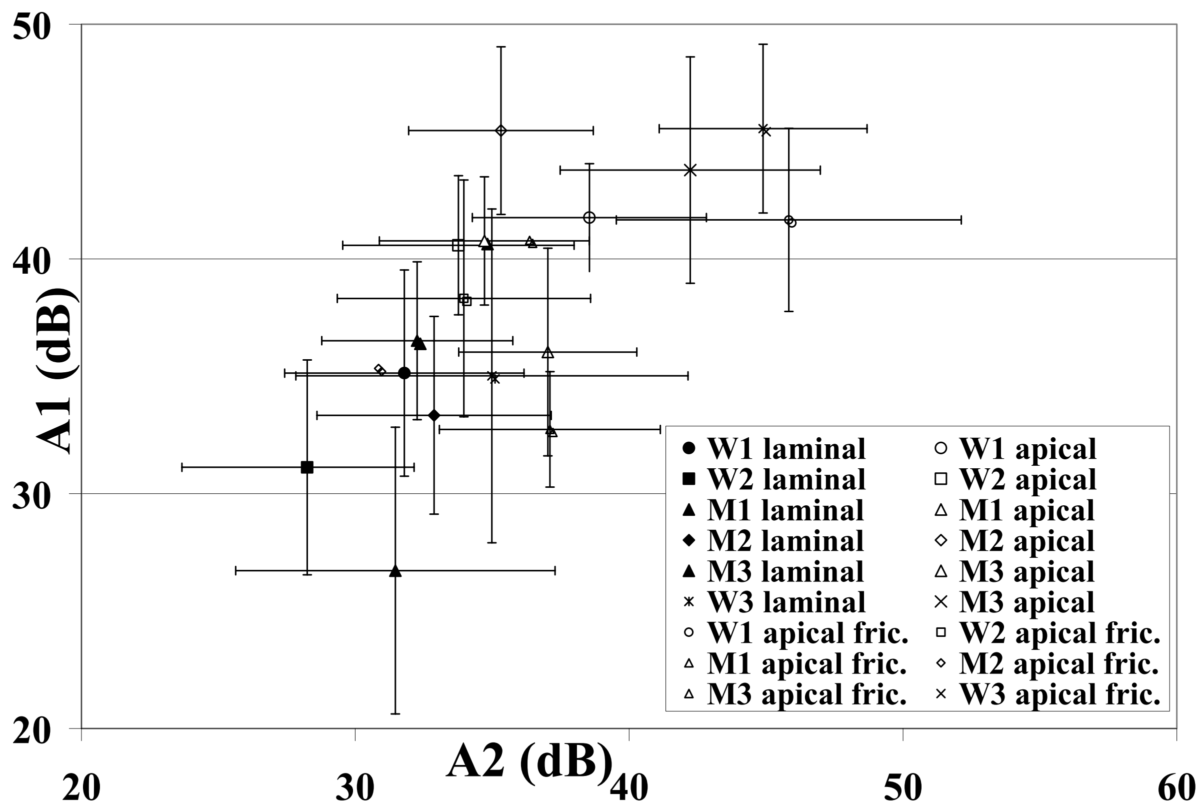

Figure 20: A1 by A2 for six Serbian speakers in word-initial accented syllable onset position

A1-A2 differences are small and inconsistent, as shown by the means and standard deviations provided in Table XIV. We have already seen that both affricates have a peak in the upper frequency range, as well as in the lower frequency range, which makes it difficult to analyze the affricates as having a stridency difference. However, if P1 were higher amplitude than P2 for one of the affricates (e.g. it had a higher A1 value), and the other affricate had a higher A2 value (e.g. the amplitude of P2 was higher than the amplitude of P1), this would present evidence for a stridency difference. That is, a positive A1-A2 value would correspond to a non-strident affricate, and a relatively small or negative A1-A2 value would correspond to a strident affricate, that has higher concentration of energy in the upper frequency range. As can be seen, individual subjects have different peaks being higher. Four out of the six subjects included in this experiment have a louder higher frequency peak (e.g. A2 is greater than A1) for [tʃ]. However, Subject M1 displays the opposite pattern, and Subject W1 does not show a consistently higher peak for either P1 or P2 for either affricate. Five out of the six subjects (all except Subject M1) display negative A1-A2 values for both affricates, suggesting that both affricates are strident.

Subject	[ts̑]	[tʃ]
W1	-3.34752 (-0.02182)	-3.21381 (1.975526)
W2	-2.88233 (-0.67182)	-6.81832 (1.257554)
W3	-0.03116 (0.052627)	-1.55496 (-0.07354)
M1	4.738042 (-0.27874)	0.99425 (-1.16935)
M2	-0.46433 (0.065094)	-10.1458 (-0.19692)
M3	-4.25872 (0.125563)	-6.05492 (1.101918)

Table XIV. A1-A2 values for the two affricates in word-initial accented syllable onset position

In the next section, we look at the results in medial unaccented syllable onset position.

4.3.2 Word-medial unaccented syllable onset position

Figure 21 displays P1 results plotted against P2 for all subjects in word-medial unaccented syllable onset position. As can be seen, the peaks are similarly distinguished in this position from those in word-initial accented syllable onset position for most of the subjects. Subject M1, however, seems to distinguish them less well in this prosodic context. Just as in word-initial position, the two affricates are categorically distinguished by the second peak frequency (P2) for all subjects, while the values for [tʃ] and [ʃ] are highly similar and overlapping.

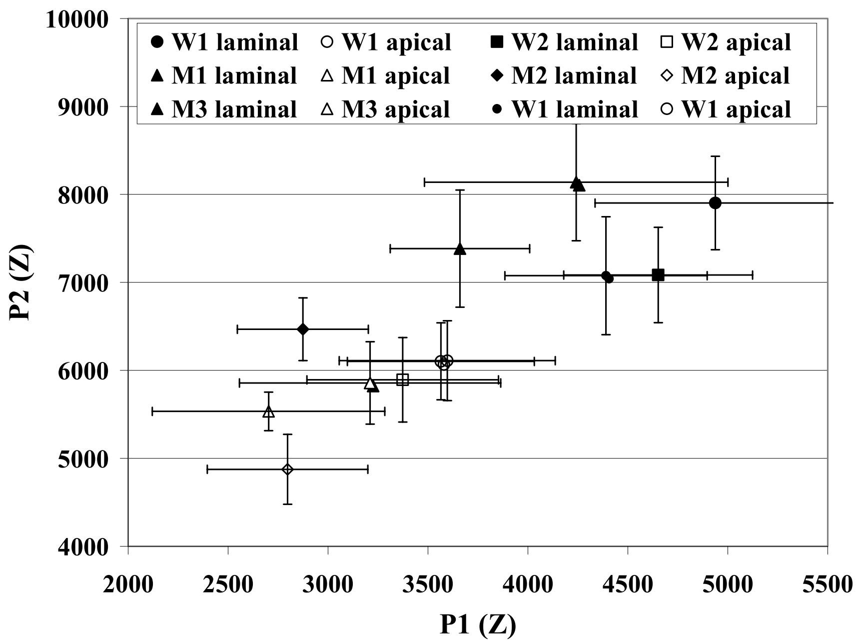

Figure 21. P1 plotted against P2 for six Serbian speakers in word-medial unaccented syllable onset position

Table XV provides p-values associated with post-hoc pair-wise comparisons using Bonferroni adjustments for the factor CONSONANT TYPE for both P1 and P2 values.

Subject	p-values P1 frequency			p-values P2 frequency		
W1		[tʃ]	[ʃ]		[tʃ]	[ʃ]
	[tŝ]	p<.001	p<.001	[tŝ]	p<.001	p<.001
	[ʃ]	NS	--	[ʃ]	p<.005	--
W2		[tʃ]	[ʃ]		[tʃ]	[ʃ]
	[tŝ]	p<.001	p<.001	[tŝ]	p<.001	p<.001
	[ʃ]	NS	--	[ʃ]	NS	--
W3		[tʃ]	[ʃ]		[tʃ]	[ʃ]
	[tŝ]	p<.001	p<.001	[tŝ]	p<.001	p<.001
	[ʃ]	NS	--	[ʃ]	NS	--
M1		[tʃ]	[ʃ]		[tʃ]	[ʃ]
	[tŝ]	p<.001	p<.001	[tŝ]	p<.001	p<.001
	[ʃ]	NS	--	[ʃ]	NS	--
M2		[tʃ]	[ʃ]		[tʃ]	[ʃ]
	[tŝ]	NS	NS	[tŝ]	p<.001	p<.001
	[ʃ]	NS	--	[ʃ]	NS	--
M3		[tʃ]	[ʃ]		[tʃ]	[ʃ]
	[tŝ]	p<.001	p<.001	[tŝ]	p<.001	p<.001
	[ʃ]	NS	--	[ʃ]	p<.001	--

Table XV: P-values associated with post-hoc pair-wise comparisons for the factor CONSONANT TYPE

Figure 22 provides values for a spectral peak found in the noise-bursts of [t] and [k] in the 2000-5000 Hz range. As in word-initial position, we again notice that the peak found in [t] is in a higher frequency range than the peak in [k]. However, the difference is smaller and the distributions are overlapping in this prosodic context for all subjects, due to the fact that consonants are more subject to coarticulatory effects in this context, and to the fact that the preceding vowel context is varied along with the following vowel context for this data set. Subject W2 in fact shows a difference in the opposite direction.

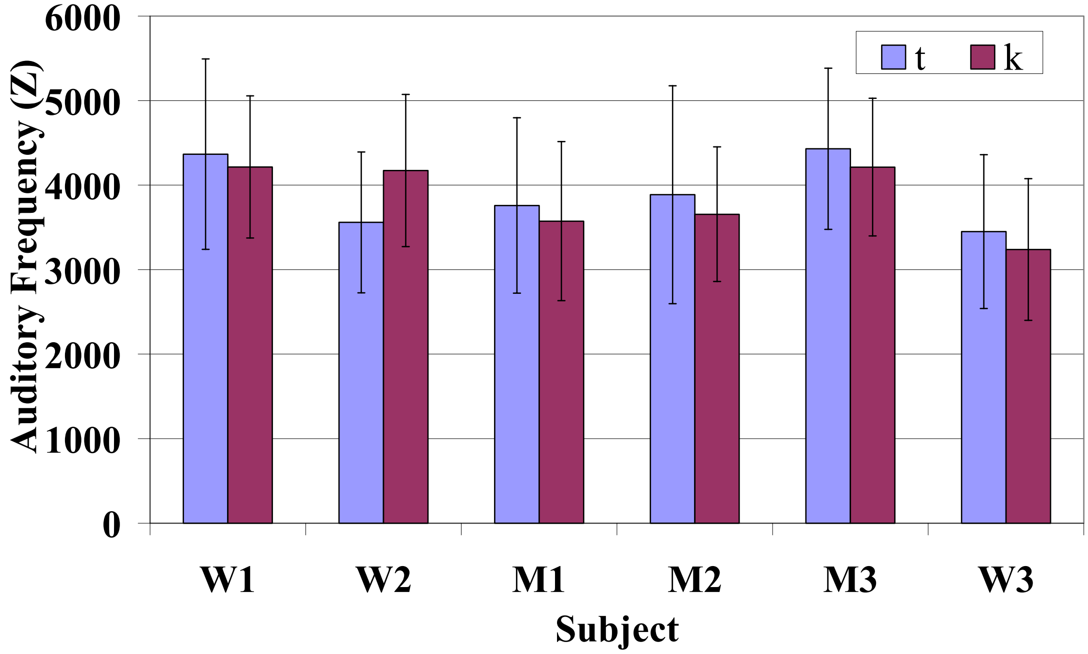

Figure 22. Frequency of a prominent spectral peak found in the noise-bursts of the stops [t] and [k] in word-medial unaccented syllable onset position

Figure 23 provides the P1-P2 values in word-medial position. Similar to the results found in word-initial position, the difference between the frequencies of the two peaks is very small for the two affricate types, except for Subject W3, whose productions display almost no difference. Subject W2's productions, which were not distinct in word-initial accented syllable onset position, display opposite differences in medial position.

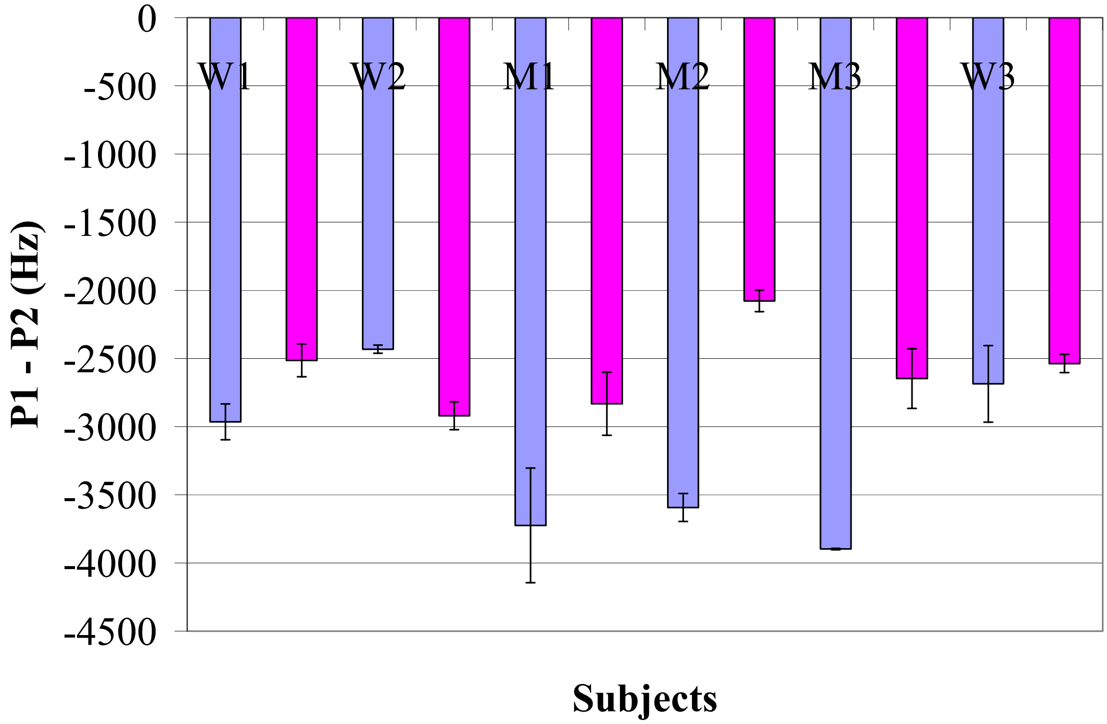

Figure 23: P1-P2 frequency values in word-medial unaccented syllable onset position

The centroid frequency results are quite similar for most of the speakers in word-initial and word-medial positions, as shown in Figure 24. However, subject M2, whose productions displayed considerable overlap for the two affricates in the word-initial accented syllable onset position, produces a marked contrast with little overlap for the two affricates in word-medial position. This leaves only subject W3 whose productions display little distinction between the two affricates on this measure.

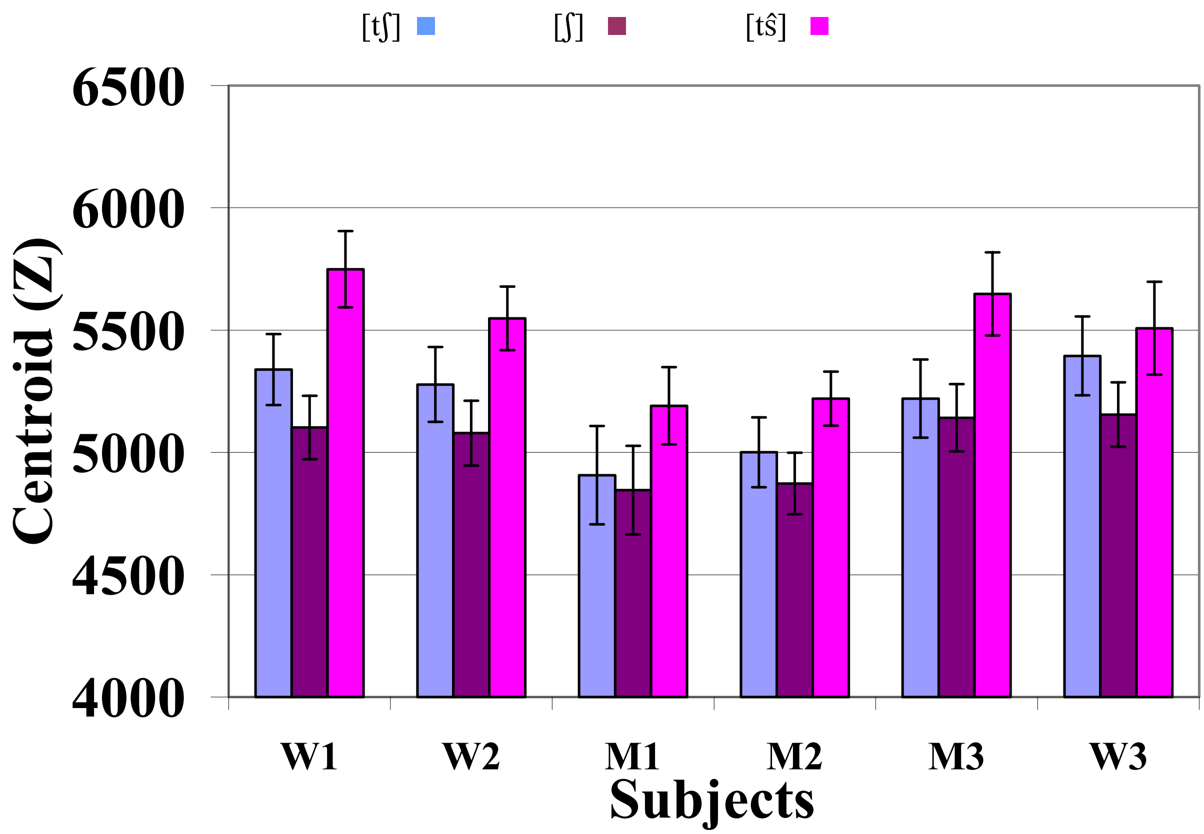

Figure 24: Centroid frequency for six Serbian speakers in word-medial unaccented syllable onset position in the i#_[a,e,i,o,u] context

A two-factor ANOVA with SUBJECT and CONSONANT TYPE as factors revealed that both factors were significant for all subjects, as well as an interaction between the two factors. P-values associated with post-hoc pair-wise comparisons using Holm's (1979) adjustment method are provided in Table XVI.

Subject	P-values CONSONANT TYPE		
W1		[tʃ]	[ʃ]
	[tŝ]	p<.001	p<.005
	[ʃ]	NS	--
W2		[tʃ]	[ʃ]
	[tŝ]	p<.001	p<.005
	[ʃ]	NS	--
W3		[tʃ]	[ʃ]
	[tŝ]	NS	NS
	[ʃ]	NS	--
M1		[tʃ]	[ʃ]
	[tŝ]	p<.001	p<.001
	[ʃ]	NS	--
M2		[tʃ]	[ʃ]
	[tŝ]	p<.001	p<.001
	[ʃ]	NS	--
M3		[tʃ]	[ʃ]
	[tŝ]	p<.001	p<.001
	[ʃ]	p<.001	--

Table XVI: P-values associated with post-hoc pair-wise comparisons of CONSONANT TYPE

In word-medial position, F2 values distinguish the two affricates less well than they did in word-initial position. As seen in Figure 25, the differences are very small. We attribute this to the more centralized formant frequency targets found in the prosodically weak (unaccented) second syllable position.

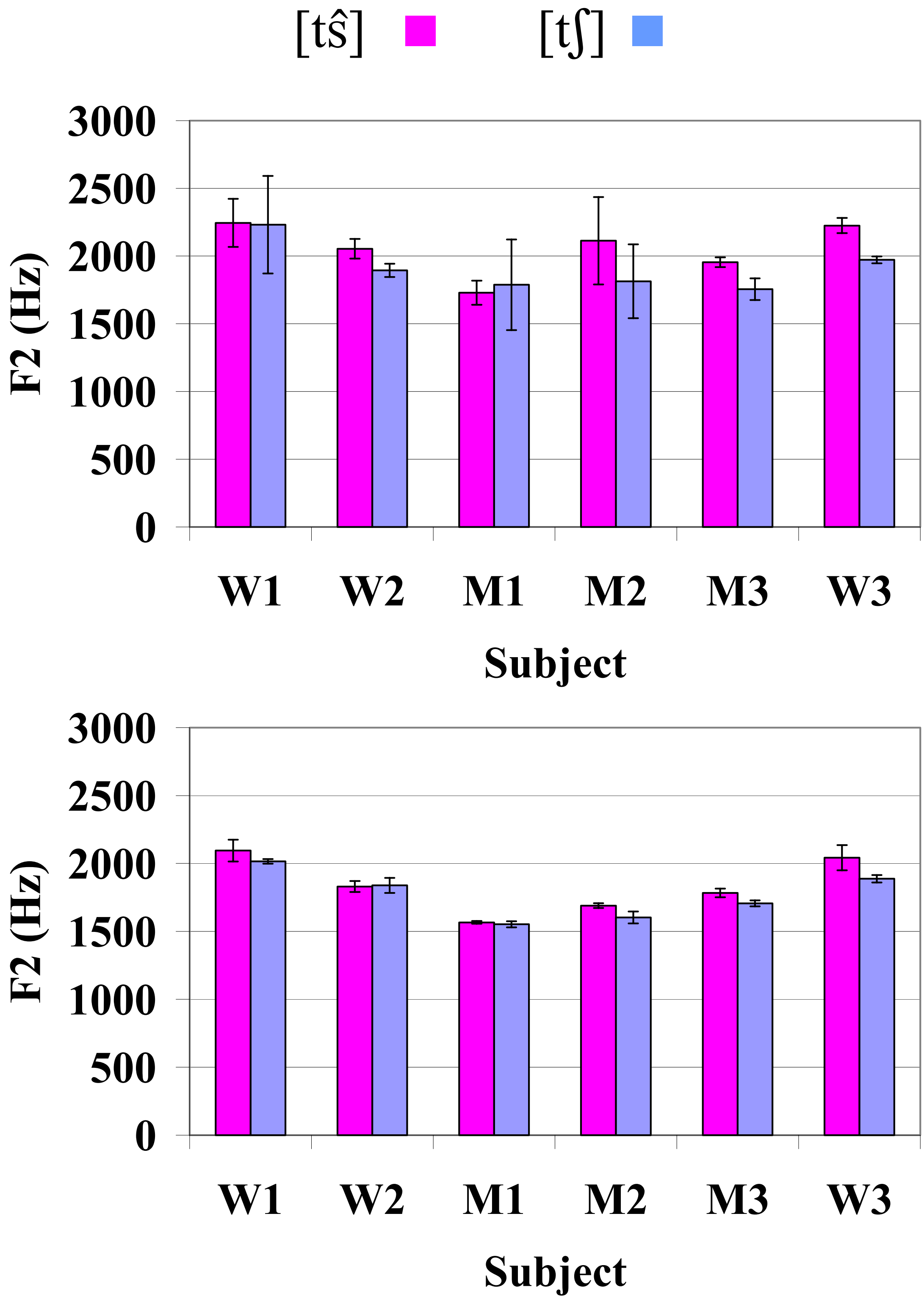

Figure 25: F2 values for six Serbian speakers in word-medial unaccented syllable onset position (begin of vowel, top; and 30 ms into the vowel, bottom)

The amplitudes of peaks 1 and 2, abbreviated as A1 and A2, are provided in Figure 26. Just as in word-initial accented syllable onset position, in the word-medial unaccented syllable onset position, there is considerable overlap in the amplitude values for the first prominent spectral peak, but little overlap for the amplitude of the second prominent spectral peak.

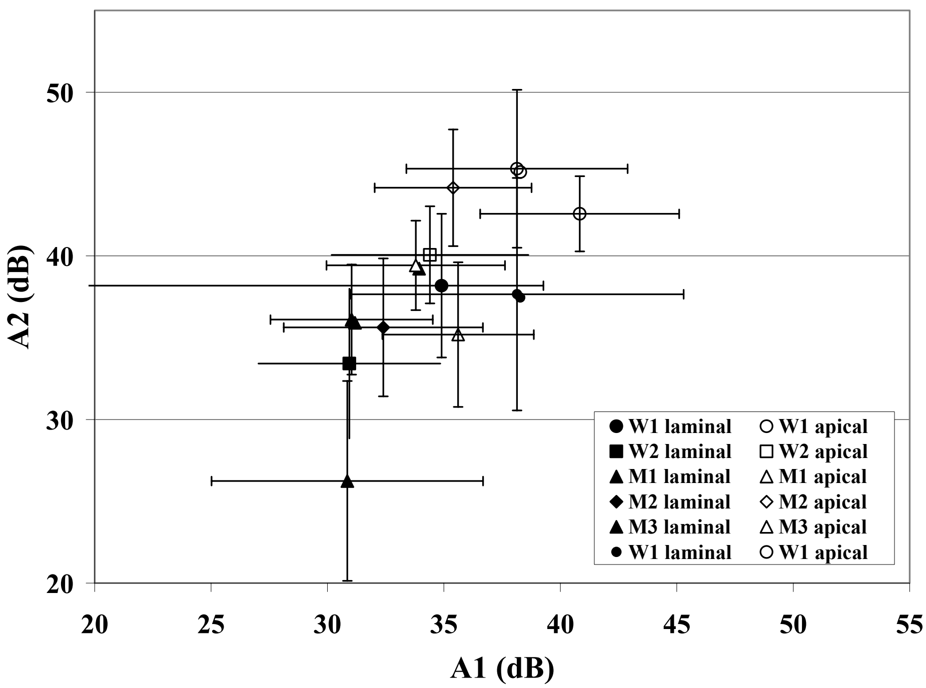

Figure 26: Peak 1 amplitude for six Serbian speakers in word-medial unaccented syllable onset position

Table XVII displays the A1-A2 values for the two affricates. Just as in initial position, there is no consistent difference in the height of the two peaks for the two affricates across subjects.

Subject	[tŝ]	[tʃ]
W1	-3.28219 (-0.05933)	-1.73887 (0.667801)
W2	-2.4822 (0.035404)	-5.66264 (0.864406)
W3	0.479328 (-0.27865)	-3.24954 (-0.14304)
M1	4.604345 (-0.01654)	0.418234 (-0.29371)
M2	-3.2317 (1.045878)	-8.76546 (-0.29145)
M3	-5.07501 (0.294288)	-5.63064 (-0.55244)

Table XVII: A1-A2 values for the two affricates in word-medial unaccented syllable onset position

4.4 Discussion

The results of Experiment 2 show that there are two distinct salient spectral peaks found for the two affricates under investigation, and that the fricative [ʃ] displays similar P1 and P2 values to the apical affricate [tʃ]. While the peak found in the lower part of the spectrum (P1) is categorically distinct for four out of six of the subjects in word-initial accented syllable position, and categorically distinct for five out of the six speakers in word-medial unaccented syllable onset position. The second peak located in the upper frequency range of the spectrum (P2) is categorically distinct for all speakers in both prosodic contexts studied. The P2 difference is consistent with our hypothesis that a host of articulatory attributes cumulatively result in front cavity volume differences in the two affricates.

The centroid frequency captures the overall lower frequency range employed for [tʃ] than for [tŝ] by four out of six speakers in word-initial accented syllable onset initial position, and five out of the six speakers in word-medial unaccented syllable position. However, the similarity of the two peaks, and the similar frequency differences found between the two spectral peaks result in a lack of difference for two speakers in word-initial accented syllable onset position, and for a single speaker in word-medial unaccented syllable onset position.

An acoustically-based feature often applied to fricatives is [compact] vs. [diffuse] (Jakobson, Fant and Halle, 1961). As shown by the measure of P1-P2, the frequency range between the two peaks is similar across the two affricates. Thus, the affricates are not distinguished in the range of noise they exhibit, and the phonological feature [compact] vs. [diffuse] would not be available to capture the differences between them.

The amplitude of the first peak (A1) overlaps quite substantially for the two affricates, while the amplitude of the second peak (A2) categorically distinguishes the two affricates for four out of six of the speakers in word-initial accented syllable onset position, but only for two of the subjects in word-medial unaccented syllable onset position.

While the two affricates display two similar frequency peaks, it might have still been possible to deem the difference one of stridency, if the lower peak had more energy than the higher peak for one of the affricates. However, there does not seem to be any consistent difference in amplitude of the two peaks, as shown by the A1-A2 values. F2 values in the following vowel, however, are distinct for all four speakers, but the ranges are overlapping.

In sum, the most reliable cue for distinguishing the two affricates is P2, the frequency of a prominent high frequency spectral peak, as it differentiates the two affricates consistently for all subjects in this study in all prosodic contexts. The apical affricate [tʃ] displays lower frequencies for each of the two peaks than does the laminal affricate [tŝ]. The stop noise bursts of [t] and [k] display a comparable difference in the frequency of a lower frequency spectral peak. Notice that the range of energy differs for the affricates and the stops, given the difference in constriction degree between frication in an affricate and a burst in a stop. This is because the resonances of the vocal tract in frication noise are realized while constriction is still in place, while stop resonances are acoustically realized only upon release of the constriction, during the opening phase. Thus, the frequency range employed in the frication interval of affricates is much higher due to fact that only the front cavity volume contributes to the overall resonances of the vocal tract. In stop bursts, which are acoustically realized in the opening phase, the resonances of both cavities are employed, leading to a lower frequency range.

In the next section, we turn to the phonological patterning of the two affricates, and show that the frequency of the second spectral peak in affricates, which corresponds to a lower frequency peak in stop noise-bursts, allows us to understand the phonetic similarity of [tŝ] and [t] on the one hand, and [tʃ] and [k] on the other hand. We propose a phonological feature based on this acoustic and physiological similarity that accounts for the phonological patterns seen.

5. Phonological patterning and analysis

The two palatal affricate classes participate in phonological alternations, most notably, the process of iotization, whereby non-palatal consonants become palatal in the environment of several suffixes, as exemplified by the passive and comparative markers shown in (2) and (3) respectively. Iotization is triggered by a [coronal, - anterior] feature of the front vowel in the passive suffix in (2), and by a covert [coronal, - anterior] feature value in (3). A non-palatal consonant becomes palatal by adopting the place specification of the vowel.

(2)

	Past	Passive	English Gloss
a.	ukro[t]io	ukro[tŝ]en	‘tame’
b.	re[k]ao	re[tʃ]en	‘say’
c.	bra[n]io	bra[ɲ]en	‘defend’
d.	ko[s]io	ko[ʃ]en	‘mow’

(3)

	Adjective	English Gloss	Masculine Comparative	Feminine Comparative	English Gloss
a.	žu[t]	‘yellow’	žu[tŝ]i	žu[tŝ]a	‘more yellow’
b.	ja[k]	‘strong’	ja[tʃ]i	ja[tʃ]a	‘stronger’
c.	cr[n]	‘black’	cr[ɲ]i	cr[ɲ]a	‘more black’
d.	vi[s]ok	‘tall’	vi[ʃ]i	vi[ʃ]a	‘taller’
e.	br[z]	‘fast’	br[ʒ]i	br[ʒ]a	‘faster’

We will remain non-committal as to how this front vowel feature is manifested in alternating palatal consonants. It is crucial for us that, through the process of iotization, the apical coronal [t] alternates with the laminal affricate [tŝ], as in (2a) and (3a), while the dorsal [k] alternates with the apical affricate [tʃ], as in (2b) and (3b). The phonetic differences between the two Serbian affricate classes established by our experimental results provide an explanation of the [t]/[tŝ], and [k]/[tʃ] alternating pattern. Results of Experiment 2 in particular show that the [t]/[tŝ] and [k]/[tʃ] alternations can be characterized in terms of the same acoustic cues that sustain the ć [tŝ]/č [tʃ] contrast. It has been shown that the apically articulated [tʃ] has a lower frequency range employed than the laminally articulated [tŝ], and this is mirrored by relative frequency differences found in the stop bursts of [k] and [t] respectively. That is, the frication noise of the apical affricate [tʃ] has a lower frequency resonance than [tŝ], just like [k] which has a lower frequency stop burst resonance than [t]. The frequency differences employed are achieved mainly through place of articulation differences in the stops, but through a much more complex combination of gestures involving two different articulators: the lips and the tongue, for the affricates. This, we propose, gives rise to a physiologically-based phonetic difference in the volume of the front cavity, which is larger for [tʃ] than for [tŝ]. While crucially contributing to the phonetic differentiation of the two palatal affricates, the volume of the front cavity also ensures a proper characterization of the process of iotization.

With this in mind, we posit a physiologically-based phonological feature [front cavity volume], with values [increased] and [decreased], in order to capture the phonological significance of this physiological distinction. The feature [front cavity volume] elucidates the phonological contrast among sounds that employ lower, and those that employ higher, resonance frequencies, and the auditory basis of this contrast. Miletić (1960) in fact notes that the acoustic impression of [tŝ] is higher than that of [tʃ]. This is indeed what we find in our quantitative study. Moreover, the feature [front cavity volume] provides the basis for the phonological alternation of iotization. Both [t] and laminal palatals such as [tŝ], which are characterized by a relatively smaller front cavity volume, have the value [decreased] for this feature, while [k] and apical palatals such as [tʃ], characterized by a relatively larger volume, have the value [increased]. Generally, the value of the feature [front cavity volume] remains unaltered through the process of iotization, and non-palatals alternate with those palatal sounds with which they share the marking for [front cavity volume].

The feature [front cavity volume] serves a broader classificatory role within the system of Serbian consonants. In what follows, we focus on how this feature accounts for the alternating pairs in the process of iotization, listed in Table XVIII.

		PALATALS	
		Laminal Palato-alveolar	Apical Alveolo-palatal
Dentals	t	tŝ	
	d	dẑ	
	ts		tʃ
	s		ʃ
	z		ʒ
	n	ɲ	
	l	ʎ	
Velars	k		tʃ
	g		ʒ
	x		tʃ

Table XVIII: Alternations triggered by iotization [5]

The fact that [t] alternates with [tŝ], and [k] with [tʃ] strongly suggests that, in the overall pattern of iotization, dentals alternate with laminal palatals, and velars alternate with apical palatals. Thus, as already noted, pairs of alternating sounds share the value for [front cavity volume], which corresponds to [decreased] for dentals and laminal palatals, and to [increased] for velars and apical palatals. This is fully supported by the dental stop [d], and the dental sonorants [n] and [l], which alternate with their counterparts in the class of laminal palatals. This is also supported by the velar segments [g] and [x], which alternate with their counterparts in the class of apical palatals.

However, there are several departures from this patterning. This is because iotization is a lexical phonological process, as proposed in Lexical Phonology and Morphology (Kiparsky 1982), and as such, is bounded by the principle of structure preservation. Lexical phonological processes are structure preserving in the sense that they do not create "new" segments. Only those segments that belong to the underlying

[5] [r] is not included into Table II because it has no palatal counterpart. So, this sound does not participate in iotization, as shown by the forms *udariti* 'hit' vs. *udaren*, the passive form of 'hit'.

segment inventory, which in turn sets the range of possible lexical contrasts, may participate in lexical alternations. The initial inventory for Serbian lexical processes is that given in Table I, and iotization takes effect within the range of this inventory. The dental fricatives [s] and [z] have no counterparts in the class of laminal palatals, as shown by the gap in Table I, and therefore pattern with the closest alternative, which in this case are the apical palatal fricatives, [ʃ] and [ʒ]. Another case of structure preservation is the [g]/[ʒ] alternation. While [dʒ] would be the natural counterpart for [g], [dʒ] entered the language via borrowing, and never became a participant in lexical alternations. Due to this, [g] alternates with its closest alternative counterpart, [ʒ], which is not a stop but is nonetheless an apical palatal sound with the same [front cavity volume] specification. The status of the [ts]/[tʃ] alternation remains something of a puzzle, as the expected alternating pair in this case should have been [ts]/[tŝ], since these sounds share the value [decreased] of the [front cavity volume] feature. While we have no explanation for this alternation, we could at least mention that, historically, *c* [ts] entered the language through an alternation with [k], and may have idiosyncratically inherited this sound's alternating pattern (Ivić 1967, 1968).

The iotization pattern exhibited by labials brings in further complexities. Iotization in labials would either need to be manifested as secondary articulation, or as a separate segment.[6] In this case, iotization is manifested as the segment *lj* [ʎ] immediately following the alternating segment. Thus, the stem *ljub* [ʎub] (inf. *ljubiti* 'kiss') has the passive form *ljubljen* [ʎubʎen], and the stem *kup* (inf. *kupiti* 'buy') has the passive form *kupljen* [kupʎen]. Labials with secondary articulations are excluded by structure preservation. There are no consonants with secondary articulations in the inventory of Serbian consonants (shown in Table I), and so neither iotization, nor any other lexical phonological process, can create palatalized consonants.

The next issue is: why is the iotization on labials realized as *lj* [ʎ]? The resulting segment should be a sonorant rather than an obstruent, for reasons of optimal sonority

[6] The third possiblity, a palatal labial resulting from iotization, is excluded on acoustic grounds. If the lip and tongue tip gestures coincided temporally, there would be no acoustic effect, since only the effect of the most forward constriction will be acoustically relevant in an oral stop. The end result would be that iotization would not be recoverable auditorily.

sequencing in prevocalic syllable onset position, and the choice of *lj* [ʎ] is consistent with this.[7] Moreover, it may well be that [ʎ], with the value [decreased] for [front cavity volume], shares this value with labials. Since labials have no front cavity volume whatsoever, they may well have the value [decreased] for this feature, and as such, pair with laminal palatals in iotization.

In conclusion, while the phonological process of iotization has its complexities, as is generally the case with lexical phonological processes, it elucidates the relation of the two palatal affricates with other sounds in the consonantal inventory, and specifically points at the phonological closeness of [t] and [tŝ], and of [k] and [tʃ]. It further points at the relevance of the phonological feature [front cavity volume] which plays a contrasting as well as classificatory role in the Serbian consonantal inventory, and provides a basis for phonological patterning.

References

Catford, J.C. (MS) *The phonetics of Caucasion languages*, University of Michigan-Ann Arbor.

Dart, S. (1998) Comparing French and English coronal consonant articulation, *Journal of Phonetics* **26**, 71-94.

Hála, B. (1962) Uvedení do fonetiky češtiny: Na obecně fonetickém základě, Prague, Nakladatelství československé Akademie Věd.

Halle, M. and Stevens, K.N. (1971) The postalveolar fricatives of Polish, In *Speech production and language: in honor of Osamu Fujimura* (Kiritani, S., Hirose, H. and Fujisaki, H., editors), Berlin, Mouton de Gruyter.

Hermansky, H. (1990) Perceptual linear predictive (PLP) analysis of speech, *Journal of the Acoustical Society of America* **87** (4), 1738-1752 (April 1990).

[7] While *lj* [ʎ] is not the only possibility, there are few others. The prime candidate would of course be [j]. However, clusters with [j] as the second member are either completely absent, as in the so-called Ekavian dialects, or highly restricted, as in the Jekavian dialects (see Ivić 1958). In either case, lexical phonological processes may not produce clusters with [j] as the second member. With [j] excluded, the choice remains between [ʎ] and [ɲ], and it may well be that the former is a more natural replacement for [j] than the latter. It is interesting that the laminal affricates have a high degree of lateral tongue contact, and thus from a phonetic point of view, it is not at all surprising that this lateral contact is interpreted as a lateral sonorant consonant in absence of another place of articulation feature.

Holm, S. (1979). A simple sequentially rejective multiple test procedure, *Scandinavian Journal of Statistics*, **6**, 65-70.

Howell, P. and S. Rosen (1983) Production and perception of rise time in the voiceless affricate/fricative distinction, *JASA.* **73** (3), 976-984.

Ivić, P. (1958) *Die serbokroatischen Dialekte. Ihre Struktur und Entwicklung*. Erster Band. Allgemeines und die stokavischen Dialektgruppe. The Hague: Mouton.

Ivić, P. (1967) Milan Surdučki, The Distribution of Serbo-Croatian Consonants, *Zbornik Matice srpske za filologiju i lingvistiku* **10**, 227-236.

Ivić, P. (1968) Razvoj principa distribucije fonema u srpskohrvatskom jeziku, *Književnost i jezik* **15**, 13-32.

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

Johnson, K. (1997) *Acoustic and Auditory Phonetics*, Oxford, Blackwell.

Keating, P. (1988) Palatals as complex segments: X-ray evidence, *UCLA WPP* **69**, 77-91.

Kiparsky, P. (1982) Lexical phonology and morphology. In *Linguistics in the Morning Calm*, vol 2, (I.S.Yang, editor), Seoul: Hanshin, 3-91.

Ladefoged, P. and Maddieson, I. (1996) *The Sounds of the World's Languages*. Oxford, Blackwell.

Maddieson, I. (1997) Phonetic Universals, In *The Handbook of Phonetic Sciences* (Hardcastle, W. and J. Laver, editors), Oxford, Blackwell.

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

Miletić, B. (1960) *Osnovi fonetike srpskog jezika*, Beograd, Naučna knjiga.

Miller-Ockhuizen, A. and Zec, D. (2003) Acoustics of contrastive palatal affricates predict phonological patterning, *Proceedings of ICPhS,* Barcelona.

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

Ohnesorg, K. and Svarny, O. (1955) *Études experimentales des articulations chinoises*, Prague, Rozpravy Ceskosloveské Akademie Věd, Volume **65**, 5.

Perkell, J., Boyce, S. and Stevens, K. N. (1979) Articulatory and Acoustic correlates of the [s-š] distinction, *Speech Communication Papers presented at the 97th meeting of the Acoustical Society of America (J. Wolff and D. H. Klatt, editors)*, 109-113.

Recasens, D. (1990) The articulatory characteristics of palatal consonants, *Journal of Phonetics* **18**, 267-280.

Sekey, A. and Hanson, B. A. (1994) Improved 1-Bark bandwidth auditory filter, *Journal of the Acoustical Society of America* **75** (6), 1902-1904 (June 1984).

Wang, S., Sekey, A. and Gersho, A. (1992) An objective measure for predicting subjective quality of speech coders, *IEEE Journal on Selected Areas in Communications*, SAC-**10** (5), 819-829 (June 1992).

Weigelt, L.F., Sadoff, S. J. and Miller, J.D. (1990) Plosive/fricative distinction: the voiceless case, *Journal of the Acoustical Society of America* **8** (6), 2729-2737.

Appendix

1. ćela (N,Nom,Sg)
2. kače (V,Pres,3,Pl)
3. meće (V,Pres,3,Pl)
4. meče (N,Nom,Sg)
5. kuće (N,Nom,Pl)
6. kesa (N,Nom,Sg)
7. ćopa (V,Pres,3,Sg)
8. vičan (N,Gen,Sg)
9. Kaću (N,Acc,Sg)
10. ćuka (N,Gen,Sg)
11. laka (A,Nom,Sg)
12. ćifta (N,Nom,Sg)
13. toče (V,Pres,3,Pl)
14. biće (V,Fut,3,Sg)
15. kuče (N,Nom,Sg)
16. kopa (V,3,Sg)
17. ćasa (N,Nom,Sg)
18. miče (V,Pres,3,Sg)
19. voće (N,Nom,Sg)
20. čela (N,Gen,Sg)
21. tuča (N,Nom,Sg)
22. kapa (N,Nom,Sg)
23. šefa (N,Gen,Sg)
24. časa (N,Gen,Sg)
25. veće (A,Neut,Nom,Sg)
26. jaču (A,Acc,Fem,Sg)
27. laku (A,Acc,Fem,Sg)
28. seči (N,Ins,Sg)
29. tati (N,Dat,Sg)
30. jače (A,Nom,Neut,Sg)
31. kašu (N,Acc,Sg)
32. šina (N,Nom,Sg)
33. kupa (N,Nom,Sg)
34. teža (A,Fem,Nom,Sg)
35. Kaće (N,Gen,Sg)
36. kaši (N,Dat,Sg)
37. seču (N,Acc,Sg)
38. kaše (N,Gen,Sg)
39. laki (A,Nom,Pl)
40. tate (N,Gen,Sg)
41. veću (A,Fem,Acc,Sg)
42. kašom (N,Ins,Sg)
43. Kaće (N,Gen,Sg)
44. tatu (N,Acc,Sg)
45. kiša (N,Nom,Sg)
46. veći (A,Masc,Nom,Sg)
47. kaša (N,Nom,Sg)
48. većom (A,Fem,Inst,Sg)
49. seča (N,Nom,Sg)
50. Kaćom (N,Ins,Sg)
51. lake (A,Fem,Nom,Pl)
52. jača (A,Nom,Fem,Sg)
53. seče (N,Gen,Sg)
54. sečom (N,Ins,Sg)
55. jačom (A,Ins,Fem,Sg)
56. čila (A,Fem,Nom,Sg)
57. lakom (A,Fem,Ins,Sg)
58. šoka (N,Gem,Sg)
59. čuka (N,Nom,Sg)
60. sićan (A,Masc,Nom,Sg)

61. šuma (N,Nom,Sg)
62. topa (N,Gen,Sg)
63. jači (A,Nom,Masc,Sg)
64. tata (N,Nom,Sg)
65. tata (N,Nom,Sg)
66. tiša (A,Fem,Nom,Sg)
67. Kaća (N,Nom,Sg)
68. kuća (N,Nom,Sg)
69. čoka (N,Nom,Sg)
70. seća (V,Pres,3,Sg)
71. šaka (N,Nom,Sg)
72. voća (N,Gen,Sg)
73. mača (N,Gen,Sg)
74. luča (N,Nom,Sg)
75. seča (N,Gen,Sg)
76. Goča (N,Gen,Sg)
77. Kaći (N,Dat,Sg)
78. tatom (N,Ins,Sg)
79. veća (A,Fem,Nom,Sg)