Original research article
UDC 811.111'354:811.163.41'35
DOI 10.21618/fil2021095p
COBISS.RS-ID 129049601

Borivoje T. Petrović[1]
University of Novi Sad
Faculty of Education, Sombor
Language and Literature Department

SPECTRAL MOMENTS OF FRICATIVE CONSONANTS IN SERBIAN (AN ACCOUNT OF FEMALE SPEAKERS' PRODUCTION)

Abstract: *This paper deals with the measurement of spectral moments of Serbian fricatives. The four spectral moments: centre of gravity, spectral variance, skewness, and kurtosis were measured at the onset of the fricative, on the stable-state part and at the fricative offset. The results show that the highest values for the centre of gravity can be found in non-sibilants, except for the sound [v]. The measurements were taken in both Hertz and Bark scales, and the results show that the Bark scale serves better in distinguishing fricative types. The energy which is generated during the fricatives' production is not evenly allocated in the time-course of the fricative length. Voiced and voiceless fricatives differ in force of articulation within the fricative duration. The greatest factor effects were found to be the place of articulation and phonological voicing. The vowel that follows does not statistically influence the spectral moments' values, except for the difference in fricative [x] which surfaces as an acoustically different segment before the vowels [i] and [u].*

Keywords: *centre of gravity, spectral variance, skewness, kurtosis, spectral moments, fricatives, Serbian.*

Introduction

There exist many phonetic cues that can be used to distinguish between various phonemes and their features. Plenty of research deals with temporal properties of sounds. Hence, it is a fact that the duration of a vowel before an obstruent may signal

[1] borivoje@gmail.com

whether the obstruent is voiced or voiceless (Ladefoged and Disner 2012). Spectral cues, such as formant transitions, may differentiate between phonetic features of voice or place of articulation (House and Fairbanks 1953, among others). Rare progressive voicing assimilation in Serbian on the boundary between two plosives occurs if the word-final plosive is realised without occlusion (Batas 2014).

In this paper, we have set out to explore the acoustic properties of fricative consonants, i.e., their spectral properties. This paper aims to examine spectral moments that may signal the difference between some phonological contrasts, namely the features of voicing and place of articulation, as well as some non-linguistic factors such as the difference between speakers.

Fricative sounds are generated by a turbulent air stream that comes to constriction in the oral cavity which serves as the main resonator for fricatives' articulation. Fricatives in languages are of two kinds, sibilant, and non-sibilant. Sibilant fricatives feature the main constriction after which the air stream comes upon another obstacle (generally the teeth) on their way out (Ladefoged and Maddieson 1996). The non-sibilants are produced with the airflow passing through the main constriction only. During the production of sibilant fricatives, it is required that the constriction be made as precise as possible to allow the turbulent flow of air pressure, but wide enough not to articulate a closure (Iskarous et al. 2011).

The aerodynamic difference in production of fricatives influences the overall spectral picture of sounds. Flat spectra and low energy characterise front cavities, the mid fricatives have high intensity and spiky spectra, while the back fricatives have the medium intensity and a visible formant structure. (Tabain 1998: 100).

The values of spectral moments are numeric representations of parameters of inherent properties of the spectra. They are computed as the random distribution of values (Forrest et al. 1988) shows the mean frequency of the spectral slice (the first spectral moment or centre of gravity, [COG]), divergence of the values from the mean (the second spectral moment, spectral variance, or standard deviation of the spectrum), (as)symmetry of the distribution of the random values (the spectral skewness or spectral tilt) and the peakedness of a distribution (the spectral kurtosis).

The first spectral moment (M1, spectral mean, the centre of gravity or centroid) is a measure which shows the arithmetic mean of all frequency values that exist in a particular window and is calculated as:

$$\mu = \frac{1}{N} \times \sum_{i=1}^{N} x_i$$

where μ is the spectral mean, N – the number of randomly distributed values x added and divided by their number.

It is a fact that dental (alveolar) fricatives exhibit higher spectral energy than palatals (Haley 2002; Maniwa et al. 2009, among the others). Jongman et al. (2000) have calculated spectral means for the fricatives of American English. The [s, z] had the highest mean frequency (6133 Hz). On the other side, there were the palatals [ʃ, ʒ] (4229 Hz). The labiodental [f, v] had a spectral mean of 5108 Hz and interdental [θ, ð] 5137 Hz. Results for Serbian homologs will be given in the figures in the Results section.

Among sibilants, dental fricatives have the highest centre of gravity. Such behaviour is due to the present obstacle – the lower teeth – some 3 cm apart from the noise origin. The cavity that is formed is smaller, the air stream is precisely directed towards the teeth, and the channel during articulation is narrower (Haley et al. 2010). The non-sibilants do not have such obstacles and therefore consist of larger cavity sizes which absorbs the acoustic energy (except for the labials) (Tabain 2001). This state of facts tells that posterior places of articulation among sibilants should have lower COG values. We have mentioned that the spectral shape of non-sibilants is flatter than that of non-sibilants. The energy is more spread along the frequency range, and the centre of gravity has somewhat higher values (somewhere around 10 kHz).

Voiced fricatives, in general, have lower values for the COG. The reason for this is the fact that the voice source trace, which exists in the lowest parts of the frequency range, rearranges the spectral picture adding information to the low-frequency area.

The second spectral moment (M2, the spectral variance or standard deviation of the spectrum) measures the range of the random distribution (Jongman, Wayland, et al. 2000). It is calculated as:

$$\sigma = \sqrt{\frac{1}{N} \times \sum_{i=1}^{N} (x_i - \mu)^2}$$

where σ is the spectral variance, N is the number of random values x, and the μ is the spectral mean.

The second spectral moment primarily distinguishes sibilants from non-sibilants. The spectra of non-sibilants have higher dispersion (higher M2 value) than sibilants, which means that the energy is spread more extensively in the frequency dimension (Koenig et al. 2013). Also, the standard deviation is significant in defining the "weak" fricatives (fricatives with low centroid) (dentals and velars) (Shadle and Mair 1996). Spectral variance is considered by some authors the only dimension among which all four places of articulation could be distinguished (Jongman, Wayland, et al. 2000; Nissen and Fox 2005).

The skewness (M3, the spectral tilt) is a measure which determines the asymmetry of energy distribution (Jongman, Wayland, et al. 2000). The skewness values can have positive, negative and zero values, depending on the side of an abscissa toward which the tails of the distribution are extending. Positive skewness means that the energy concentration is in lower frequencies with tails towards the higher frequencies (tails towards the positive side of the x-axis). Positive skewness means that the energy is concentrated in the left side of the distribution (tails towards the negative side of the x-axis).

The skewness is calculated as:

$$Sk = \frac{\mu - Mo}{\sigma}$$

where *Sk* is the skewness, *Mo*, the mode of the distribution and σ the standard deviation.

The palatals have high skewness because these sounds have the most of their energy distributed in the lower frequency ranges while dentals/alveolars are expected to have negative skewness since the energy of its spectrum is located in the higher frequencies (Haley et al. 2010). It is considered that American English fricatives can be distinguished by spectral skewness (all four places of articulation) as by Jongman, Wang, et al. (2000).

Kurtosis is the measure which represents the peakedness of the spectral distribution. The higher the kurtosis, the more peaked is the spectrogram, that is, the values of high intensity (of frequency) are concentrated more around the mean. Flatter spectrograms display lower values of kurtosis.

The kurtosis of a spectrum is calculated as:

$$K = N \times \frac{\sum_{i=1}^{N} (x_i - \mu)^4}{\left(\sum_{i=1}^{N} (x_i - \mu)^2\right)^2}$$

where *K* is the kurtosis, N – the number of values (x) in a random distribution, and μ the spectral mean.

Kurtosis is expected to be high in the sounds which have higher intensity, while the low-intensity spectra should have low kurtosis.

Variability

Variability in the spectral moments is affected by the quality of the following vowel, speaker's gender and by the individual speakers' differences.

Postalveolar (palatal) fricatives are considered to be the most stable in the acoustic measures, and least prone to coarticulatory effects, unlike the dentals/alveolars, which exhibit greater variability of the acoustic signal. (Tabain 2001). However, various factors influence the overall results of the spectral moments' measurement. Lip rounding in some sibilant fricatives, vowel context, the tempo of speech, gender and inter-speaker differences all contribute to the variability.

Vowel quality is a compelling factor for the peak frequencies, and therefore the spectral moments' measures. For example, the peak frequency in the vicinity of [u] is lower than the peak frequency before [i] or [a]. (Jongman, Wayland, et al. 2000; Koenig et al. 2013; Shadle and Mair 1996). Vowels also affect non-sibilants more than sibilants, and also more voiceless than voiced fricatives (L. M. Jesus & Shadle, 2002).

Gender is another non-linguistic factor that causes variation. It is regarded that female speakers' speech has higher values of the first spectral moment than the males', which may be due to smaller cavity size, different degree of lip-rounding or different place of constriction formation. (Koenig et al. 2013). As well, females' speech yields more consistent results than the males' (e.g., Jones and Nolan 2007) regarding the first spectral moment measurement. Higher values for the first two spectral moments are also recorded for females while the third and fourth spectral moments are lower in females than in males (Jongman, Wayland, et al. 2000).

The procedure

The recordings were obtained from four female speakers, GLf, NBf, SSf, TKf, the same speech material was used in Petrović (2017). The material which consisted of real words and phonotactically correct nonsense words was examined with Praat Speech Analysis Software (Boersma and Weenink 2016). Nonsense words were chosen in places where there were no dictionary words for some consonant-vowel sequences. The choice of words, in the prosodic sense, follows the pattern of a fricative, plus a vowel with a short-falling accent. Some fricative-vowel combinations were taken as the fricative plus long-falling accent and the author considered them to serve the purpose well, having that some combinations of fricatives and vowels are non-existent, and should not affect the picture of the real properties of speech.

The 30ms Hamming windows were taken on three parts of the fricative: at the very onset (B), on the most prominent stable state part (M) and at the fricative offset (E) (cf. Koenig et al. 2013). Spectrogram pictures were extracted for each window. All spectral moments are computed at the power of 2. Before taking the measures, the extracted windows were pre-emphasised for frequencies above 80 Hz (Smith). The measured values were converted to Bark scale as:

$$Bark = 13 \times \arctan(0.00076 \times F) + 3.5 \times \arctan((F \div 7500)^2)$$

where F is a frequency in Hertz of the results obtained.

The Bark scale, proposed by Zwicker (1961), is a logarithmic function that divides the audible frequency into 24 bands and gives much more space for analysis in the higher frequency regions. "This causes greater weighting of energy in higher frequency ranges and an upward shift in spectral peaks" (Forrest et al. 1988). There are reports that frequencies converted to the Bark scale are more suitable for differentiating fricative types (Flipsen et al. 1999).

Some considerations need to be kept in mind while measuring spectral moments. Iskarous et al. (2011) claim that during the time-course the first spectral moment rises during the first half of the fricative duration and then falls towards its end. Also, some variability in measures can occur due to the higher signal/noise ratio and pre-emphasis. (Koenig et al. 2013).

The statistical computations were performed in IBM SPSS Statistics version 24. The multi-factorial ANOVAs were performed for both, strictly phonetic factors (voicing, place of articulation, quality of the following vowel), and also some prosodic (tempo of speech) and non-linguistic factors (speakers' performance). We have applied the narrow and wide scope of conditions required for asserting the contrasts we observed. In the narrow scope, we looked only for the significance of the interaction between the phonological voicing and the place of articulation in which way we were able to determine whether the phoneme with its main phonetic features behaves differently from other phonemes. In the broader scope, we looked at the interaction between voicing, place and the following vowel. This interaction would give us the proof that the phoneme depends not only on its primary phonetic features but also on the segment they precede. Also, it enabled us to check the claims made in previous research that vowels have a significant role in spectral measures of fricatives. The conclusions were drawn from both groups of results.

The factors with more than two values were computed in a post-hoc analysis to see if there was any significance in between-factor relations. Hierarchical regression analysis at the end of this paper included the R^2 change which was the primary procedure for weighing the cue's significance, while ANOVA tables in the regression procedure were used as an additional method. Tables with mean values are presented in the paper.

Results

Central of Gravity –- B

Hertz Scale

The first spectral moment reached a high significant main effect for the phonological voicing ($F(1, 8) = 141.514$, $p < 0.001$), place of articulation ($F(3, 8) = 75.836$, $p < 0.001$) and the mode of speech ($F(1, 8) = 45.862$, $p < 0.001$). The quality of the vowel showed to be significant ($F(4, 8) = 6.979$, $p = 0.01$), while the inter-speaker contrast was almost non-existent ($F(3, 8) = 1.028$, $p = 0.43$).

The post-hoc analysis of speakers' performance did not give significant results among any of the four speakers ($p > 0.05$), which allows for a claim that the first spectral moment is not speaker-dependent. The analysis of place rendered all classes as different, except for the contrast between the labials and velars ($p = 0.157$). These results can be seen on images of typical spectra for labials and velars, which show great dispersion in the frequency domain, and also high COG. Fricatives were different when the following vowels were [i] and [u] after [x] ($p = 0.023$).

There were also interactions between the factors: voicing and place of articulation had a high significance result ($F(2, 8) = 65.19$, $p < 0.001$) which is a signal that at the onset of fricatives there exists the difference with respect to the first spectral moment. The interaction between voicing and the following vowel was not obtained ($p = 0.502$), allowing for a claim that the vowel type does not influence the COG parameter at the beginning of the frication process.

Bark Scale

According to the results on the Bark-transformed scale, there was significance in the following factors: voicing ($F(1, 8) = 31.268$, $p = 0.001$), the place of articulation ($F(3, 8) = 7.180$, $p = 0.012$), and their interaction ($F(2, 8) = 10.277$, $p = 0.006$). The similarity exists in the mode of speech ($p = 0.803$), speakers' performance ($p = 0.937$), and vowel type ($p = 0.445$).

The post-hoc analysis of the place of articulation showed a significant difference between the labials and palatals ($p = 0.022$) and palatals and velars ($p = 0.011$). There was no difference in either speakers' performance or vowel type.

The difference in the co-effect of voicing and the place of articulation existed ($F(2, 8) = 10.277$, $p = 0.006$), while the difference in co-effect of the voicing, place, and the following vowel did not ($p = 0.788$), which is in line with the Hertz-measured values that the fricatives as units have a significant effect on each other, while their immediate phonological surrounding does not.

Centre of Gravity –- M

Hertz Scale

At the most prominent part of the fricative, on its stable-state, the ANOVA results showed that there exist differences in the voicing ($F(1, 8) = 62.58$, $p < 0.001$) and the place of articulation ($F(3, 8) = 110.664$, $p < 0.001$).

The post-hoc analysis of the place of articulation showed different results for the dental – velar opposition ($p = 0.046$). Mode of speech also yielded different results for tokens spoken in isolation and those spoken in the carrier sentence ($F(1, 8) = 23.912$, $p = 0.001$). Significance was obtained also in the speakers' performance ($F(3, 8) = 14.514$, $p = 0.001$). The post-hoc analysis of the speaker factors revealed a similarity between the performance of GLf and SSf ($p = 0.574$), while the other speakers performed differently. The vowels [a] and [e] were different before [v] ($p = 0.002$).

There was an interaction between the voicing and place of articulation ($F(2, 8) = 37.629$, $p < 0.001$) as well as between the voicing, place and vowel type ($F(7, 8) = 6.31$, $p = 0.009$). This means that coarticulatory effects are present in the middle part of the fricative.

Bark Scale

The ANOVA of Bark scale values resulted in the distinctive voicing ($F(1, 8) = 64.73$, $p < 0.001$), place of articulation ($F(3, 8) = 89.923$, $p < 0.001$), mode of speech ($F(1, 8) = 15.211$, $p = 0.005$) and speakers' performance ($F(3, 8) = 13.646$, $p = 0.002$).

The post-hoc analyses for the place of articulation showed that neutralisation of the Bark-scaled centre of gravity in the steady-state windows was only between the labial and dental ($p = 0.279$) and dental and velar ($p = 0.064$) fricatives. For the speaker factor, there were differences in performance between GLf and TKf ($p = 0.008$), NBf and TKf ($p = 0.001$), and SSf and TKf ($p = 0.024$), which means that only TKf performed differently from the other three speakers. There were differences in vowel factors for vowels [e] and the three other recorded vowels ($p \leq 0.032$) after [v].

There was a significant effect of the interaction between the voicing and the place of articulation ($F(2, 8) = 27.629$, $p < 0.001$) as well as the interaction between the voicing, place and the following vowel ($F(7, 8) = 7.041$, $p = 0.007$).

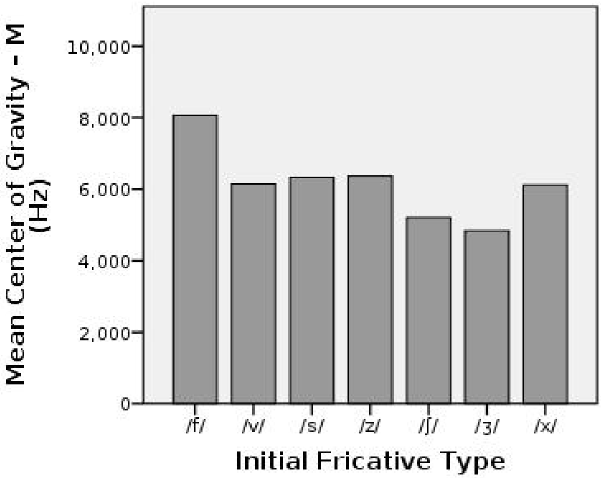


Figure 1*: Mean COG for each fricative in Bark scale –- the second window*

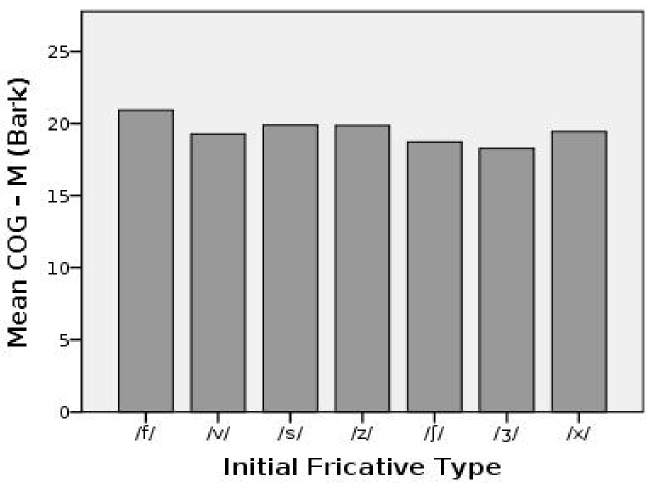


Figure 2*: Mean COG for each fricative in Bark scale –- the second window*

Centre of Gravity –- E

Herz Scale

ANOVAs of Hertz values showed the distinctive glottis mode ($F(1,8) = 98.1$, $p < 0.001$), the difference in the place of articulation ($F(3, 8) = 18.352$, $p = 0.001$), as well as mode of speech ($F(1, 8) = 111.772$, $p < 0.001$) and the vowel quality ($F(4, 8) = 12.013$, $p = 0.002$). Speakers' performance was also significant ($F(3, 8) = 54.641$, $p < 0.001$). The effect of interaction between the voicing and the place of articulation also existed ($F(2, 8) = 39.226$, $p < 0.001$), while the interaction between voicing, place and vowel type was insignificant ($p = 0.416$). Considering the high significance of interactions between the voicing, place, and vowel separately, voicing and place together and homogeneity of voicing, place and the following vowel contrast, we may conclude that, at the junction between a fricative and a vowel, the voicing contrast neutralises, which may be the reason for considering this part of the fricative to be the subject of coarticulation.

The post-hoc analyses of the significant factors between the groups showed the difference – for the place of articulation the similar were labial and dental fricatives ($p = 0.957$). At other places of articulation the results were within the proximity numbers. Vowel quality numbers differed [u] and the rest of the vowels ($p \leq 0.040$) except for [o] ($p = 1$). The different performance for the COG in the last window had speaker SSf.

Bark Scale

The Bark scale of the final portion of the fricative's first spectral moment gave significant differences for phonological voicing ($F(1, 8) = 36.409$, $p < 0.001$), mode of speech ($F(1, 8) = 51.044$, $p < 0.001$), vowel type ($F(4, 8) = 5.841$, $p = 0.017$). Speakers' performance also differed from one speaker to another ($F(3, 8) = 27.281$, $p < 0.001$) (the post-hoc showed the difference of SSf from all other three informants (GLf ($p = 0.001$), NBf ($p < 0.001$) and TKf ($p < 0.001$)). The different places of articulation were marginally similar ($p = 0.075$), but the interaction between the voicing and the place of articulation was highly significant ($F(2, 8) = 15.566$, $p = 0.002$). There were also no effects in the interaction between the voicing, place and the following vowel ($p = 0.714$).

Standard Deviation –- B

Hertz Scale

When measured in Hertz, the spectral variance yielded different results for the following factors: voicing ($F(1,8) = 165.72$, $p < 0.001$), place of articulation ($F(3, 8) = 262.145$, $p < 0.001$), mode of speech ($F(1, 8) = 83.99$, $p < 0.001$), vowel type ($F(4, 8) = 10.971$, $p = 0.002$), and speaker ($F(3, 8) = 25.658$, $p < 0.001$).

The post-hoc analyses of the place of articulation revealed differences between all groups (all places of articulation, with the *p*-value reaching at most 0.014 between the palatal and the dental). The vowel quality difference was not neutralised between [a] and [e] ($p = 0.049$) and [a] and [i] ($p = 0.024$) both after [x]. Speakers NBf and SSf performed approximately the same, and considerably different from the other two speakers. The other two speakers GLf and TKf also performed differently from each other.

There was no interaction between the voicing and place of articulation ($p = 0.203$). The interaction between voicing, place and the vowel was also insignificant ($p = 0.414$), which means that the fricatives as phonological units, cannot be distinguished solely on the parameter of spectral variance, nor is the combination of a fricative and a vowel a distinguishing feature when the spectral variance is regarded, at least at the beginning of the fricative noise.

Bark Scale

The bark scale gave similar results regarding the dependence of the second spectral moment on the factors: the voicing was highly significant ($F(1, 8) = 357.512$, $p < 0.001$), as well as the place of articulation ($F(3, 8) = 426.774$, $p < 0.001$) and the mode of speech ($F(1, 8) = 186.196$, $p < 0.001$). The vowel type was significant with the $F(4, 8) = 14.216$, $p = 0.001$. Speaker performance was highly significant ($F(3, 8) = 66.554$, $p < 0.001$).

The post-hoc analyses for the place of articulation showed differences between all four places ($p \leq 0.003$).

There was a vowels difference between [a] and [e] ($p = 0.016$) before [v]. Among other vowel combinations, *p-value* was not lower than 0.3.

The speaker's performance was significant, and the inter-speaker variability showed that SSf and NBf performed in the same manner, while the other two speakers, TKf and GLf performed differently from the other two and amongst themselves.

There also existed the effect for the interaction between the voicing and the place of articulation ($F(2, 8) = 13.535$, $p = 0.003$), but not for the voicing, place and the type of the vowel ($p = 0.253$). As opposed to the Hertz scale, calculations for the same fricative portion, this analysis yielded a significant interaction between the voicing and place, which means that Bark-transformed values have a greater impact on the distinction between this spectral moments measure of fricatives. However, the wider scope of significance was not obtained, since the vowel type did not add to the contrast.

Standard Deviation –- M

Hertz Scale

At the steady-state 30 ms Hamming windows, the significant results for the measures of standard deviation are for voicing ($F(1, 8) = 194.491$, $p < 0.001$), place of articulation ($F(3, 8) = 487.745$, $p < 0.001$), the mode of speech ($F(1, 8) = 47.009$, $p < 0.001$), vowel quality ($F(4, 8) = 14.045$, $p = 0.001$) and speaker ($F(3, 8) = 20.488$, $p < 0.001$).

The post-hoc analyses for the place showed differences between all fricative place categories ($p \leq 0.002$).

For vowel type the difference was between [a] and [i] after [x] ($p = 0.001$), [e] and [o] ($p = 0.029$) and [e] and [u] ($p < 0.001$) and [i] and [o], also after [x] ($p < 0.001$). GLf performed differently from all other speakers ($p \leq 0.002$), while the other speakers' performance was similar. ($p \geq 0.906$).

The effect of the interaction between voicing and place of articulation was also achieved ($F(2, 8) = 29.306$, $p < 0.001$), while with the added vowel the significant main effect was not obtained. These results point us to the conclusion that at the most prominent and stable part of the fricative, the phonemes themselves are distinguishable from one another, while the fricatives do not coarticulate with the following vowel.

Bark Scale

Bark-transformed frequencies, provided the following results: the difference existed between voicing ($F(1, 8) = 620.263$, $p < 0.001$), place of articulation ($F(3, 8) = 1644.654$, $p < 0.001$, type of the vowel ($F(4, 8) = 36.389$, $p < 0.001$), speaker ($F (3, 8) = 113.814$, $p < 0.001$) and mode of speech ($F(1, 8) = 118.213$, $p < 0.001$).

A significant difference is shown for all place groups ($p < 0.001$ for all groups); there were differences between each vowel and [u] ($p < 0.001$) except for [o] ($p \leq 0.016$), as well as between [a] and [i] ($p = 0.044$) before [x]; the speaker GLf performed differently from all other speakers ($p < 0.001$), while they amongst themselves performed within insignificance bounds ($p \geq 0.111$).

Interaction between the voicing and place of articulation was highly significant ($F(2, 8) = 9.607$, $p = 0.007$). There was also an interaction between voicing, place and vowel quality ($F(7, 8) = 5.161$, $p = 0.017$).

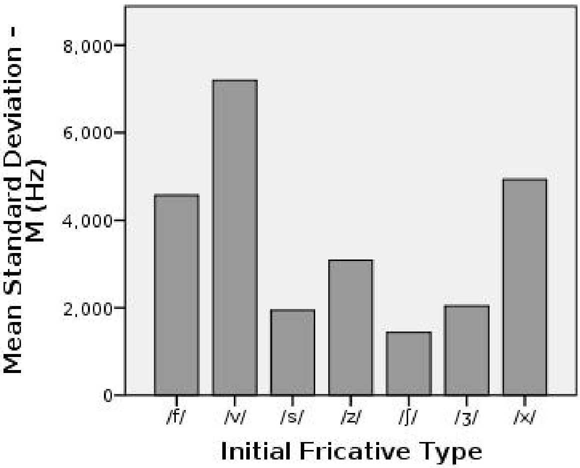


Figure 3*: Mean Spectral Variance in Hertz for each fricative –- the second measurement*

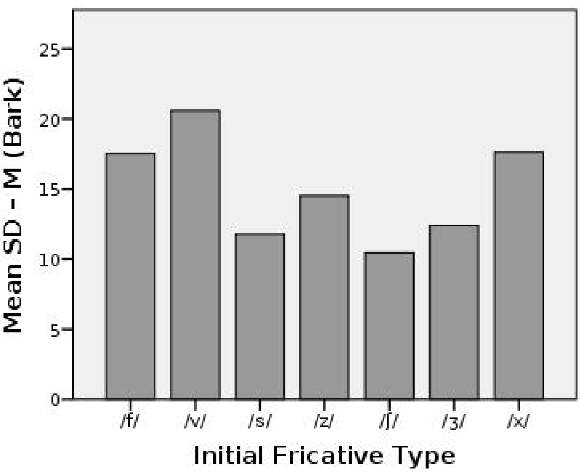


Figure 4*: Mean Spectral Variance in Barks for each fricative –- the second measurement*

Standard Deviation –- E

Hertz Scale

Standard deviation of the spectra distinguished between the following factors: the place of articulation ($F(3, 8) = 242.371$, $p < 0.001$), mode of speech ($F(1, 8) = 81.048$, $p < 0.001$), vowel type ($F(4, 8) = 13.889$, $p = 0.001$), and speaker ($F(3, 8) = 64.666$, $p < 0.001$). The voicing contrast was neutralised ($p = 0.617$).

The second spectral moment was neutralised for labials and velars ($p = 0.08$), while the other interactions showed a significance within $p \leq 0.018$; there was a significant difference between [a] and [u] ($p = 0.038$), [e] and [u] ($p = 0.001$), [i] and [o] ($p = 0.047$) and [i] and [u] ($p = 0.001$) all after [x]. [u] was different from all other vowels ($p \leq 0.038$) except for [o] ($p = 0.442$); the difference between speakers existed between TKf and GLf on one side and NBf and SSf on the other. These groups of speakers, respectively, did not differ from one another ($p \geq 0.706$).

The interaction between voicing and the place of articulation was non-significant ($p = 0.1$). Neither was the interaction between voicing, place and vowel type ($p = 0.401$). Such results point to the conclusion that the final part of the fricative does not give enough information about the fricatives differing amongst themselves, regarding the second spectral moment.

Bark Scale

Similar results are obtained in the Bark scale measurement. There was a non-significant voicing effect ($p = 0.464$). The other four factors had a highly significant main effect: the place of articulation ($F(3, 8) = 452.993$, $p < 0.001$), mode of speech ($F(1, 8) = 110.909$, $p < 0.001$, speaker ($F(3, 8) = 156.908$, $p < 0.001$) and vowel type ($F(4, 8) = 26.23$, $p < 0.001$).

In the post-hoc analyses there were significant differences between all groups ($p \leq 0.005$); vowels [i] and [u] differed from the others after [x] ($p = 0.008$), other groups – $p \geq 0.089$; GLf and TKf performed similarly from SSf and NBf ($p = 0.795$), which were different from each other ($p \leq 0.001$).

There was a significant main effect for the interaction between the place and voicing ($F(2, 8) = 5.546$, $p = 0.031$). The same is for the interaction between voicing, place, and type of vowel ($F(7, 8) = 4.191$, $p = 0.031$). This showed that the Bark-scaled values which are rather low for the spectral variance yielded a significant interaction between all the observed factors.

Skewness –- B

Hertz Scale

ANOVA results for skewness at the onset of frication noise showed significant results for voicing ($F(1, 8) = 6.081$, $p = 0.039$), place of articulation ($F(3, 8) = 60.262$, $p < 0.001$) and the following vowel ($F(4, 8) = 6.295$, $p = 0.014$).

The post-hoc results for skewness in the first window, show that for the place groups the contrast was neutralised for the labials and velars ($p =0.825$), as well as dentals and palatals ($p = 0.312$); for the following vowel there were no significant differences within-groups ($p \geq 0.183$) (except for the difference between [i] and [u] after [x] ($p = 0.013$)) as well as for the difference between the speakers ($p \geq 0.328$).

There was also an interaction between the voicing and place ($F(2, 8) = 17.909$, $p = 0.001$), but not between the voicing, place and the type of the following vowel ($p = 0.335$). which shows that skewness at the onset of the fricative is significant only for the fricative, while the interaction with a vowel type at this point does not exist.

Bark Scale

Bark-transformed values for skewness at the beginning of the fricative gave the following significant results: voicing ($F (1, 8) = 6.081$, $p = 0.039$), place of articulation ($F (3, 8) = 60.260$, $p < 0.001$) and vowel type ($F(4, 8) = 6.295$, $p = 0.014$). Results for the speaker and the mode of speech were insignificant ($p \geq 0.35$).

The post-hoc analyses showed that for skewness values, the labials and velars are on one side, while dentals and palatals are on the other. There were no significant differences between the vowels following the fricatives themselves ($p \geq 0.206$) except for the difference between [i] and [u] after [x] ($p = 0.013$), nor between the speakers amongst one another ($p \geq 0.328$).

The observed interactions, those between the voicing and the place and voicing, were significant ($F(2, 8) = 17.908$, $p = 0.001$) and the interaction between the place and the following vowel was not ($p = 0.335$).

Skewness –- M

Hertz Scale

The stable portion of fricatives showed the significant values below the 0.001 boundary for almost all the factors and their interactions: the voicing ($F(1, 8) =$

706.4, $p < 0.001$), the place of articulation ($F(3, 8) = 2803.478$, $p < 0.001$), the following vowel ($F(4, 8) = 55.788$, $p < 0.001$) and difference between the speakers ($F(3, 8) = 308.466$, $p < 0.001$). Somewhat greater *p*-value is recorded only for the mode of speech ($p = 0.026$), and some interactions with the vowel type.

The post-hoc tests reveal the difference in the place of articulation: only velars and palatals were behaving the same ($p = 0.999$). Other contrasts are all below 0.001 significance. The following vowel types were all neutralised, except for the contrast between [i] and [u] and [i] and [o] after [x] ($p \leq 0.005$). Speakers GLf and SSf performed the same ($p = 0.926$), while TKf and NBf performed differently from the previous group ($p = 0.039$ and $p < 0.001$) and differently amongst themselves ($p < 0.001$).

The interaction between the voicing and the place was highly significant ($F(2, 8) = 635.438$, $p < 0.001$), as well between voicing, place and the following vowel ($F(7, 8) = 57.913$, $p < 0.001$). The most stable part of the fricative is different among fricatives concerning both the fricative phoneme and the vowel that follows it.

Bark Scale

At the midpoint of the fricative, the main effect was obtained for the following: voicing ($F(2, 8) = 706.439$, $p < 0.001$), place of articulation ($F(3, 8) = 2803.616$, $p < 0.001$), following vowel ($F(4, 8) = 55.709$, $p < 0.001$, speaker ($F(3, 8) = 308.481$, $p < 0.001$) and mode of speech ($F(1, 8) = 7.5$, $p = 0.026$).

Among the fricatives with different places of articulation, there were differences among all places ($p < 0.001$) except for the difference between the palatals and velars ($p = 0.999$). Vowels after the fricatives were all below the significance threshold, except for the difference between [i] and [u] ($p = 0.02$) after [x]. Speakers were divided in performance between GLf, NBf and SSf and TKf ($p \geq 0.093$).

The interactions of interest were significant: the voicing and place ($F(2, 8) = 635.471$, $p < 0.001$) and voicing and place and the following vowel ($F(7, 8) = 57.915$, $p < 0.001$).

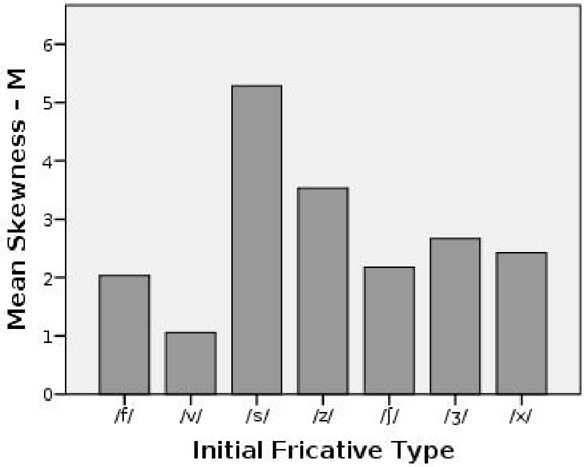


Figure 5*: Mean Spectral Skewness for each fricative –- the second window measures (Hertz).*

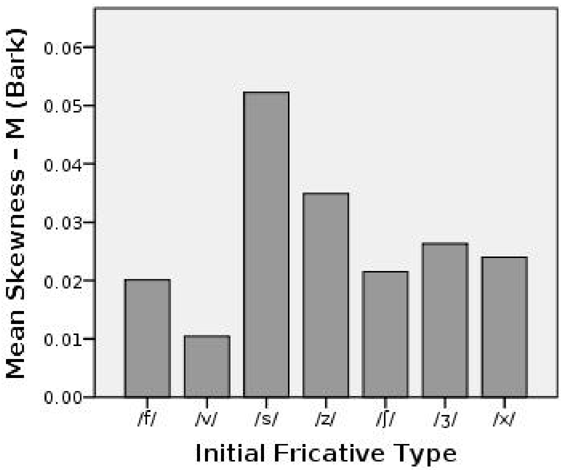


Figure 6*: Mean Spectral Skewness for each fricative –- the second window measures (Bark).*

Skewness –- E

Hertz Scale

At the end of the fricative noise, the difference between the voiced and voiceless fricatives almost neutralised ($F(1, 8) = 5.514$, $p = 0.047$). Place of articulation remained distinct ($F(3, 8) = 65.188$, $p < 0.001$). The difference existed for the vowel type ($F(4, 8) = 8.544$, $p = 0.005$), as well as for the mode of speech ($F(1, 8) = 29.505$, $p = 0.001$) and speaker ($F(3, 8) = 20.712$, $p < 0.001$).

The post-hoc analyses of factors are – for the place of articulation: labials and velars neutralised ($p = 0.206$), as well as dentals and palatals ($p = 0.872$); There was no between-group vowel significance ($p \geq 0.293$), except for the difference between [i] and [u] after [x] ($p = 0.02$); speakers: GLf and TKf performed similarly ($p = 0.234$), NBf performed similarly to SSf ($p = 0.067$) and TKf ($p = 0.094$)

The interaction between the voicing and the place of articulation was distinguishing ($F(2, 8) = 5.341$, $p = 0.034$), while the added vowel factor did not yield a significant interaction ($p = 0.185$).

Bark Scale

At the end of the fricative, the skewness value was on the boundary for voicing distinction ($p = 0.047$). The place of articulation ($F(3, 8) = 65.188$, $p < 0.001$), succeeding vowel ($F(4, 8) = 8.544$, $p = 0.005$), speaker ($F(3, 8) = 20.713$, $p < 0.001$) and mode of speech ($F(1, 8) = 29.505$, $p = 0.001$) all displayed relevant contrast.

The difference between labials and velars and dentals and palatals was obtained ($p \geq 0.206$). Vowels were not different from each other ($p \geq 0.293$). It is only after [x] that [i] and [u] were different ($p = 0.02$).

The voicing and the place of articulation together were significant ($F(2, 8) = 5.341$, $p = 0.034$) while the voicing, place and the following vowel were not ($p = 0.185$).

Kurtosis –- B

Hertz Scale

Results for kurtosis of the spectra at the onset of the fricative showed distinction for the following factors: voicing ($F(1, 8) = 75.173$, $p < 0.001$), place of articulation ($F(3, 8) = 102.105$, $p < 0.001$), vowel type ($F(4, 8) = 5.332$, $p = 0.022$), speaker ($F(3,8) = 13.521$, $p = 0.002$) and the mode of speech ($F(1, 8) = 11.871$, $p = 0.009$).

Distinction between labials and velars was non-significant ($p = 0.229$), as well as between dentals and palatals ($p = 0.085$). There was no between-group significance in the vowel type ($p \geq 0.27$); GLf performed the same as TKf ($p = 0.110$), NBf performed similarly to SSf ($p = 0.892$). The similar groups of speakers reached no more than 0.036 in *p* value significance.

There was a significant interaction between voicing and place of articulation (F(2, 8) = 17.832, p = 0.001), but not between voicing, place and the following vowel (p = 0.236), which shows the dependence of the contrast on the class of the fricative phoneme exists, but not on the neighbouring segments.

Bark Scale

Bark-transformed values for kurtosis at the beginning of the fricative noise have achieved the main effect for voicing (F(1, 8) = 75.069, p < 0.001), place of articulation (F(3, 8) = 102.036, p < 0.001), vowel type (F(4, 8) = 5.392, p = 0.022), speaker (F(3, 8) = 13.508, p = 0.002) and mode (F(1, 8) = 11.872, p = 0.009).

This parameter split labials and velar and dental and palatal fricatives (p = 0.229). Vowels amongst themselves were not significantly different (p ≥ 0.269). Speakers were divided in performance between GLf and TKf and NBf and SSf (p ≤ 0.036).

Interaction between voicing and place of articulation was significant (F(2, 8) = 17.815, p = 0.001). That between voicing, place and vowel type was not (p = 0.412).

Kurtosis –- M

Hertz Scale

At the steady-state window all the factors had significant results, as well as their interactions: voicing (F(1, 8) = 10757.897, p < 0.001), place of articulation (F(3, 8) = 36455.54, p < 0.001), vowel type (F(4, 8) = 876.875, p < 0.001), speaker (F(3, 8) = 2641.011, p < 0.001) and mode of speech (F(1, 8) = 175.107, p < 0.001).

All the differences were significant (p < 0.001); There was a difference between [i] and [u] after [x]. All speakers performed differently (p < 0.001).

The interaction between voicing and place of articulation was significant (F(2, 8) = 3898.047, p < 0.001). So is the interaction between the voicing, place and vowel (F(7, 8) = 598.103, p < 0.001): the phoneme and its environment contribute to the distinction between fricatives' kurtosis value on the steady-state portion of the fricative.

Bark Scale

In the Bark scale all factors reached highly significant values: voicing (F(1, 8) = 10745.079, p < 0.001), place of articulation (F(3, 8) = 36413.4, p < 0.001), following vowel (F(4, 8) = 876.668, p < 0.001), speaker (F(3, 8) = 2639.23, p < 0.001), mode of speech (F(1, 8) = 174.682, p < 0.001).

All places of articulation were different between themselves ($p < 0.001$) as well as vowels ($p < 0.001$) (except for the liminal significance between [i] and [u] after [x] ($p = 0.043$)). Each speaker performed differently from the other ($p < 0.001$).

There was interaction between the voicing and place ($F(2, 8) = 3888.65$, $p < 0.001$) and the voicing, place and vowel type ($F(7, 8) = 596.842$, $p < 0.001$).

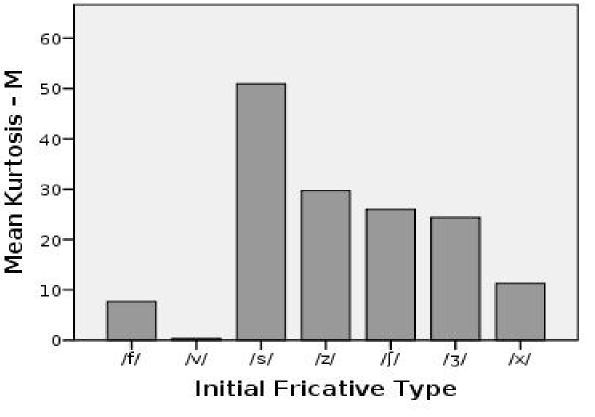


Figure 7*: Mean kurtosis for each fricative –- the second measured window (Hertz).*

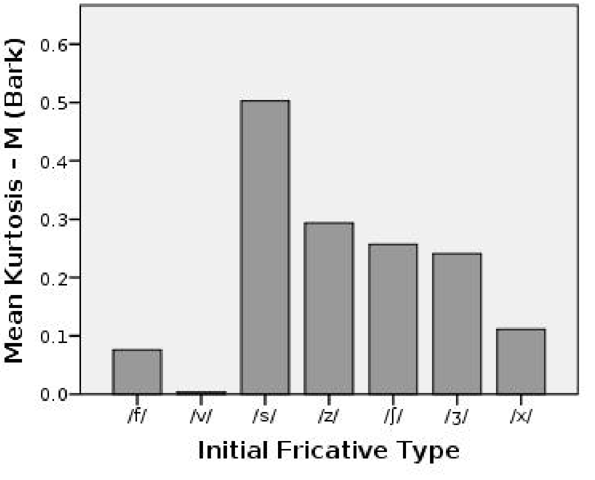


Figure 8*: Mean kurtosis for each fricative –- the second measured window (Bark).*

Kurtosis –- E

Hertz Scale

ANOVA results for kurtosis at the junction of a fricative with a vowel yielded significant differences for the place of articulation ($F(3, 8) = 381.942$, $p < 0.001$), vowel type ($F(4, 8) = 35.664$, $p < 0.001$), speaker ($F(3, 8) = 2258.067$, $p < 0.001$) and the mode of speech ($F(1, 8) = 54.033$, $p < 0.001$).

There was a boundary neutralisation between labials and palatals ($p = 0.079$), while the other groups were different ($p \leq 0.002$); There were no significant differences between vowels as a group after each fricative; GLf and TKf performed the same ($p = 0.092$), while the other two informants performed differently from the first two ($p \leq 0.001$) and differently amongst themselves ($p \leq 0.002$).

Voicing was insignificant on the third measured spectral window ($p = 0.081$) as well as the interaction between the voicing and place of articulation ($p = 0.186$). On the other side, the interaction between the voicing, place, and vowel was significant ($F(7, 8) = 10.951$, $p = 0.002$), which leads to the conclusion that although the voicing and the whole phoneme are neutralised, the impact of the following vowel is strong and shows different kurtosis at the end of the fricative.

Bark Scale

Bark-scaled kurtosis values at the end of the fricative yielded non-significant results for the phonological voicing ($p = 0.082$). The other four factors were significant: place of articulation ($F(3, 8) = 381.637$, $p < 0.001$), vowel type ($F(4, 8) = 35.648$, $p < 0.001$), speaker ($F(3, 8) = 181.487$, $p < 0.001$) and mode of speech ($F(1, 8) = 54.098$, $p < 0.001$).

Dental and palatal fricatives were different from the labial and velar ($p = 0.078$). There were no significant differences when the vowel was taken as a factor. GLf and TKf performed similarly; differently from NBf and SSf ($p = 0.092$).

The interaction between voicing and place was insignificant ($p = 0.187$), and between voicing, place and vowel type significant ($F(7, 8) = 10.920$, $p = 0.002$).

The movement of spectral moments' curves

Jongman et al. (2000) have presented us with the trajectories of curves that display the values of spectral moments during the time-course of the fricative. It is said that the line for spectral mean starts low at the beginning of the fricative, at the most prominent part has the highest value, while it descends towards the end of the fricative and more so at the interface between the fricative and the vowel. Somewhat similar results are obtained in our study. The only missing part is the results from the window that consist of parts of the fricative and the vowel, but, considering the low frequencies where the vowels' spectral energy is concentrated, the result would be much alike. The following graphs (9 through 12) have been obtained by averaging the values across all speakers and modes of speech.

The ANOVA analyses with the three windows for the four spectral moments and the type of the fricative as a factor gave a significant difference between all windows and all moments, meaning that values for the four spectral moments vary with the fricative progression, and depend on the type of fricative.

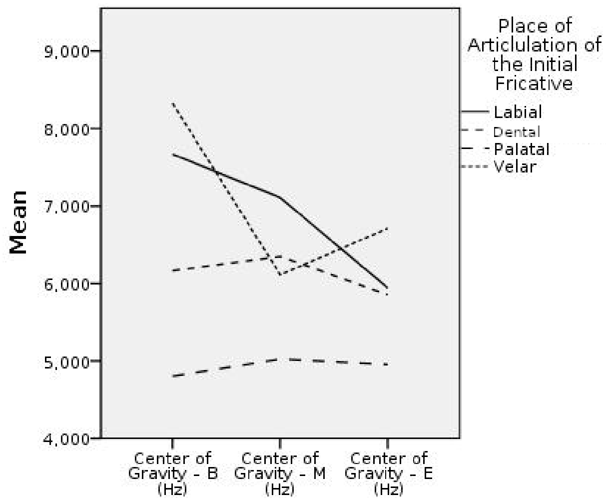


Figure 9*: The center of gravity movement during the time-course of fricative duration –- approximated by all speakers and speech modes*

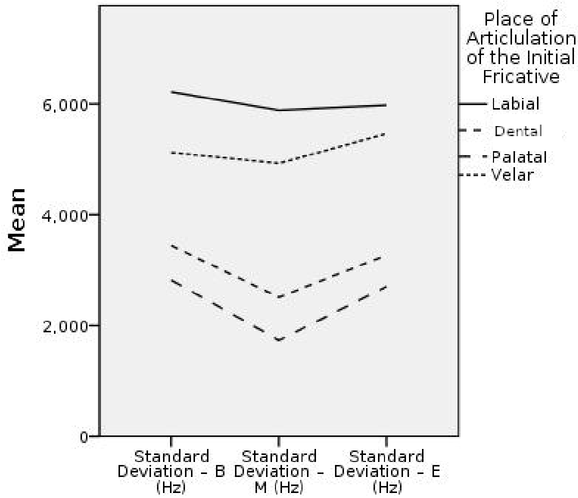


Figure 10*: The spectral variance movement during the time-course of fricative duration –- approximated by all speakers and speech modes*

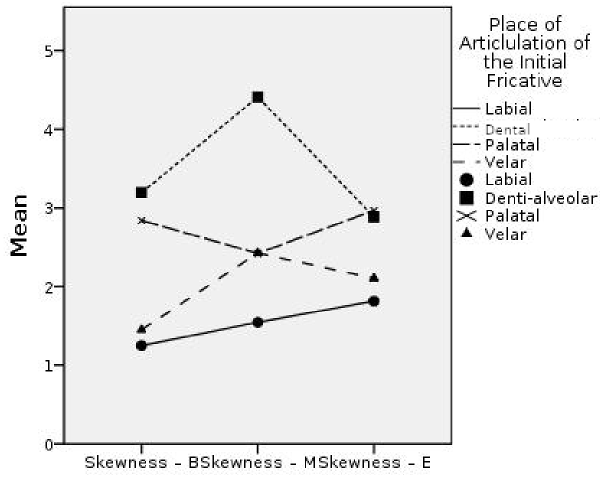


Figure 11*: The skewness movement during the time-course of fricative duration –- approximated by all speakers and speech modes*

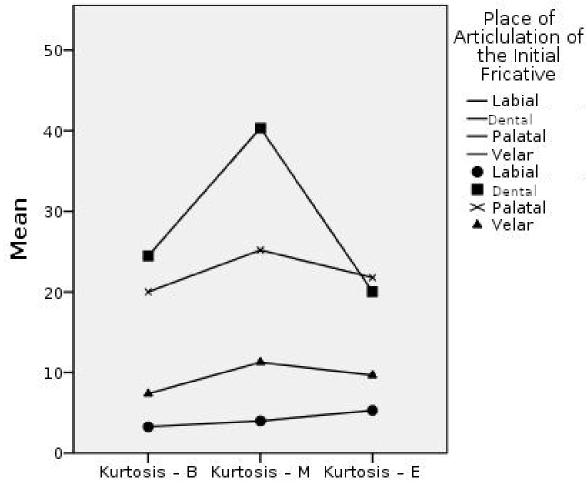


Figure 12*: The kurtosis movement during the time-course of fricative duration –- approximated by all speakers and speech modes*

We may see that apart from the notable difference between the four spectral moments for each place of articulation, the claims of (Iskarous et al. 2011) that the value of spectral moments changes as the fricative progresses are correct. The curve is boomerang-shaped, with values rising to the most prominent part and then falling towards the end of the vowel. The exception here is the velar fricative, which is represented only with the voiceless cognate. The following graphs represent the split projection of COG curves, pooled across all speakers and speech modes.

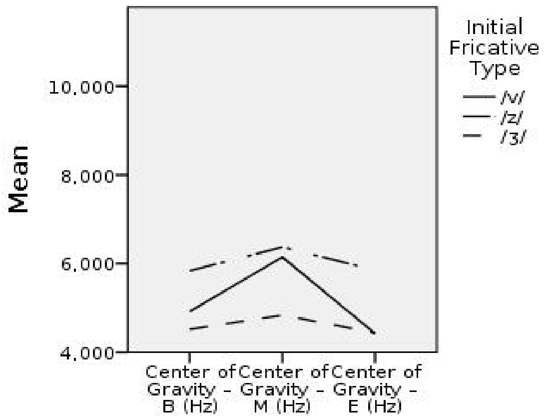


Figure 13*: The time curve of voiced fricatives' COG pooled across all speakers and speech modes*

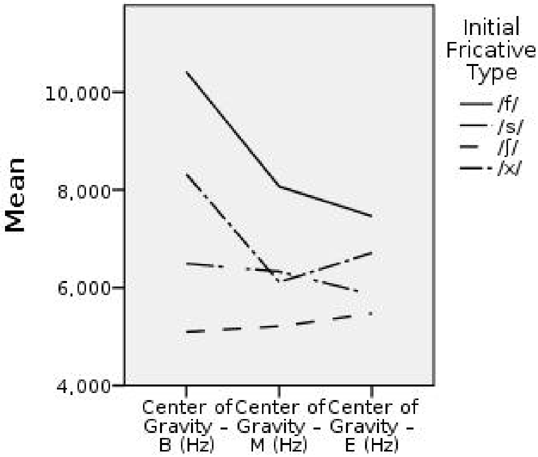


Figure 14*: The time curve of voiceless fricatives' COG pooled across all speakers and speech modes*

The additional results make a distinction between the voiced and voiceless fricatives. Voiceless fricatives all surfaced with a rise-fall curve, unlike the voiceless where only the sibilants have a rise-fall shape, while the non-sibilants show a fall-rise line. What we may say so far is that the energy concentration of voiceless fricatives is high at the beginning, and descending towards the steady-state part of the vowel. The voiced fricatives show the opposite tendency: the energy at the onset of the fricatives is low, while at the most prominent part the spectral energy rises. The fall in voiced and the rise in voiceless fricatives after the steady-state window is the leveling with the following vowel's spectral energy.

Regression analysis

In order to examine the most dominant factor in differentiation between fricatives in Serbian, a series of hierarchical regression analyses were conducted on all three windows and all four spectral moments. It is considered that the R^2 parameter of a factor which has a higher value makes it the most dominant in affecting the behaviour of segments in great extent.

The factors that were put in the regression test were: mode of speech, speaker, post-fricative vowel type, place of articulation and voicing, respectively.

The beginning of the fricative (the first measured window) yielded the R^2 change coefficient 0.154 for voicing which was significant below 0.001 value. All other factors had much smaller values of R^2 change.

The steady-state portion of the fricative (the second measured window) gave somewhat greater results and greater significance for the place of articulation than for voicing, although the ANOVA table shows the greater dependence of the second window's COG on the voicing than on the place of articulation.

The juncture between the fricative and a vowel, logically, gave the greatest R^2 parameter for the voicing, having in mind that at this point in an utterance, effects of coarticulation are most present. The place of articulation had a great insignificance value.

The most significant impact on the change of spectral variance at the onset of fricative noise was recorded for a place of articulation (R^2 change = 0.094, $p < 0.001$), and after it for the voicing (R^2 change = 0.02, $p = 0.01$).

For the central part of the fricative, the results were similar. The place of articulation had the greatest impact (R^2 change = 0.099, $p < 0.001$), after which came voicing (R^2 change = 0.011, $p = 0.063$).

The third measured window also gave the most significant results for the place of articulation.

Spectral skewness gave somewhat ambiguous results when it comes to the dominant factor of dependence. In the first windows, the most significant was the place of articulation (R^2 change = 0.013, $p = 0.062$). All other factors were far below the significance level. The central part of the fricative's skewness resulted in the most significant voicing (R^2 change = 0.015, $p = 0.042$) and from random factors, the most prominent was the speakers' influence (R^2 change = 0.023, $p = 0.011$). The third measured window (at the boundary between a consonantal and vocalic segment) showed the most significant dependence on the place of articulation (R^2 change = 0.016, $p = 0.031$).

The dominant factor for the spectral kurtosis was a phonological voicing of the target fricative. The second and third windows showed a great impact for both the place of articulation and voicing, although the ANOVA table gave greater significance to the voicing. Only at the fricative offset did the place of articulation have greater significance than the glottal state of the fricative.

Summing up the results of the above analyses, we may conclude that for the COG, the factors that most effectively distinguish fricatives were voicing and the place of articulation. The spectral variance depends mostly on the place of articulation followed by voicing. Skewness seems to behave randomly when the distinctive features are regarded, but we may say that, besides the variability that is caused by the speakers' performance, the place of articulation and voicing pertain mostly to the difference between the skewness in fricatives. And finally, the spectral kurtosis is mostly dependent on voicing, after which comes the place of articulation.

Conclusion

The statistical results have shown that the general distinction between the dentals and palatals on the one side and labials and velars on the other exists in almost all aspects of measurement. It is a phenomenon observable from the sole spectral images. Produced with high energy, the sibilant fricatives show the lower centre of gravity, smaller spectral variance, and greater kurtosis. On the other hand, the dispersed energy in the spectra of non-sibilants implies a higher standard deviation and lower kurtosis. The skewness is positive in the sampling rate of 22.050 Hz, except for some occurrences of sibilants which had the negative skew.

If we look further into the details, the highest centre of gravity at the onset of the fricative is recorded for the voiceless labial [f], which is located around 8 kHz. The voiced counterpart of the labial fricative had the COG at the level of 6 kHz. The value for the velar [x] is next in magnitude and is at about 8 kHz. The sibilant fricatives are all below the non-sibilants. The [s] has the highest spectral mean, somewhat higher than 6 kHz. The next is its voiced counterpart, which shows the COG value below 6 kHz. The palatal voiceless and voiced fricatives seemed to have the lowest first spectral moment, which is around 5 kHz (for the voiceless) and 4 kHz (for the voiced). At the middle part of the fricative, we can see some leveling of the labials. The frequency for the COG for the voiced labial is at about 6 kHz. Among sibilants, the dentals have higher frequencies than the palatals for, approximately, 1 kHz, which is consistent with the previous findings. The order of values of the Hertz scale COG at the mid-point of the fricative is approximately the same as it was at the beginning of the fricative. The values for the sibilants seem to have leveled, which may be due to the voiceless pronunciation of a voiced fricative at the stable state (mostly for dentals). The presented results are along the claims of Haley et al. (2010) and Maniwa et al. (2009) that dental fricatives have higher COG values than the palatals. Also, as presented in Tabain (1998) the non-sibilant [f, v] have higher COG values that the sibilants. The velar non-sibilants' results are at the level of non-sibilant labials.

It seems that the fricatives that have the acoustic energy spread mostly across the spectrum are the voiced fricatives of all places. The order of values for spectral variance of the voiced fricatives is [v] (above 7 kHz), [z] (around 4 kHz), [ʒ] (around 3.,5 kHz). The voiceless fricatives possess far lower values for the spectral variance. The order of values is [f] (somewhat below 6 kHz), [x] (around 5 kHz), [s] (something somewhat above 2 kHz), [ʃ] (at about 2 kHz). The medial and final windows showed the same order of values for each fricative, except that the values (variation) are somewhat lower in the former than the values at the beginning and end of the fricatives. The second spectral moment seems to discriminate well between the sibilants and non-sibilants (Koenig et al. 2013; Shadle and Mair 1996) the sibilant fricatives having the less dispersed energy (thus, less spectral variance) in the frequency domain than the non-sibilants.

The fricatives with the highest skew are the sibilants. The greater skew value is recorded for the dental (at around 4) than for the palatal (something less than 3) voiceless fricatives. The voiced counterparts of the sibilant fricatives are next in skewness with 2.5 for the voiced palatal and somewhat less for the voiced dental fricative. The flat spectra of [f], [v] and [x] have significantly lower values. At the central window, the third spectral moment for dentals (voiceless and voiced) exceeds the values of both palatal fricatives. Somewhat greater skewness is recorded in the velars, while the values for labials are on the level of palatal sibilants. The claims that dental fricatives are expected to have negative skewness did not show in our experiment (cf. Haley et al. 2010), although the author is a witness of some skewness values tailing towards the negative side.

The most peaked spectrum can be found in sibilant voiceless fricatives, dentals having significantly greater value than the palatals. Voiced sibilants are next in, respectively to the order of the place of articulation. At non-sibilants, the velar [x] has the highest spectral kurtosis, after which [f], and [v] follow. The middle window, as for skewness, showed greater values for the dentals than palatals including both voicing modes. The voiceless have higher kurtosis than the voiced. Next are, as in the first part of the fricative [x], [f] and [v]. The third window values seem alike similar to those in the first window.

Returning to the question we asked at the beginning of the paper, on the narrow and the wide scope of the dependence of the spectral moments on the factors and their interactions, we may look at Table 1, which shows the complete listing of factors and spectral moments with labels for significance.

Table 1: Significance of the four spectral moments on the all the factors and interactions at the beginning (B), middle (M) and end (E) part of the fricative (Hertz measurements).

Moments	Centre of Gravity			Spectral Variance			Spectral Skewness			Spectral Kurtosis		
Factors	B	M	E	B	M	E	B	M	E	B	M	E
Voicing	sign.	sign.	sign.	sign.	sign.	insign.	sign.	sign.	sign.	sign.	sign.	insign.
Place	sign.	sign.	sign.	sign.	sign.	sign.	sign.	sign.	sign.	sign.	sign.	sign.
Following Vowel	sign.	insign.	sign.	sign.	sign.	sign.	sign.	sign.	sign.	sign.	sign.	sign.
Speaker	insign.	sign.	sign.	sign.	sign.	sign.	insign.	sign.	sign.	sign.	sign.	sign.
Mode of Speech	sign.	sign.	sign.	sign.	sign.	sign.	insign.	sign.	sign.	sign.	sign.	Sign.
Voicing x Place	sign.	sign.	sign.	insign.	sign.	insign.	sign.	sign.	sign.	sign.	sign.	insign.
Voicing x Place x Following Vowel	insign.	sign.	insign.	insign.	insign.	insign.	insign.	sign.	insign.	insign.	sign.	insign.

Table 2: Significance of the four spectral moments on all the factors and interactions at the beginning (B), middle (M) and end (E) part of the fricative. (Bark measurements).

Moments	Centre of Gravity			Spectral Variance			Spectral Skewness			Spectral Kurtosis		
Factors	B	M	E	B	M	E	B	M	E	B	M	E
Voicing	sign.	sign.	sign.	sign.	sign.	insign.	sign.	sign.	sign.	sign.	sign.	insign.
Place	sign.	sign.	sign.	sign.	sign.	sign.	sign.	sign.	sign.	sign.	sign.	sign.
Following Vowel	insign.	insign.	sign.	sign.	sign.	sign.	sign.	sign.	sign.	sign.	sign.	sign.
Speaker	insign.	sign.	sign.	sign.	sign.	sign	insign.	sign.	sign.	sign.	sign.	sign.

Mode of Speech	insign.	sign.	sign.	sign.	sign.	sign.	insign.	sign.	sign.	sign.	sign.	sign.
Voicing x Place	sign.	sign.	sign.	sign.	sign.	sign.	sign.	sign.	sign.	sign.	sign.	insign.
Voicing x Place x Following Vowel	insign.	sign.	insign.	insign.	sign.	sign.	insign.	sign.	insign.	insign.	sign.	sign.

Regarding the narrow scope of significance, where only the voicing and place features interaction needs to be significant in order to call the two fricatives distinct (with respect to spectral moments), we see that the interaction between voicing and place is significant in all parts of the centre of gravity, middle part of the spectral variance, all parts of the spectral skewness and the first two parts of the spectral kurtosis. If we say that fricatives can be distinguished by all four moments, it is the middle, the most stable part of the fricative, that carries information about all fricative types for all four spectral moments. We can also claim that fricatives can be distinguished by the centre of gravity and spectral skewness and almost spectral kurtosis in every part of the fricative, but not by the spectral variance.

The broader scope, where we declared that, besides voicing and place, the vowel type is required to be included in the interaction, we may say that the part carrying the most information about the following vowel is the most stable portion of the fricative. Also, the picture is not complete, since there was no interaction at any part of the fricative for these three factors with regard to the spectral variance.

On the other hand, the Bark scale of values made a far more significant distinction between the fricative types (as in Flipsen et al. 1999). The two-factor interaction that between the voicing and place seemed to be important in all parts of all four spectral moments except for kurtosis at the end of the fricative spectrum. The stricter condition of importance still gives more significant results since, besides the central part of the fricative, the spectral variance and spectral kurtosis give, also, significant results.

Table 3: Mean spectral moments' values approximated between all three measured windows

Place	Voicing	1st (Hz)	2nd (Hz)	3rd (Hz)	4th (Hz)	1st (Bark)	2nd (Bark)	3rd (Bark)	4th (Bark)
Labial	Voiceless	8650.22	5366.82	1.4981	5.4658	20.9682	18.5044	0.0148	0.0540
	Voiced	5162.25	6687.35	1.5746	2.8869	17.8217	20.0756	0.0156	0.0285
Dental	Voiceless	6226.52	2614.72	4.1166	35.2087	19.5514	13.5110	0.0407	0.3477
	Voiced	6020.57	3539.52	2.8758	21.3393	19.2588	15.4844	0.0284	0.2108
Palatal	Voiceless	5259.16	2143.84	2.7460	24.3176	18.6689	12.3819	0.0271	0.2402
	Voiced	4598.29	2700.69	2.7404	20.3328	17.9392	13.9817	0.0271	0.2009
Velar	Voiceless	7050.29	5174.55	1.9929	9.4264	19.7208	18.0093	0.0197	0.0931

The factors that influence the spectral moments the most are voicing and place of articulation. The first spectral moment is more dependent on the voicing than on the place of articulation. Spectral variance varies mostly with changes in the state of the glottis. Skewness is mostly speaker-dependent, but also is connected to voicing. Kurtosis depends primarily on voicing.

The values of spectral moments change in the time-course of the fricative duration. After the initial low values, the curve rises at the middle part and falls as the fricative progresses towards the end and the next vowel. There exists a voicing distinction in the concentration of spectral energy at the beginning and the middle part of the fricatives. Voiced fricatives have low energy at the onset, which rises at the steady-state portion. The opposite is noticed in voiceless fricatives.

The vowel that follows the initial fricative seemed to differentiate spectral moments of mostly non-sibilant fricatives, more precisely the velar fricative [x]. This is by claims that the /h/-sounds are the voiceless counterparts of the following vowels (Ladefoged and Maddieson 1996; Tabain 1998). This may be partially applied to the Serbian velar [x]. Partially, because the significant differences were found only in spatially distant high vowels [i] and [u].

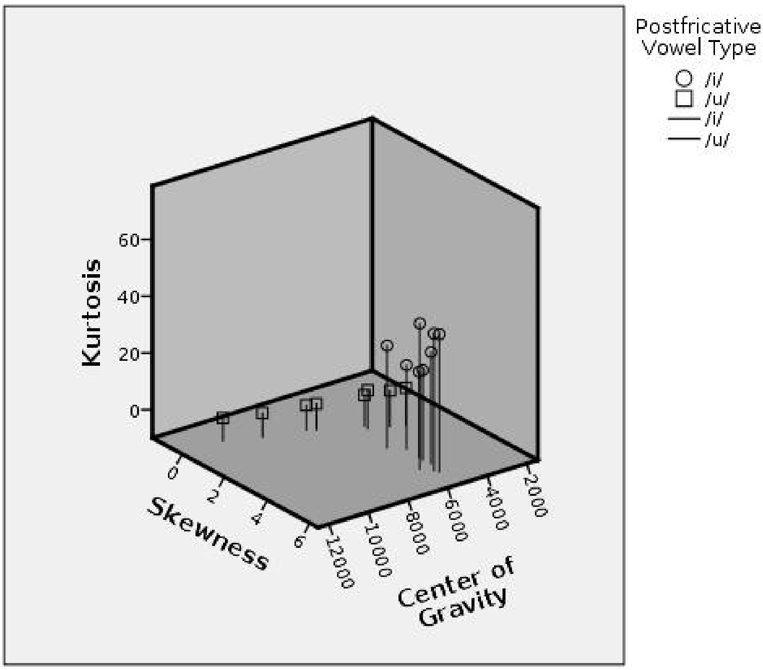


Figure 15: Three-dimensional location of spectral moments in function of the type of the following vowel for velar fricative [x]

Figure 15 represents the location of the fricative and the adjoined vowel in the three-dimensional system with COG, skewness, and kurtosis as the axes (as in (Forrest et al. 1988).

In the end, we may say that the results obtained were an approximation of our four female speakers (for inter-gender variability see the *Variability* section). The statistical difference between the interviewed speakers divided them, on most occasions, into two groups: GLf and TKf on one side, and SSf and NBf on the other. It is a perceived impression that the former group produced fewer quality data than the latter group regarding loudness and articulatory precision.

At the end one point must be made. In the results presented, the fricatives before short and long accents were treated as being the same. The future research should give answers to the question of whether the different length of the following vowel plays any significant role in the manifestation of four spectral moments in the fricatives in Serbian. Another problem that needs to be addressed in the future research is the high discrepancy between [x] sound before different vowel qualities, although there are opinions (Ladefoged, 1996) that the /h/-sounds are voiceless counterparts of the vowels they precede.

References

1. Batas, Ana (2014), 'Progresivno obezvučavanje eksploziva u proklizi i enklizi kao individualna fonetska pojava u srpskom jeziku', *Zbornik Matice srpske za filologiju i lingvistiku* LVII(2), 61–75.
2. Boersma Paul, and David Weenink (2016), *Praat: Doing Phonetics by Computer [Computer Program]*, http://www.praat.org/%0A%0A.
3. Flipsen JR, Peter et al. (1999), 'Acoustic Characteristics of/s/in Adolescents', *Journal of Speech, Language, and Hearing Research* 42(3), 663–77.
4. Forrest, Karen, Gary Weismar, Paul Milenkovic and Ronald N Dougall (1988), 'Statistical Analysis of Word-Initial Voiceless Obstruents: Preliminary Data', *The Journal of the Acoustical Society of America* 84(1), 115–23.
5. GUDURIĆ, Snežana and Dragoljub PETROVIĆ (2007), 'O Prirodi i Statusu Foneme [v] u Srpskom Jeziku', *Zbornik radova Filozofskog fakulteta u Prištini* (36), 321–40.
6. Haley, Katarina L. (2002), 'Temporal and Spectral Properties of Voiceless Fricatives in Aphasia and Apraxia of Speech', *Aphasiology* 16(4–6), 595–607.
7. Haley, Katarina L, Elizabeth Seelinger, Kerry Callahan Mandulak and David J Zajac (2010), 'Evaluating the Spectral Distinction between Sibilant Fricatives through a Speaker-Centered Approach', *Journal of Phonetics* 38(4), 548–54.
8. House, Arthur S and Grant Fairbanks (1953), 'The Influence of Consonant Environment upon the Secondary Acoustical Characteristics of Vowels', *The Journal of the Acoustical Society of America* 25(1), 105–13.
9. Iskarous, Khalil, Christine H Shadle and Michael I Proctor (2011), 'Articulatory–Acoustic Kinematics: The Production of American English/s', *The Journal of the Acoustical Society of America* 129(2), 944–54.
10. Jesus, Luis MT and Christine H Shadle (2002), 'A Parametric Study of the Spectral Characteristics of European Portuguese Fricatives', *Journal of Phonetics* 30(3), 437–64.

11. Jones, Mark J and Francis J Nolan (2007), 'An Acoustic Study of North Welsh Voiceless Fricatives', In *Proc. 16th ICPhS*. 873–6.
12. Jongman, Allard, Yue Wang and Joan Sereno (2000a), 'Acoustic and Perceptual Properties of English Fricatives', In *Sixth International Conference on Spoken Language Processing*.
13. Jongman, Allard, Ratree Wayland and Serena Wong (2000b), 'Acoustic Characteristics of English Fricatives', *The Journal of the Acoustical Society of America* 108(3), 1252–63.
14. Koenig, Laura L, Christine H Shadle, Jonathan L Preston and Christine R Mooshammer (2013), 'Toward Improved Spectral Measures of/s: Results from Adolescents', *Journal of Speech, Language, and Hearing Research*.
15. Ladefoged, Peter and Sandra Ferrari Disner (2012), *Vowels and Consonants*, John Wiley & Sons.
16. Ladefoged, Peter and Ian Maddieson (1996), *The Sounds of the World's Languages*, vol. 1012, Blackwell Oxford.
17. Maniwa, Kazumi, Allard Jongman and Travis Wade (2009), 'Acoustic Characteristics of Clearly Spoken English Fricatives', *The Journal of the Acoustical Society of America* 125(6), 3962–73.
18. Marković, Maja and Bojana Jakovljević (2011), 'Some controversies about/v/ in Serbian, transfer in English', *Exploring English Phonetics* 13.
19. Nissen, Shawn L and Robert Allen FOX (2005), 'Acoustic and Spectral Characteristics of Young Children's Fricative Productions: A Developmental Perspective', *The Journal of the Acoustical Society of America* 118(4), 2570–8.
20. Petrović, Borivoje (2016), 'Jednačenje po zvučnosti opstruenata u proklizi i enklizi u srpskom jeziku', *Radovi Filozofskog fakulteta u Istočnom Sarajevu* 18, 221–36.
21. Petrović, Borivoje (2017), 'F0 Perturbation Effect in Initial Fricatives of Serbian – an Acoustic Analysis', *Filolog* VIII(15), 89–112.
22. Shadle, Christine H and Sheila J Mair (1996), 'Quantifying Spectral Characteristics of Fricatives', In *Proceeding of Fourth International Conference on Spoken Language Processing, ICSLP'96*, 1521–4.
23. Tabain, Marija (1998), 'Non-Sibilant Fricatives in English: Spectral Information above 10 KHz', *Phonetica* 55(3), 107–30.
24. Tabain, Marija (2001), 'Variability in Fricative Production and Spectra: Implications for the Hyper-and Hypo-and Quantal Theories of Speech Production', *Language and speech* 44(1), 57–93.
25. Zwicker, Eberhard (1961), 'Subdivision of the Audible Frequency Range into Critical Bands (Frequenzgruppen)', *The Journal of the Acoustical Society of America* 33(2), 248–248.

Borivoje T. Petrović
Univerzitet u Novom Sadu
Pedagoški fakultet Sombor
Katedra za jezik i književnost

SPEKTRALNI MOMENTI KAO FAKTORI DIFERENCIJACIJE FRIKATIVA U SRPSKOM JEZIKU KOD ŽENSKIH GOVORNIKA

Rezime

Rad ispituje spektralne momente kod frikativa u srpskom jeziku. Četiri spektralna momenta: centar gravitacije, spektralna varijansa, iskošenost i oštrina mereni su na početku, u najstabilnijem delu i na kraju frikativa. Rezultati pokazuju da je centar gravitacije najveći kod nesibilantnih frikativa, osim foneme [v]. Merenja su vršena u dva merna sistema – Hercovoj i Bark skali. Zaključak je da se tip frikativa lakše može odrediti korišćenjem Bark skale. Akustička energija koja se generiše za vreme izgovaranja frikativa nije podjednako raspoređena za sve vreme trajanja izgovora. Vidljiva je razlika između zvučnih i bezvučnih frikativa. Faktori koji najviše utiču na razliku između frikativa u odnosu na spektralne momente su fonološka zvučnost i mesto izgovora foneme. Vokal koji sledi inicijalni frikativ ne utiče znatno na razliku u spektralnim momentima, osim razlike iza frikativa [x], koji se realizuje na akustički različite načine ispred vokala [i] i [u]. Ovaj fenomen je potrebno malo podrobnije ispitati.

▸ **Ključne reči**: centar gravitacije, spektralna varijansa, iskošenost, oštrina, spektralni momenti, frikativi, srpski jezik.

Preuzeto: 3. 5. 2019.
Korekcije: 18. 8. 2019.
Prihvaćeno: 15. 12. 2019.

Appendices

Mean spectral moments of fricatives in Serbian in function of the following vowel

Fricative	Succ. Vowel	COG (Hz)	SD (Hz)	Skew (Hz)	Kurtosis (Hz)	COG (B)	SD (Bark)	Skew (B)	Kurtosis (B)
[f]	[a]	8153	5255	1.667	6.632	20.501	18.399	0.017	0.066
	[e]	8361	5134	1.495	4.758	20.795	18.362	0.015	0.047

	[i]	8540	5208	1.696	7.502	20.882	18.151	0.017	0.074
	[o]	9287	5744	1.275	4.490	21.308	18.920	0.013	0.044
	[u]	8910	5493	1.358	3.947	21.356	18.690	0.013	0.039
[v]	[a]	5359	6799	1.761	5.231	17.605	20.032	0.017	0.052
	[e]	4479	6442	1.846	3.918	16.915	19.846	0.018	0.039
	[i]	5313	6720	1.461	2.140	18.101	20.124	0.014	0.021
	[o]	5346	6756	1.345	1.006	18.386	20.253	0.013	0.010
[s]	[a]	6338	2803	4.232	36.607	19.627	13.548	0.042	0.361
	[e]	6822	2956	3.614	28.549	20.069	14.348	0.036	0.282
	[i]	6275	2179	4.044	38.627	19.460	12.581	0.040	0.382
	[o]	6052	2639	4.260	34.994	19.421	13.650	0.042	0.346
	[u]	5646	2497	4.434	37.266	19.181	13.429	0.044	0.368
[z]	[a]	5479	3745	2.691	16.650	18.565	16.107	0.027	0.165
	[e]	6849	3502	2.964	24.063	20.150	15.297	0.029	0.238
	[i]	6184	3304	2.945	26.966	19.455	14.704	0.029	0.266
	[o]	5534	3189	3.042	23.598	18.731	14.950	0.030	0.233
	[u]	6056	3957	2.737	15.419	19.394	16.365	0.027	0.152
[ʃ]	[a]	5094	2239	2.663	20.984	18.515	12.664	0.026	0.207
	[e]	5495	2337	2.579	21.268	18.697	12.681	0.026	0.210
	[i]	4876	1975	3.143	27.852	18.341	11.988	0.031	0.275
	[o]	5360	2220	2.642	23.750	18.794	12.627	0.026	0.235
	[u]	5471	1949	2.704	27.734	18.997	11.950	0.027	0.274
	Total	5259	2144	2.746	24.318	18.669	12.382	0.027	0.240
/ʒ/	[a]	4489	2837	2.826	18.590	17.843	14.285	0.028	0.184
	[e]	4112	2474	3.306	23.620	17.266	13.699	0.033	0.233
	[i]	4891	2806	2.533	19.169	18.330	14.000	0.025	0.189
	[o]	4400	2702	2.659	17.553	17.628	14.197	0.026	0.173
	[u]	5100	2683	2.378	22.733	18.630	13.728	0.024	0.225
[x]	[a]	6404	5611	1.785	5.344	19.394	18.792	0.018	0.053
	[e]	6531	4043	2.567	14.273	19.466	16.524	0.025	0.141
	[i]	5684	3154	3.532	19.962	19.164	15.223	0.035	0.197
	[o]	8006	6082	1.374	5.552	20.151	19.257	0.014	0.055
	[u]	8626	6983	0.707	2.000	20.429	20.250	0.007	0.020

The list of recorded tokens

san /san/ –- 'dream'

senf /senf/ –- 'mustard'

sir /sir/ –- 'cheese'

sob /sob/ –- 'reindeer'

sup /sup/ –- 'vulture'

zalf /zalf/ –- 'saucer'

zet /zet/ –- 'son/brother-in-law'

zip /zip/ – the name of archiving software

zort /zort/ – 'strength'

zum /zum/ – onomatopoeic exclamation of a quick movement

šav /ʃav/ –- 'stitch'

šeh /ʃex/ –- a move in chess

šik /ʃik/ –- 'chic'

šor /ʃor/ –- 'village street'

šut /ʃut/ –- 'shot, kick'

žal /ʒal/ –- 'mourning'

žeg /ʒeg/ –- a nonsense word

Žid /ʒid/ –- 'a Jew'

žok /ʒok/ – a nonsense word

žup /ʒup/ – a nonsense word

fah /fax/ –- 'pigeonhole'

fes /fes/ –- 'fez'

fil /fil/ –- 'filling, custard'

fond /fond/ –- 'fund'

fuš /fuʃ/ –- 'of poor quality'

Vat /vat/ –- 'Watt'

veš /veʃ/ –- 'laundry'

vic /vit͡s/ –- 'joke'

vođ /vod͡ʑ/ –- alike to 'leader'

half /xalf/ –- archaic term for a player position in football

herc /xert͡s/ –- a suite in cards, 'hearts'

hit /xit/ –- 'a hit'

hop /xop/ –- a sound of jumping

huk /xuk/ –- onomatopoeic of birds, owl, e.g.

Spectrograms of fricatives in the Serbian language (Middle Window) after vowel [a]

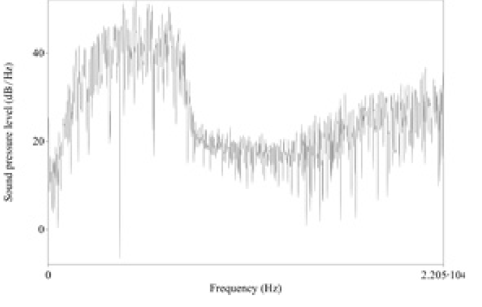
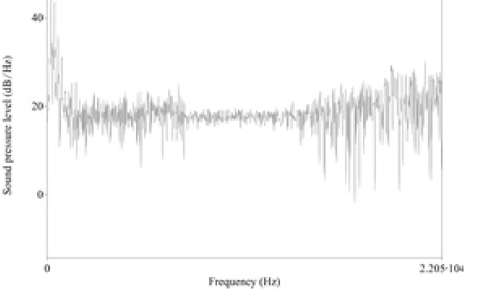
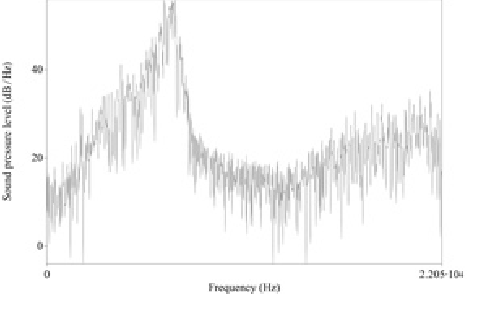


Figure 16: Spectral slice of the fricative [f] in the word 'fah' [pigeon-hole] at the beginning of the fricative noise.





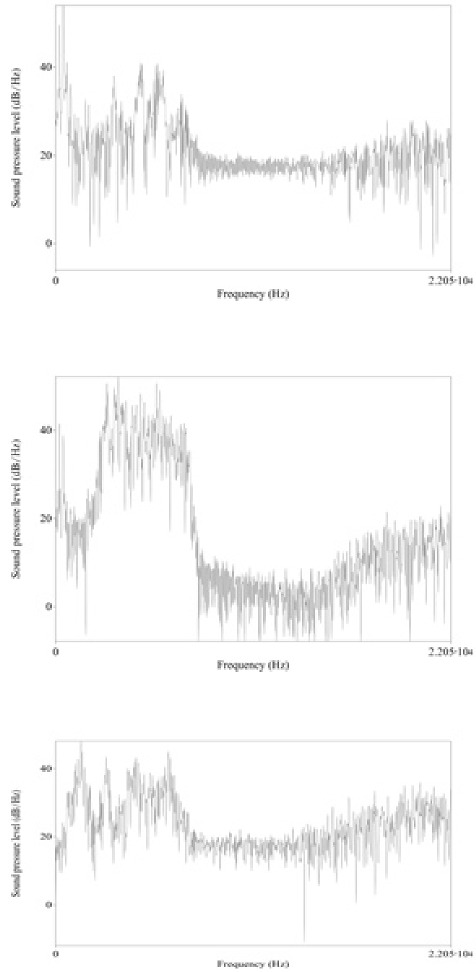


Figure 22: Spectral slice of the fricative /h/ in the word 'half' [a player position in football] at the beginning of the fricative noise.