FACTA UNIVERSITATIS
Series: **Linguistics and Literature** Vol. 18, Nº 2, 2020, pp. 235-256
https://doi.org/10.22190/FULL2002235T

ACOUSTIC ANALYSIS OF PITCH ACCENT AS A REGIONAL FORENSIC MARKER IN SERBIAN

UDC 811.163.41'342.1

Kristina Tomić

University of Niš, Faculty of Philosophy, Serbia

Abstract. *The current research compares the acoustic correlates of pitch accent of two urban dialects of spoken Serbian, urban Niš and Novi Sad. We analyzed spontaneous speech of nineteen native speakers of Serbian with regard to vowel duration and fundamental frequency ratios. It was shown that the tone between the stressed and post-stressed vowel is generally falling in the speech of Niš, while in the speech of Novi Sad it reflects the tone of the expected pitch accent in that particular word. The same is true of the interval between the end of the stressed and the beginning of the post-stressed vowel, which is always rising in the speech of Niš. In Niš, speakers tend to produce vowels in words with falling accents as longer than in words with rising accents. On average, vowels are longer in the speech of Novi Sad. Bearing in mind that regional variations provide important forensic markers (Kašić and Đorđević 2009), this research aims to contribute to the discipline of forensic phonetics, in particular to speaker profiling. Its significance also lies in the fact that it examines spontaneous speech, and is thus relevant for forensic casework (Rose 2002; Nolan, de Jong, and McDougall 2006).*

Key words: *forensic phonetics, speaker profiling, regional markers, pitch accent, fundamental frequency*

1. INTRODUCTION

Standard Serbian belongs to the class of languages with hybrid prosodic systems, those that manipulate both stress and tone (Lehiste and Ivić 1986; Zec and Zsiga 2009). Its prosodic system has two pitch accents, falling and rising. Each accent is defined by a characteristic pitch shape, as well as by stress, whose correlate is an increase in duration (Zec and Zsiga 2009; Subotić, Sredojević, and Bjelaković 2012; Sredojević 2017). The four realizations of Serbian pitch accent are conventionally transcribed with the following symbols above the vowel: short falling [̏], short rising [̀], long falling, [̂] and long

Submitted February 25, 2020; Accepted September 26, 2020
Corresponding author: Kristina Tomić
University of Niš, Faculty of Philosophy
E-mail: kristinatomic89@hotmail.com

rising [ˊ]. In addition, the prosodic system of standard Serbian is said to have two more prosodemes: a post-stressed syllable length [¯] and shortness [˘] (Zec and Zsiga 2009; Subotić, Sredojević, and Bjelaković 2012; Sredojević 2017).

Pitch accent of standard Serbian is based on the prosodic system of New Shtokavian dialects. The classification of Shtokavian speech communities was performed by relying on accentuation and the "yat" reflex (Ivić 1956; Petrović 2009). According to Ivić (1956), Shtokavian speech communities can be divided into those in which progressive (younger) dialects with the new accentuation system are spoken, such as Šumadija-Vojvodina, Eastern Herzegovinian and Younger Ikavian dialect, and archaic (older) dialects with an old accentuation system, such as Kosovo-Resava and Zeta-Sjenica dialect (Ivić 1956; Petrović 2009). In his first classification, Ivić (1956) did not define Balkanized speech of Prizren-Timok as a dialect; instead, he considered it to be an entirely different variety, since speech communities of this region are quite distinct from the rest of Shtokavian speech communities. The distinction is observed both on the structural and on the prosodic level (Ivić 1956). For instance, as opposed to new Shtokavian dialects, in the Prizren-Timok area, there is no pitch accent as such – there is only an expiratory stress, which is the result of the elimination of all quantitative and qualitative differences (Ivić 1956). It was not until later that Prizren-Timok was recognized as a full-fledged Shtokavian dialectological region (Ivić 1988, 68-72, as cited in Petrović 2009, 12).

When speakers from different dialectological backgrounds adopt the standard orthoepy, they tend to realize some linguistic features in a novel way; however, the vestiges of their original articulation base may still be present in their pronunciation. The varieties of language that come into existence in this way are known as urban varieties or urban dialects, and these have different characteristics across regions (Kašić and Đorđević 2009). The reluctance to adapt the prescriptive rules of standard Serbian to language in use often results in discrepancies between the prescribed norm and actual realization of certain linguistic features. For instance, recent research indicates that certain aspects of the standard language, such as post-stressed vowel length, have been reduced or have disappeared entirely (Subotić, Sredojević, and Belaković 2012; Sredojević 2017). Furthermore, the realization of the four pitch accents may vary significantly across regions (Kašić and Đorđević 2009; Lončar Raičević and Sudimac 2017a).

Pronunciation differences which exist across regions are important in forensic linguistic practice (Kašić and Đorđević 2009; Lončar Raičević and Sudimac 2017a). Namely, regional markers are seen as effective discriminants in cases of speaker profiling and language analysis for determination of origin. When dealing with proper representatives of a certain dialect, especially from rural areas, auditory analysis may be enough for successful determination of origin. However, it is not always obvious how speakers from different dialectological regions acquire standard pronunciation, or how their language changes in contact with speakers of other dialects. As Kašić and Đorđević (2009) and Lončar Raičević and Sudimac (2017) point out, the area provides fruitful ground for research into acoustic description of different urban varieties of the Serbian language.

2. THEORETICAL FRAMEWORK

2.1. Forensic Phonetics and Speaker Profiling

Forensic phonetics is a discipline which employs knowledge and methodology of phonetics to legal and forensic purposes. A forensic linguist is usually hired when there is a voice recording of a perpetrator, and their tasks may range from transcription or authentication of recordings, compiling of voice lineups for speaker identification by ear witnesses, to speaker profiling and speaker comparison (Rose 2002; Hollien 2002; 2012; Jessen 2010; French 2017). Speaker profiling refers to a set of procedures performed in forensic cases when there is a recording of a voice but no suspect. Commonly requested by the police during an investigation, it implies the determination of age, sex, region, socio-economic status and foreign language background of the recorded speaker (Rose 2002; Jessen 2010). In Serbian, forensic linguists often rely on regional markers to determine the speaker's origin and areas where they have lived; these include segmental (vowel and consonant quality), as well as prosodic features (pitch accent, tempo, loudness and intonation patterns) (Kašić and Đorđević 2009).

2.2. Previous Acoustic Research on Serbian Accents

Ivić and Lehiste (2002) made the first attempts at acoustic measurements of Shtokavian dialect characteristics, and they published their results in a series of papers from 1963 to 1984. However, research of the prosodic characteristics of modern urban dialects appeared much later.

The most comprehensive study of pitch accent in the urban speech of Novi Sad was conducted by Sredojević (2017), who examined the fundamental frequency movement in the four pitch accent types relative to the position in the intonation phrase (Sredojević 2017). Sredojević confirms that the speakers from Novi Sad use duration (long and short) and tone movement (rising and falling) distinctively. He notes, however, that, in some cases, the realization of rising accents in the urban speech of Novi Sad is different from that described in previous research for standard Serbian (Sredojević 2017, 71). Sredojević (2017, 199, 221) emphasizes that pitch accent in Serbian is realized over the entire syllable and that the pitch movement between the end of the stressed and beginning of the post-stressed vowel is the most indicative of accent type. Bjelaković and Marković (2009) and Dragin (2009) examined the post-stressed vowel length, concluding that the given prosodeme has been significantly reduced in some contexts, while in others it has entirely disappeared (Bjelaković and Marković 2009; Dragin 2009). Both studies provide very few examples where the post-accented syllable length can be observed in the modern language. Similarly, Sredojević (2017) claims that there is no significant difference between the length of vowels in post-stressed length and post-stressed shortness (Sredojević 2017, 154).

As far as the Prizren-Timok dialect is concerned, the first accounts in the literature were based on the auditory analysis (see Belić 1905; Ivić 1956). The pitch accent of the Prizren-Timok dialect was described as lacking any distinction in duration and tone, with the prominence produced using only respiratory effort, which is why this distinction is referred to as the 'expiratory accent' (Ivić 1985, as cited in Lončar Raičević and Sudimac 2017a, 210), or stress (Lehiste 1970). Recently, in a series of research papers, Lončar Raičević and Sudimac (2017a; 2017b; 2018; 2019) have provided acoustic description of the speech of five urban centers of the Prizren-Timok dialectological area. The authors

measured duration, intensity and fundamental frequency of stressed and post-stressed vowels in disyllabic words in the urban speech of Niš, Leskovac, and Vranje (Lončar Raičević and Sudimac 2017a), Svrljig (2017b) and Pirot (2018), concluding that the general tonal movement in realization of all four accent types is falling in all cities (Lončar Raičević and Sudimac 2019, 211). The speech of Niš, Leskovac, Vranje and Pirot is characterized by a slight rising interval between the end of the stressed and beginning of the post-stressed syllable, resembling the short rising accent in standard Serbian (Lončar Raičević and Sudimac 2017a). On the other hand, the speech of Svrljig is characterized by a slight fall in f_0 at this point, which is similar to the short falling accent in the standard language (Lončar Raičević and Sudimac 2017b). The rising interval that appears in the speech of Prizren-Timok urban dialects can be described as region-specific since there is a significant difference in the values of tonal movement between the cities (Lončar Raičević and Sudimac 2017a, 223). As far as the vowel duration is concerned, these authors note that speakers from Niš on average produce longer stressed vowels than speakers from other cities they examined (Lončar Raičević and Sudimac 2018, 435)

To our knowledge, research on pitch accents in Serbian has so far mostly been performed on tokens in frame sentences (see Sredojević 2017; Ivić and Lehiste 2002; Lončar Raičević and Sudimac 2017a; 2017b; 2018) or individual words (Marković and Bjelaković 2009). Bearing in mind that most of the forensic casework involves the analysis of spontaneous speech, there is a growing tendency of using spontaneous speech corpora in forensic phonetic experiments (Rose 2002; Nolan, de Jong, and McDougall 2006). With this in mind, the current research aims to present the pitch accent realization in the given dialects from a new perspective.

3. Research Aims and Methodology

3.1. Research Aims

In their paper on speaker articulation base, Kašić and Đorđević (2009) explain that through education, migration and language contact, speakers of a certain dialect start acquiring standard pronunciation. However, depending on their articulation base, some segmental or prosodic features may be pronounced in a different, novel way. The varieties which come into existence during the adoption of standard pronunciation are known as urban varieties. Kašić and Đorđević (2009) point out that one can easily be misled regarding the origin of a native speaker of Serbian, especially if the speaker has adopted standard accentuation; therefore, they encourage further research on acoustic patterns of realization of the four pitch accents in various urban dialects of Serbia (Kašić and Đorđević 2009).

The goal of the current research is to compare the acoustic correlates of pitch accent of native speakers of two Serbian urban dialects, i.e. of Niš and Novi Sad, and to explore their differences. We will examine the duration and fundamental frequency ratios of stressed and post-stressed vowels in disyllabic words in spontaneous speech. We will test the following hypotheses[1] by performing acoustic and statistical analyses:

[1] The hypotheses presented here are alternative hypotheses (H_1), bearing in mind that the statistical tests performed below are designed in such a way that the null hypohesis (H_0) assumes no difference between the sets of data.

(1) Speakers from Niš distinctively pronounce all four accent types.
(2) Speakers from Niš distinctively pronounce long and short accents.
(3) Speakers from Niš distinctively pronounce rising and falling accents.
(4) Speakers from Novi Sad distinctively pronounce all four accent types.
(5) Speakers from Novi Sad distinctively pronounce long and short accents.
(6) Speakers from Novi Sad distinctively pronounce rising and falling accents.
(7) Speakers from Niš and Novi Sad realize the four pitch accents in Serbian in a different way.

3.2 Participants

The sample for the current research was based on the recordings of spontaneous speech of ten female native speakers of Serbian from Niš and nine from Novi Sad. The speakers have lived in these cities since their birth or a very young age. All of them were either undergraduate students or graduates, aged between 20 and 23 (Niš), and 21 and 28 (Novi Sad) at the time of recording[2]. The average age of speakers from Niš was 21.2 with SD of 1.32, while the average age of those from Novi Sad was 24.78 with SD of 3.23. T-test for two independent means with the score of -3.22421 and p-value of .004981 indicate that there is a significant age difference between the two groups of speakers. Since fundamental frequency of female voice tends to lower with aging (Russel, Penny, and Pemberton 1995; Nishio and Niimi 2008; Eichhorn et al. 2018), it would be methodologically incorrect to compare the absolute values of fundamental frequency of these speakers. Therefore, in this research, we will only compare fundamental frequency ratios and frequency movement expressed in semitones.

3.3. The sample

To elicit spontaneous speech, we created two custom maps that contained 55 unique token words in Serbian, selected from *Dictionary of Serbian language* (Vujanić et al. 2017) in such a way as to ensure the presence of all four pitch accents. The speakers were instructed to give directions on the map and mention as many landmarks as they can on the way. Even though the participants were primed before the experiment, not all of the tokens appeared in the corpus, and some were encountered more than once. Only disyllabic words that appeared in phrase-medial positions were selected for the analysis, which resulted in the total of 314 words, 16.53 per speaker (SD 2.93). Token words in phrase-initial or phrase-final positions were disregarded in our research since the tone of the stressed and unstressed syllable largely depends on the location in the intonation unit (Ivić and Lehiste 2002, 202-20; Sredojević 2017, 46-9, 220).

While it is known that the voicing of flanking consonants may affect the length of the vowels (Lehiste 1970: 19; Sovilj-Nikić 2010: 602; Sredojević 2017: 134), as well as the fundamental frequency in vowels (Ohde 1984; Hanson 2009), in our research, the consonant environment in the token words was not controlled for voicing. Namely, pitch accent in Serbian is a suprasegmental feature realized over a syllable or several syllables; thus, it is supposed to override the intrinsic length of the segments. In addition, as the current research is interested in the dynamics of pitch movement rather than raw f_0 values of the given vowels at given points, it was not considered necessary to control the voicing parameter.

[2] The recording of the corpus was conducted in April and May 2017.

The recording process was performed in an acoustically treated room using *Audacity*, audio software, and *Rode NT-USB* cardioid condenser microphone with a pop-filter. It lasted about 5 minutes per speaker, which resulted in approximately 90 minutes of spontaneous speech. Auditory analysis of the recorded material was performed to select the relevant tokens. Portions of speech pronounced with laughter or hesitation were not taken into consideration.

3.4. Acoustic Analysis

The following variables were measured in the stressed and the post-stressed vowel of each token word using *Wavesurfer* (Sjölander and Beskow 2000):

- Duration (milliseconds)
- Fundamental frequency (f_0) (Hz) at the beginning and the end of the vowel
- Peak f_0 (Hz)
- f_0 peak location (percent)

The measurements of the fundamental frequency were taken into six points in each token word: a) at the beginning of the stressed vowel (A), b) at the peak f_0 of the stressed vowel (B), c) at the end of the stressed vowel (C), d) at the beginning of the post-stressed vowel (D), e) at the peak f_0 of the post stressed vowel (E), and f) at the end of the post stressed vowel (F). The points that mark the beginning and the end of the vowel were chosen to be places where the waveform starts vibrating and the voicing begins. If any distortions were noted at the onset/end of the vowel, the measurements were taken at the next/previous point where the pitch contour is stable.

Table 1 below shows the summary of the parameters derived and compared across the two dialects.

Table 1 The parameters tested

Parameter	Explanation
Vowel duration, str	duration of the stressed vowel
Start-end R str	the ratio of f_0 at the starting and ending point in the stressed vowel
Peak-end R str	the ratio of f_0 at the peak and the ending point in the stressed vowel
Peak f_0 % str	the location of peak f_0 in the stressed syllable in percentage
Vowel duration, p-s	duration of the post-stressed vowel
Start-end R p-s	the ratio of f_0 at the starting and ending point in the post-stressed vowel
Peak-end R p-s	the ratio of f_0 at the peak and the ending point in the post-stressed vowel
Peak f_0 %, p-s	the location of peak f_0 in the post-stressed vowel in percentage
Vowel duration R (%)	the ratio of the post-stressed and the stressed vowel expressed in percentage
f_0 ratio – peaks	the ratio of f_0 at the peak in the stressed and the peak in the post-stressed vowel
f_0 R end (str)-start (p-s)	the ratio of f_0 at the ending point of the stressed and the starting point of the post-stressed vowel
f_0 ratio – starts	the ratio of f_0 at the starting point in the stressed and post-stressed vowel
f_0 ratio – ends	the ratio of f_0 at the ending point in the stressed and post-stressed vowel
Word tone – start-end	the ratio of f_0 at the starting point in the stressed and the ending point in the post-stressed vowel
Peak (str) – end (p-s)	the ratio of f_0 at peak in the stressed and the ending point in the post-stressed vowel
Word pitch range	the ratio of the highest f_0 point and the lowest f_0 point in the word

3.5. Statistical Analysis and Graphical Representation

We performed the statistical analysis in *Microsoft Excel* and *iNZight*, a free data analysis software developed by the Department of Statistics at the University of Auckland (Elliott and Kuper 2017). The statistical significance of the difference in means between the four pitch accents (within one dialect) is measured using One Way ANOVA (F test), while the difference in means between the two dialects is measured with Welch's two-sided t-test, an adaptation of Student's t-test that is more reliable when the samples have unequal variances and unequal sample sizes (Welch 1938).

For the purposes of clear graphical representation and a better understanding of pitch movement in the stressed and post-stressed vowel for every accent type, the distances between the six measured points were expressed in semitones. A semitone (ST) is a twelfth part of an octave, which is a minimal possible distance between two tones and it represents a subjective feeling of difference in pitch height (Jovičić 1999: 248; Sredojević 2017: 33). Namely, even though a listener can perceive the pitch of one sound as "higher" than that of the other, the distance between two frequencies is not always perceived in the same manner. For instance, to a listener, frequency movement from 100Hz to 200Hz is perceived as greater than that from 250Hz to 350Hz (Sredojević 2017: 33). Expressing pitch movement into semitones is a common part of pitch accent research methodology (cf. Sredojević 2017; Lončar Raičević and Sudimac 2017a) as it normalizes f_0 differences between speakers ('t Hart, Collier and Cohen 2006: 24).

4. The Results

In the first two parts of this section, we examine the selected variables in the urban Niš and urban Novi Sad speech, respectively, for each pitch accent type. It is important to note, however, that, throughout the paper, we do not claim each pitch accent type was actually realized in either of the dialects. In fact, the values we provide below are *actual realizations* of dynamic accent for *expected* accent types. The results are presented in tables with group means, statistical test scores and p-values. In addition, there are graphs presenting pitch movement for each expected pitch accent type, expressed in semitones. In the third part of the section, we compare the relevant parameters and pitch movement across the two dialects.

4.1. Pitch Accent Realization in Niš

In Table 2 below, we can observe the mean values of urban Niš speakers for the selected parameters in four expected pitch accents: long falling (L-F), long rising (L-R), short falling (S-F) and short rising (S-R), while Table 3 lists the statistical test results for differences in parameters between long and short and falling and rising accents, respectively. The p-values marked in bold indicate that there is a statistically significant difference in the realization of the expected accent types.

Table 2 Mean values of acoustic parameters of pitch accent for urban Niš speakers with ANOVA (F test), $p < 0.05$

Variable	L-F	L-R	S-F	S-R	F-test	p-value
Vowel duration, str.	0.0982	0.08467	0.09226	0.08461	5.5043	**.001**
f_0 start, str.	232.8	233.6	241.5	228.9	1.1813	.318
f_0 end, str.	215.5	222.4	227.8	217.9	1.2142	.306
f_0 peak, str.	236	242.2	249.3	232.7	1.9856	.117
Peak f_0 %, str	20.88	34.6	28.15	27.43	1.3772	.251
Start-end R (tone), str.	1.084	1.059	1.068	1.057	1.0561	.369
Peak-end R, str.	1.098	1.093	1.098	1.072	1.263	.288
Vowel duration, p-s	0.05887	0.05452	0.04974	0.05384	2.5466	.057
f_0 start, p-s	221.5	239.6	236.4	230.6	2.5441	.057
f_0 end, p-s	203.8	218.3	217.1	207.4	2.4111	.068
f_0 peak, p-s	222.8	241.9	237.1	232.2	2.7683	**.043**
Start-end R p-s	1.088	1.097	1.090	1.115	1.1991	.311
Peak-end R, p-s	1.094	1.108	1.093	1.124	1.7639	.155
Peak f_0 %, p-s	8.95	10.46	7.58	12.49	0.4694	.704
Vowel duration R (%)	60.91	65.91	56.91	66.8	2.2761	.081
f_0 ratio – peaks	1.063	1.012	1.057	1.018	3.9038	**.009**
f_0 R end (str.)-start (p-s)	0.977	0.9366	0.9672	0.9528	2.2152	.088
f_0 ratio - starts	1.0559	0.9924	1.0323	1.0068	3.3514	**.020**
f_0 ratio - ends	1.062	1.026	1.054	1.061	1.3084	.273
Word tone – start - end	1.148	1.085	1.124	1.121	2.2266	.086
Peak (str.) – end (p-s)	1.163	1.120	1.155	1.137	1.3659	.254
Word pitch range	1.176	1.133	1.122	1.139	0.39686	.755

In Tables 2 and 3, we may observe that there is a significant difference in the duration of stressed vowels for speakers from Niš. Namely, vowels in words where falling accents are expected are longer than the ones in words with expected rising accents. However, there is no difference in duration of stressed vowels between the words with expected long and short accents. These words, on the other hand, exhibit the difference in duration of post-stressed vowels, which are longer in words where long accents are expected, $t = 2.852$ (Table 3).

For each expected accent type, the speakers from Niš also exhibit differences in values of peak f_0 in the post-stressed vowel, $F = 2.7683$ (Table 2) as well in the ratios between starting points, $F = 3.3514$, and f_0 peaks, $F = 3.9038$, of the stressed and post-stressed vowel (Table 2). The Words in which the long rising accent is expected have the highest f_0 peak (241.9 Hz), while those with a long falling accent have the lowest f_0 peak in the post-stressed syllable (222.8 Hz). The ratio of the two peaks in falling accents is higher than the ratio of the two peaks in the rising accents, which can also be observed in Figures 1 and 2 below (almost 1 ST distance in falling and close to 0ST distance in rising accents).

Table 3 Difference between words with expected long and short and falling and rising accents, respectively; urban Niš speakers; Welch's t-test, $p < 0.05$

Variable	long - short		falling - rising	
	t-test	p-value	t-test	p-value
Vowel duration, str.	1.3523	.178	3.8345	**.000**
f_0 start, str.	-0.33035	.741	0.97459	.331
f_0 end, str.	-0.80707	.421	-0.18279	.855
f_0 peak, str.	-0.31288	.755	0.488	.626
Peak f_0 %, str	-0.1816	.856	-1.5444	.124
Start-end R (tone), str.	0.85329	.395	1.5697	.118
Peak-end R, str.	1.176	.241	1.3302	.185
Vowel duration, p-s	2.1852	**.030**	0.71152	.478
f_0 start, p-s	-0.68121	.497	-1.6706	.097
f_0 end, p-s	-0.33724	.736	-1.0913	.277
f_0 peak, p-s	-0.59846	.551	-1.8477	.066
Start-end R p-s	-1.0536	.294	-1.5258	.129
Peak-end R, p-s	-0.88846	.376	-2.0469	**.042**
Peak f_0 %, p-s	-0.24087	.810	-1.0189	.310
Vowel duration R (%)	-0.77268	.441	2.6057	**.008**
f_0 ratio – peaks	0.58261	.561	3.3971	**.001**
f_0 R end (str.)-start (p-s)	0.02926	.977	2.369	**.019**
f_0 ratio - starts	0.60833	.544	2.9333	**.004**
f_0 ratio - ends	-0.74101	.460	1.1626	.247
Word tone – start - end	-0.06045	.952	2.0579	**.041**
Peak (str.) – end (p-s)	-0.01792	.986	1.8875	.061
Word pitch range	0.30711	.759	0.533	.595

As presented in Table 3, apart from the difference in duration of the post-stressed vowel, the words for which long accents are expected do not differ from the words for which we expect realization of short accents. However, there are significant differences between the words in which we expect falling and rising accents, and these are:

- the ratio of the peak f_0 and f_0 at the end of the post-stressed vowel, which is higher in words with rising accents;
- the stressed and post-stressed vowel duration ratio, higher in words with falling accents;
- f_0 ratios of peaks and starting points of the stressed and post-stressed vowel, both of which are higher in words where falling accents are expected;
- the overall word tone, also higher in words with falling accents.

Similarly as in previous research (Lončar Raičević and Sudimac 2017a), words under all four expected accent types exhibit a rising interval between the end of the stressed and beginning of the post-stressed syllable. However, the interval in words with expected falling accents is significantly lower (0.5ST and 0.6ST) than in words with rising accents (1.2ST and 0.9ST) (see Table 3 and Figure 1 and 2).

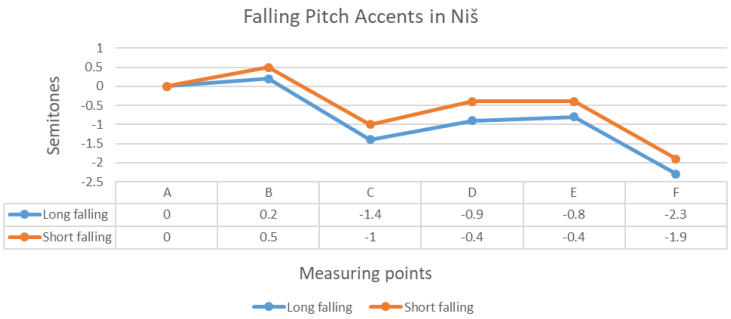

Fig. 1 Realization of falling accents in the urban Niš speech, with pitch movement expressed in semitones

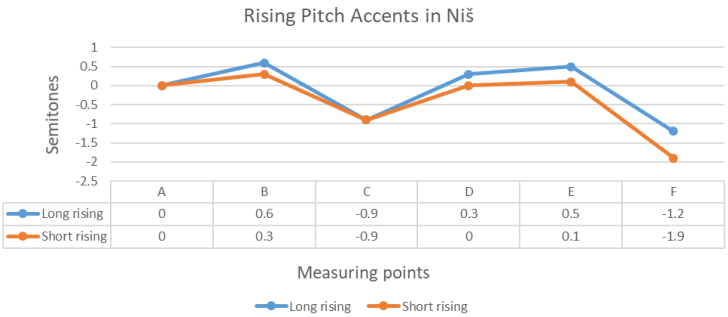

Fig. 2 Realization of rising accents in the urban Niš speech, with pitch movement expressed in semitones

Observing Figures 1 and 2 above, we may note that in Niš, the stressed syllable always has the falling tone (point C is negative relative to the reference point A). The figures also indicate that, in words with rising pitch accents, the post-stressed vowel starts at approximately the same tone as the stressed vowel, while, in falling accents, it always starts slightly lower.

4.2. Pitch Accent Realization in Novi Sad

Table 4 contains the mean values of urban Novi Sad speakers for the selected parameters in four expected pitch accents: long falling (L-F), long rising (L-R), short falling (S-F) and short rising (S-R), while Table 5 encloses the statistical test results for differences in means between long and short and falling and rising accents, respectively. The p-values marked in bold indicate a statistically significant difference.

Table 4 Mean values of acoustic parameters of pitch accent for urban Novi Sad speakers with ANOVA (F test), $p < 0.05$

Variable	L-F	L-R	S-F	S-R	F-test	p-value
Vowel duration, str.	0.1105	0.10524	0.10579	0.09971	1.1467	.332
f_0 start, str.	220.9	190.5	226.5	203	9.1756	**.000**
f_0 end, str.	193.7	180.8	252.0	192.5	23.502	**.000**
f_0 peak, str.	226.6	194.4	270.4	208.4	26.602	**.000**
Peak f_0 %, str	16.49	34.86	59.88	36.57	9.7821	**.000**
Start-end R (tone), str.	1.1553	1.0540	0.9305	1.0611	15.081	**.000**
Peak-end R, str.	1.179	1.075	1.087	1.087	8.8996	**.000**
Vowel duration, p-s	0.05577	0.0616	0.06218	0.05667	1.0479	.373
f_0 start, p-s	186.5	217.6	236.0	224.0	12.91	**.000**
f_0 end, p-s	163.9	192.6	193.3	203.0	13.144	**.000**
f_0 peak, p-s	188.3	220.5	236.5	229.0	12.716	**.000**
Start-end R p-s	1.136	1.132	1.229	1.111	4.8952	**.002**
Peak-end R, p-s	1.147	1.148	1.232	1.131	4.1912	**.007**
Peak f_0 %, p-s	12.1	20.32	5.17	16.52	2.6486	.052
Vowel duration R (%)	51.03	59.15	60.74	58.53	2.849	**.040**
f_0 ratio – peaks	1.2314	0.8955	1.1622	0.9228	29.489	**.000**
f_0 R end (str.)-start (p-s)	1.0603	0.8458	1.0731	0.8674	18.271	**.000**
f_0 ratio - starts	1.2160	0.8922	0.9943	0.9150	22.901	**.000**
f_0 ratio - ends	1.2047	0.9481	1.3188	0.9644	19.917	**.000**
Word tone – start - end	1.3762	0.9996	1.2060	1.0166	23.227	**.000**
Peak (str.) – end (p-s)	1.409	1.018	1.430	1.043	28.966	**.000**
Word pitch range	1.410	1.162	1.445	1.155	15.086	**.000**

As opposed to the speakers from Niš, who do not exhibit any difference across the four pitch accents in the great majority of the tested parameters, the speakers from Novi Sad vary fundamental frequency patterns across pitch accents significantly. However, surprisingly enough, in Novi Sad, there does not seem to be a significant difference in duration of stressed and post-stressed vowels in short and long pitch accents. Bearing in mind that the quality of the vowel may affect its length (Marković and Bjelaković 2009; Sredojević 2017), we performed the duration comparison for the five vowels in Serbian /a/, /e/, /i/, /o/ and /u/, respectively, obtaining the same results.

Words with long and short accents exhibit the difference in the following parameters:

- f_0 of the peak and ending point of the stressed vowel, namely, short vowels have higher average f_0;
- Location of the peak f_0 in percentage, which is further from the beginning for short accents;
- f_0 of the start, peak, and ending point of the post-stressed vowel, which is on average higher in short accents;
- Frequency ratios of the stressed vowel, from the beginning to the end and from the peak f_0 to the end, as well as the ratio between the beginnings of the stressed and post-stressed vowel are higher for the words in which long accents are expected;
- Overall word tone, which is sharper in words with long accents.

Table 5 Difference between words with expected long and short and falling and rising accents, respectively; urban Novi Sad speakers; Welch's t-test, $p < 0.05$

Variable	long - short		falling - rising	
	t-test	p-value	t-test	p-value
Vowel duration, str.	1.2863	.202	1.4408	.153
f_0 start, str.	-0.74709	.457	5.2074	**.000**
f_0 end, str.	-4.483	**.000**	4.1532	**.000**
f_0 peak, str.	-3.3577	**.001**	6.6665	**.000**
Peak f_0 %, str	-3.9641	**.000**	-0.6655	.507
Start-end R (tone), str.	4.6349	**.000**	0.9279	.355
Peak-end R, str.	3.1492	**.002**	3.9954	**.000**
Vowel duration, p-s	-0.612	.542	-0.3787	.706
f_0 start, p-s	-4.9043	**.000**	-2.5438	**.012**
f_0 end, p-s	-4.4096	**.000**	-4.245	**.000**
f_0 peak, p-s	-4.8765	**.000**	-2.8449	**.005**
Start-end R p-s	-1.5862	.116	2.0798	**.040**
Peak-end R, p-s	-1.6423	.104	1.8253	.070
Peak f_0 %, p-s	1.1555	.250	-2.1142	**.038**
Vowel duration R (%)	1.8303	.070	2.3403	**.021**
f_0 ratio – peaks	1.9789	.049	10.217	**.000**
f_0 R end (str.)-start (p-s)	0.50775	.613	8.08	**.000**
f_0 ratio - starts	4.1637	**.000**	7.2477	**.000**
f_0 ratio - ends	-0.63419	.527	8.4143	**.000**
Word tone – start - end	3.1083	**.002**	8.6679	**.000**
Peak (str.) – end (p-s)	0.69602	.488	10.883	**.000**
Word pitch range	0.48514	.629	8.2682	**.000**

Words in which falling and rising accents are expected exhibit differences in almost all of the examined parameters, as follows:

- In the stressed vowel, f_0 is higher at the beginning, peak and at the end in falling accents, and the frequency ratio of the peak f_0 and vowel end is more prominent.
- In the post-stressed vowel, f_0 at the beginning, peak and end in falling accents is lower than in rising accents.
- Falling accents have higher frequency ratios in the stressed vowel than rising accents.
- In rising accents, the peak f_0 of the post-stressed vowel is further from the beginning than in the falling accents
- The interval between the end of the stressed and the start of the post-stressed vowel is falling in words with falling accents and rising in words with rising accents. The same applies to the intervals between the beginnings, peaks and ends of the stressed and post-stressed vowel.
- Vowel duration ratio of the stressed and post-stressed vowel is higher in words with falling accents.

Figures 3 and 4 below illustrate the relationship between the pitch movement of short and long accents in urban Novi Sad speakers expressed in semitones.

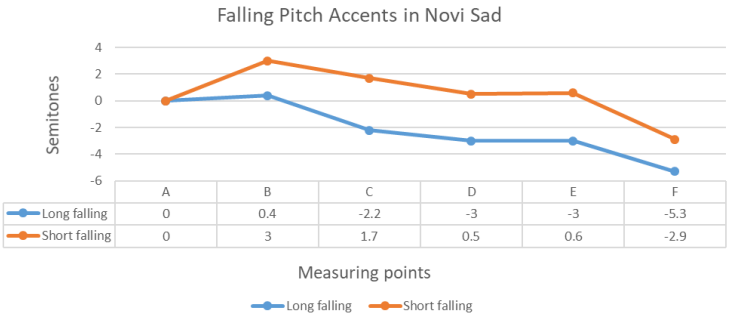

Fig. 3 Realization of falling accents in the urban Novi Sad speech, with pitch movement expressed in semitones

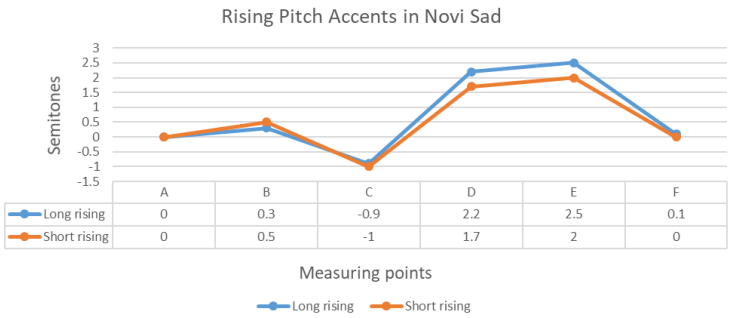

Fig. 4 Realization of rising accents in the urban Novi Sad speech, with pitch movement expressed in semitones

Observing the values in Tables 4 and 5, we may note that falling accents in the speech of Novi Sad start with a higher fundamental frequency than rising accents. The word f_0 peak is in the stressed syllables for words with falling accents, while for words with rising accents, it is in the post-stressed syllable. Words under falling accents exhibit a prominent fall between the two peaks (-3.4ST and -3.6ST), while words with rising accents exhibit a significant rise (2.2ST and 1.5ST). Finally, while long and short falling accent have remarkably different pitch shapes, the greatest difference being in the peak f_0 (0.4ST in the long falling and 3ST in the short falling accent), the contours of rising accents appear to be quite similar. We performed a statistical comparison of the means of all of the acoustic parameters tested above in words with expected long rising and short rising accent in the speech of Novi Sad sample. However, none of the parameters exhibited a statistical difference. This leads us to the conclusion that our speakers did not realize both rising pitch accents as described in the literature.

4.3. Comparing Pitch Accent Realizations in Niš and Novi Sad

We compared the duration of vowels in each expected accent type, and confirmed that the stressed vowel is significantly longer in Novi Sad than in Niš (Figure 5), while the difference in the post-stressed vowel is notable only in words for which the short falling accent is expected (Figure 6). In Novi Sad, the post-stressed vowel in the long falling accent is twice as short as the stressed vowel (51.03%), while its duration in Niš is 60.91% of the stressed vowel. Stressed – post-stressed vowel ratios in other pitch accents do not differ significantly between Niš and Novi Sad.

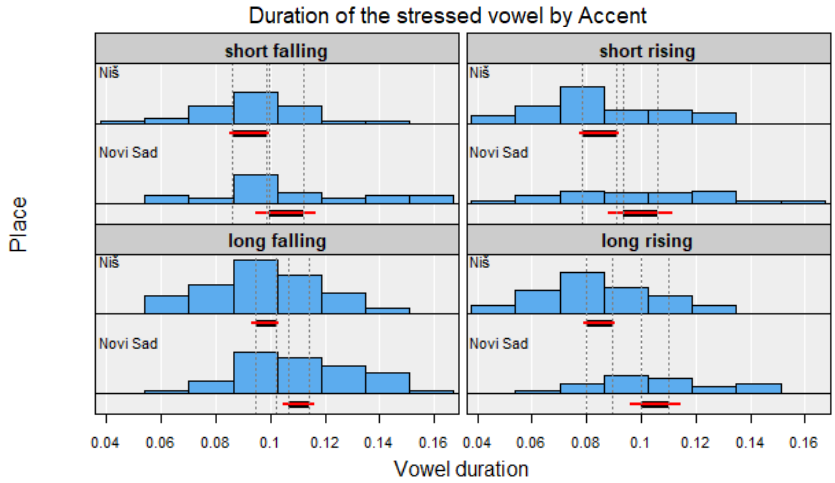

Fig. 5 Duration of the stressed vowel in four expected pitch accents in Niš and Novi Sad

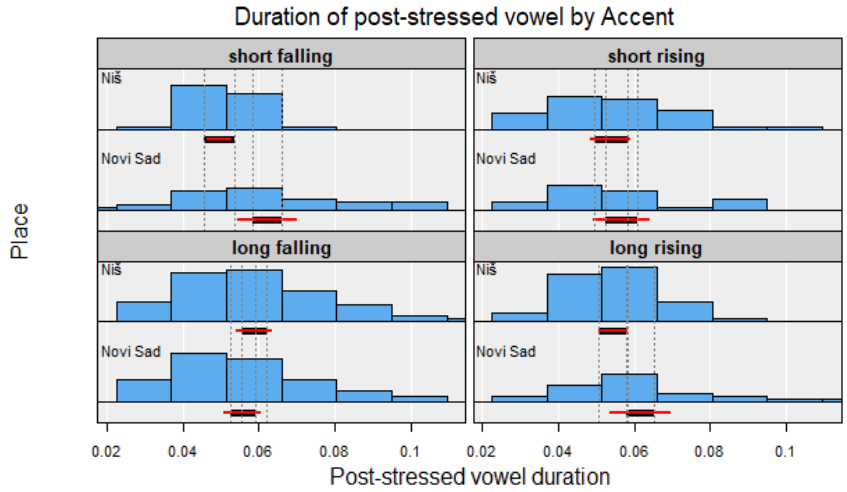

Fig. 6 Duration of the post-stressed vowel in four pitch accents in Niš and Novi Sad

Table 6 below summarizes Welch's t-test scores for the difference in means between the speakers of Niš and Novi Sad (NS) urban dialects in the four expected pitch accents. The p-values marked in bold indicate that the difference between the means is statistically significant. It should be noted, however, that due to the established age difference in the two samples, the absolute fundamental frequency values of the two groups of speakers were not compared.

Table 6 Welch's t-test scores for difference in means between the speakers from Niš and Novi Sad, significant for $p < 0.05$

Variables	Long-falling		Short-falling		Long-rising		Short-rising	
	t-test	p	t-test	p	t-test	p	t-test	p
Vowel duration, str.	-3.2218	**.001**	-2.1083	**.040**	-3.8733	**.000**	-2.2479	**.030**
Start-end R, str.	-3.1643	**.002**	3.173	**.002**	0.2296	.819	-0.2080	.836
Peak-end R, str.	-4.3898	**.000**	0.3962	.693	1.037	.304	-0.9965	.324
Vowel duration, p-s	0.9145	.362	-3.0047	**.004**	-1.641	.109	-0.6213	.537
Start-end R, p-s	-2.8859	**.004**	-4.0904	**.000**	-1.3503	.184	0.1735	.863
Peak-end R, p-s	-3.2015	**.001**	-4.1457	**.000**	-1.7044	.096	-0.3465	.730
Vowel duration R (%)	3.3043	**.001**	-0.8562	.395	1.7478	.086	1.4906	.141
f_0 ratio – peaks	-5.332	**.000**	-3.1684	**.003**	3.6992	**.000**	2.8816	**.006**
f_0 R end (str)-start (p-s)	-2.9145	**.004**	-4.4367	**.000**	3.287	**.002**	2.9181	**.006**
f_0 ratio - starts	-4.8687	**.000**	0.8232	.415	2.8342	**.006**	3.1667	**.002**
f_0 ratio - ends	-3.8772	**.000**	-5.6062	**.000**	3.0844	**.003**	2.488	**.018**
Word tone – start-end	-5.9465	**.000**	-1.5533	.128	2.5977	**.012**	2.7597	**.008**
Peak (str.) – end (p-s)	-6.5034	**.000**	-4.8991	**.000**	3.333	**.001**	2.5001	**.016**
Word pitch range	-6.3053	**.000**	-5.2781	**.000**	-0.145	.885	0.5308	.598

The results indicate that the two dialects exhibit the greatest difference in the realization of the long falling accent, followed by the short falling accent. Long rising and short rising accent differ in the same parameters in the two groups of speakers, which is expected because the recorded group of speakers from Novi Sad did not exhibit any difference in production of long rising and short rising pitch accent.

Figures 7, 8, 9 and 10[3] below compare the pitch movement for every accent type in the two cities.

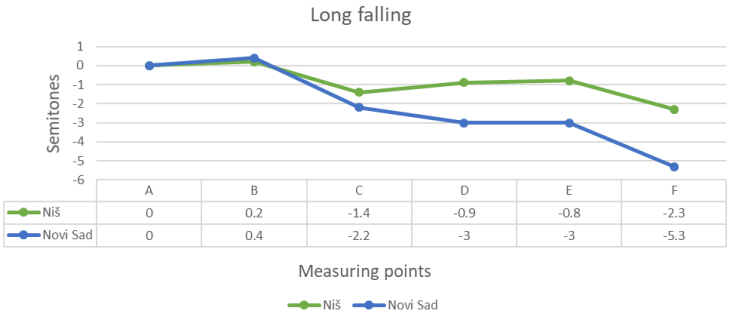

Fig. 7 Pitch movement in long falling pitch accent in the urban Niš and urban Novi Sad speech expressed in semitones

[3] NB The displays below are the same as those in Figures 1 – 4, with differently grouped data.

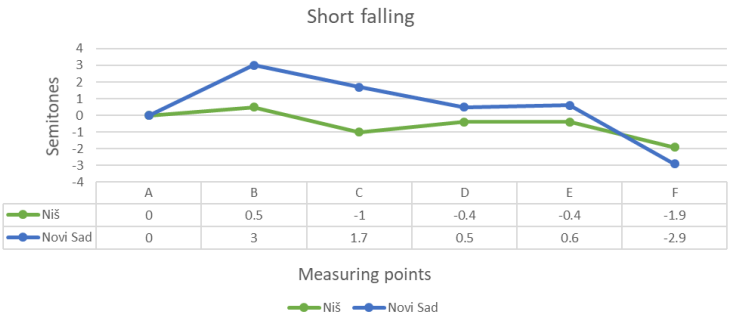

Fig. 8 Pitch movement in short falling pitch accent in the urban Niš and urban Novi Sad speech expressed in semitones

In the long falling accent in Novi Sad, the falling trend between the stressed vowel's starting and ending point as well as that of the peak f_0 and ending point is much more prominent than in Niš (-2.2ST as opposed to -1.4ST). The same is true of the post-stressed vowel (see Figure 7). Moreover, word pitch range is wider in Novi Sad than in Niš (5.7ST vs. 2.5ST).

In words with the expected short falling accent, in Novi Sad, there is a rise of 3 ST in the stressed vowel, while, in Niš, the increase in pitch in the short accent is less than 1ST and it does not differ significantly compared to its long counterpart. Moreover, even though, the tone starts to fall before the end of the stressed vowel in the spoken language of both urban areas, the ending point of the stressed vowel in Novi Sad is on average 1.7ST higher than the onset, while in Niš, it is one semitone lower than the reference point.

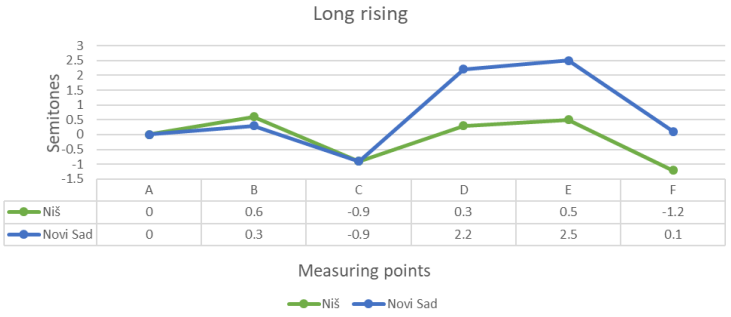

Fig. 9 Pitch movement in long rising pitch accent in the urban Niš and urban Novi Sad speech expressed in semitones

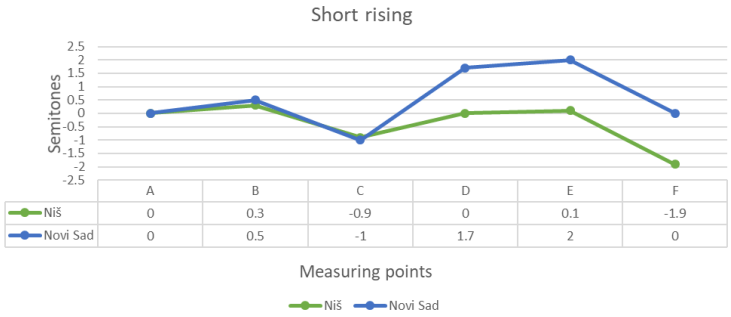

Fig. 10 Pitch movement in short rising pitch accent in the urban Niš and urban Novi Sad speech expressed in semitones

In the long rising accent in Niš, the two peaks are at approximately the same height, while in Novi Sad, there is a rise of 2.2ST. The most notable difference is in the rising interval between the stressed and the post-stressed vowel, which is 1.2ST in the speech of Niš and 3.1ST in the speech of Novi Sad. The word peak is 2.5ST higher than the stressed vowel onset in the speech of Novi Sad, while in Niš, the difference in pitch between these two points is only 0.5ST.

Finally, the pitch movement in the short rising pitch accent is very similar to the movement in the long rising accent. The speakers from Novi Sad produce a sharp rise in the interval between the stressed and the post-stressed syllable (2.7ST) in comparison to the speakers from Niš (0.9ST). For these speakers, the ending point of the post-stressed vowel occurs higher than in the stressed vowel, while speakers from Niš produce a falling tone between these two points.

5. Discussion

Comparing the results from sections 4.1 4.2 and 4.3, we may summarize the differences between the two groups of speakers with regard to production of Serbian pitch accents. Stressed vowels are typically longer in the speech of Novi Sad (106.57 ms, SD 2.45) than in Niš (90.74, SD 2.11), regardless of the expected accent type. When the duration of post-stressed vowels is concerned, the two groups of speakers differ only in words in which the short falling accent is expected. Furthermore, the speakers from Niš produce the stressed vowel in falling accents as longer than in rising, while the speakers from Novi Sad do not make such a distinction. In both cities, the duration ratio of the stressed and post-stressed vowel is higher in falling than in rising accents. In rising accents in Niš, the ratio of the peak f_0 and end of the post-stressed vowel is higher than in falling accents, while, in Novi Sad, both rising and falling accents have a similar frequency ratio of these two points. Most importantly, in Niš, the interval between the beginnings, peaks and ending points of the stressed and post-stressed vowel is falling regardless of which accent type is anticipated for that word; on the other hand, in Novi Sad, words under rising accents have a rising interval, while words under falling accents have a falling interval between these points. Similarly, speakers from Novi Sad produce a rising interval in words with rising accents and a falling interval in words with falling

accents between the ending point of the stressed and starting point of the post-stressed vowel, whereas, the speakers from Niš always produce the rising interval here. It should be noted, however, that there is a significant difference in the frequency ratios between the two points depending on the expected accent tone.

The results that we obtained for speakers of Prizren-Timok dialect from the urban area of Niš confirm some of the previous findings related to production of pitch accents in this city. In their series of studies, Lončar Raičević and Sudimac (2017a, 223; 2018, 434) recorded the rising interval that appears between the ending point of the stressed and the starting point of the post-stressed vowel in all four pitch accent patterns. These authors refer to this interval as parameter R and conclude that it is region specific as it is present in the speech of Niš, Leskovac, Vranje and Pirot, urban centers of the Prizren-Timok dialectological area (Lončar Raičević and Sudimac 2019, 211). In addition, they point out that there is a significant difference in the realization of this interval between the four groups of speakers, as well as that the rising tone is not recorded in the speech of Svrljig, a city that also belongs to the Prizren-Timok dialectological area (Lončar Raičević and Sudimac 2017a, 224; 2018, 434). With regard to vowel duration of the speakers from Niš, in our research, the average value is 90.74 ms (SD 2.11), while the mean length of vowels in the corpus of Lončar Raičević and Sudimac (2017a, 215) is 110 milliseconds. The discrepancy may be attributed to the difference in the nature of corpora used for the quoted research. Namely, the current study is performed on the sample of spontaneous speech while the aforementioned analysis of pitch accents was conducted on tokens in carrier sentences.

When the speakers of Šumadija-Vojvodina dialect from the urban area of Novi Sad are concerned, we confirmed that they produce a distinctive word tone depending on the expected accent type. There is a falling interval between the ending point of the stressed and starting point of the post-stressed vowel in words in which a falling accent is expected, while the words in which a rising accent is expected exhibit a rising interval between these two points (Ivić and Lehiste 2002, 90-104; Sredojević 2017: 223). The same is true of the tone between the starting points, peaks and ends of the stressed and post-stressed vowel. With regard to vowel duration, earlier studies of standard Serbian (Ivić and Lehiste, 2002, 31, 36) and the speech of Novi Sad (Marković and Bjelaković 2009, 152; Sredojević 2017, 139) clearly suggest a difference between the duration of stressed vowels in words with expected long and words with expected short accents. In the current research, however, such distinction was not confirmed. As mentioned before, the difference in the results may stem from the difference in the corpus (spontaneous speech versus individual words or carrier sentences) or the nature of the task itself (giving directions on the map). Bearing in mind that it was already shown that vowel duration is the only parameter that distinguishes long rising and short rising accents while all of the acoustic parameters remain quite similar (Sredojević 2017, 68-69), the lack of quantitative contrast renders all of the rising accents produced by our speakers as the same. It is noteworthy, however, that these speakers do produce a proper rising accent, as it is described in previous literature, with word f_0 peak in the second syllable (Sredojević 2017, 63-66)

6. CONCLUSION

In this paper, we compared the acoustic correlates of pitch accent in spontaneous speech in the urban centers of two Serbian dialectological areas: Prizren-Timok and Šumadija-Vojvodina. The purpose of the research was to investigate the realization of four pitch accents in spontaneous speech and establish acoustic parameters which differ across dialects, and could be used as regional markers in forensic speaker profiling.

The results have shown that the speakers from Niš do not produce the four pitch accent types as they are described in standard Serbian, which confirmed the null hypotheses in (1) and (2). Nevertheless, this group of speaker does include a qualitative difference in tone between words in which a falling accent and words in which a rising accent is expected, which provides evidence for the alternative hypothesis in (3). Speakers from Novi Sad, on the other hand, produce long falling, short falling and rising accents distinctively. Namely, since the two rising accents in the speech of Novi Sad differ only in the quantitative component (Sredojević 2017, 68-9), the absence of this component in our results implies that words in which the long rising and words in which the short rising pitch accent is expected are produced in the same manner. This implies that the alternative hypotheses in (4) and (5) were not confirmed.

On the other hand, there is a lot of evidence to confirm the alternative hypothesis in (7). The most prominent difference between the two urban varieties is reflected in the tone between the stressed and post-stressed vowel, which is generally falling in the speech of Niš, while in the speech of Novi Sad it reflects the tone of the expected pitch accent in that particular word. In addition, even though there is a significant difference in realization of falling and rising accents, speakers from Niš always produce a rising interval between the end of the stressed and beginning of the post-stressed vowel, while speakers from Novi Sad produce a falling interval in words in which the falling accent is expected. Vowels are on average produced as longer in the speech of Novi Sad and speakers from Niš tend to produce vowels in words in which falling accents are expected as longer than vowels in words in which rising accents are expected, which is not the case for the speakers from Novi Sad. Such specificities of pronunciation may be termed region-specific and could be of use in forensic linguistic investigations.

This research presents the basis for the future study that is going to deal with cross-language accent analysis, that is, determination of native dialect when the speakers are using English as a foreign language. Further research should also focus on comparison of pitch accent realization in three-syllabic and four-syllabic words in spontaneous speech for the given urban dialects. In addition, the robustness of the current parameters needs to be tested under various circumstances and with speakers of different social backgrounds and from different dialectological areas. When dealing with regional differences, another important aspect is how they are perceived by naïve listeners. Therefore, future research should also include perception experiments that would test whether naïve listeners can infer the dialectological background of the given speakers based on their pitch accent realization. Such research would require careful planning to avoid the influence of overall intonation, vowel quality, choice of vocabulary and other linguistic aspects.

As for the current paper, it can be said that the results we obtained present a possible realization of Serbian pitch accent by young educated speakers from major urban centers of two different dialectological areas. However, it is important to note some limitations of this study. Namely, bearing in mind the number of participants, it would be presumptuous

to form any generalizations. In addition, even though the fundamental frequency at the onset of vowels could be affected by microprosodic perturbations due to the consonant environment under all pitch accents (Sredojević 2017: 218), the current paper did not take the voicing of the flanking consonants into consideration. It is also worth mentioning that none of the dialectological areas in Serbia are homogenous and that representatives from other cities may have different realizations of pitch accents (Lončar Raičević and Sudimac 2017a; 2019). Using the corpus of spontaneous speech offered a new perspective on what was known to be a staple in pronunciation of speakers from Šumadija-Vojvodina dialectological area - the quantitative component in vowels. For instance, relying only on vowel duration in the speech in our corpus, someone could be led to believe that the speakers may originally be from a different dialectological area, one in which the quantitative distinction between vowels is usually not realized and that the lack of this distinction is a vestige of their original articulation base. However, what contributed to these results may well be the nature of spontaneous speech, the speaking task the participants were exposed to or even some extralinguistic parameters such as anxiety and insecurity due to participation in the experiment. This reminds us that no forensic linguistic analysis, including speaker profiling, can be conducted by relying on a single variable. Instead, the combination of parameters is required in order to provide a clear picture of someone's origin.

Acknowledgement: *The author would like to thank to Special School with Student Dormitory "Bubanj" for providing the sound booth where materials were recorded, the anonymous speakers from Niš and Novi Sad who agreed to participate in the experiment as well as to two anonymous reviewers and the editor on their valuable comments and advice. All of the mistakes remain my own responsibility.*

REFERENCES

Belić, Aleksandar. 1905. "Dijalekti istočne i južne Srbije." *Srpski dijalektološki zbornik I.*

Bjelaković, Isidora, and Maja Marković. 2009. "Posleakcenatska dužina u govoru Novog Sada." In *Govor Novog Sada. Sveska 1, Fonetkse osobine*, 110-28. Novi Sad: Filozofski fakultet, Odsek za srpski jezik i lingvistiku.

Dragin, Gordana. 2009. "Distribucija posleakcenatskih dužina u jeziku medija." In *Govor Novog Sada. Sveska 1, Fonetske osobine*, 129-40. Novi Sad: Filozofski fakultet, Odsek za srpski jezik i lingvistiku.

Eichhorn, Julie Traub, Raymond D Kent, Diane Austin, and Houri K Vorperian. 2018. "Effects of Aging on Vocal Fundamental Frequency and Vowel Formants in Men and Women." *Journal of Voice* 32 (5). doi:10.1016/j.jvoice.2017.08.003.

Elliott, Tom, and Marco Kuper. 2017. *iNZight: iNZight GUI for Data Exploration and Visualisation (version 3.1.1) {Computer program}*. Prod. University of Auckland. Auckland, NZ. Accessed August 16th, 2018. https://www.stat.auckland.ac.nz/~wild/iNZight/install/.

French, Peter. 2017. "A developmental history of forensic speaker comparison in the UK." In *English Phonetics*, 271-86. Accessed October 5th, 2018. http://eprints.whiterose.ac.uk/117763/.

Hanson, M Helen. 2009. "Effects of obstruent consonants on fundamental frequency at vowel onset in English." *The Journal of the Acoustical Society of America* 125 (1): 425-441. doi: 10.1121/1.3021306.

Hollien, Harry. 2012. "About Forensic Phonetics." *Linguistica.* Ljubljana: Tiskana. 27-53. Accessed September 1st, 2018. doi:10.4312/linguistica.52.1.27-53.

Hollien, Harry. 2002. *Forensic Voice Identification.* London and New York: Academic Press.

Ivić, Pavle. 1956. *Dijalektologija srpskohrvatskog jezika. Uvod u štokavsko narečje.* Novi Sad: Matica Srpska.

Ivić, Pavle. 1988. "Narečja." In *Jezik, srpskohrvatski/hrvatskosprski, hrvatski ili srpski*, by Dalibor Brozović and Pavle Ivič, 54-80. Zagreb: JLZ „Miroslav Krleža".

Ivić, Pavle, and Ilse Lehiste. 1963. „Prilozi ispitivanju fonetske i fonološke prirode akcenata u savremenom srpskohrvatskom književnom jeziku." *Зборник Матице српске за филологију и лингвистику* VI: 39-68.

Ivić, Pavle, and Ilse Lehiste. 2002. *O srpskohrvatskim akcentima.* Edited by Dragoljub Petrović. Novi Sad: Izdavačka knjižarnica Zorana Stojanovića.

Jessen, Michael. 2010. "The forensic phonetician: Forensic speaker identification by experts." In *The Routledge Handbook of Forensic Linguistics*, edited by Malcolm Coulthard and Alison Johnson, 378-94. Abingdon and New York: Routledge.

Jovičić, Slobodan T. 1999. *Govorna komunikacija. Fiziologija, psihoakustika i percepcija*. Beograd: Nauka.

Kašić, Zorka, and Jelena P Đorđević. 2009. „Ostaci automatizma artikulacione baze kao forenzički markeri." *ETRAN 2009 Zbornik Radova.* Vrnjačka Banja.

Lehiste, Ilse. 1970, 1977. *Suprasegmentals.* Second. Cambridge, Massachusetts and London: The MIT Press.

Lehiste, Ilse, and Pavle Ivić. 1986. *Word and Sentence Prosody in Serbocroatian.* Cambridge, MA: MIT Press.

Lončar Raičević, Aleksandra, and Nina Sudimac. 2017a. „Akustički opis naglaska u govorima Prizrensko-južnomoravskog dijalekta." *Zbornik Matice srpske za filologiju i lingvistiku*, 209-25.

Lončar Raičević, Aleksandra, and Nina Sudimac. 2017b. "Acoustic Parameters of Speech as Forensic Markers for Speakers of Prizren-Timok Dialect." *Proceedings of the International Scientific Conference "Archibald Reiss Days".* Belgrade: Academy of Criminalistic and Police Studies.

Lončar Raičević, Aleksandra, and Nina Sudimac. 2018. „Akustički opis naglaska u govorima timočko-lužničkog dijalekta." *Philologia Mediana* 10: 423-37.

Lončar Raičević, Aleksandra, and Nina Sudimac. 2019. „Belićevi pogledi na prirodu ekspiratornog akcenta iz današnje perspektive." *Philologia Mediana* 11: 203-14.

Nishio, Masaki, and Seiji Niimi. 2008. "Changes in speaking fundamental frequency characteristics with aging." *Folia phoniatrica et logopaedica* 60 (3): 120-27. doi: 10.1159/000118510.

Nolan, Francis, Gia de Jong, and Kirsty McDougall. 2006. "Introducing the DyViS project: "Dynamic variability in speech: a forensic phonetic study of British English'." *Paper presented at the British Association of Academic Phoneticians Colloquium, Edinburgh, 10-12 April 2006.*

Ohde, N Ralph. 1984. "Fundamental frequency as an acoustic correlate of stop consonant voicing." *The Journal of the Acoustical Society of America* 75 (1): 224-230. doi: 10.1121/1.390399.

Petrović, Dragoljub. 2009. „Pavle Ivić i srpska dijalektologija." *Zbornik Matice srpske za filologiju i lingvistiku*, 9-28.

Rose, Philip. 2002. *Forensic Speaker Identification.* London and New York: Taylor & Francis.

Russel, A, L Penny, and C Pemberton. 1995. "Speaking fundamental frequency changes over time in women: a longitudinal study." *Journal of Speech and Hearing Research* 38 (1): 101-9. doi: 10.1044/jshr.3801.101.

Sjölander, Kåre, and Jonas Beskow. 2000. "Wafesurfer - An Open Source Speech Tool." *A peper presented at The International Conference on Spoken Language Processing*. Accessed on February 1st, 2020. https://www.speech.kth.se/wavesurfer/wsurf_icslp00.pdf

Sovlj Nikić, Sandra. 2010. „Trajanje glasova i faktori koji ga određuju u srpskom i drugim jezicima." *18. Telekomunikacioni forum TELFOR 2010, Beograd, Srbija*. 602-605. Accessed on August 18th, 2018. http://2010.telfor.rs/files/radovi/TELFOR2010_05_09.pdf

Sredojević, Dejan. 2017. *Fonetsko-fonološki opis akcenata u standardnom srpskom jeziku: od specifičnog ka opštem.* Novi Sad: Sajnos.

Subotić, Ljiljana, Dejan Sredojević, and Isidora Bjelaković. 2012. *Fonetika i fonologija: ortoepska i ortografska norma standardnog srpskog jezika*. Novi Sad: Filozofski fakultet Novi Sad.

't Hart, Johan, René Collier, i Antonie Cohen. 2006. *A Perceptual Study of Intonation: an Experimental-Ponetic Approach to Speech Melody*. Cambridge: Cambridge University Press.

Vujanić, Milica, Darinka Gortan-Premk, Milorad Dešić, Rajna Dragićević, Miroslav Nikolić, Ljiljana Nogo, Vasa Pavković, et al. 2017. *Rečnik srpskoga jezika*. Novi Sad: Matica srpska.

Welch, Bernard L. 1947. "The generalization of Student's problem when several different population variances are involved." *Biometrika* 34 (1): 28-35. Accessed September 1st, 2018. doi:10.2307/2332510.

Zec, Draga, and Elizabeth Zsiga. 2009. "Interaction of Tone and Stress in Standard Serbian: The Cornell Meeting 2008." Edited by Wayles Browne, Adam Cooper, Alison Fisher, Esra Kesici, Nikola Predolac and Draga Zec, 537-55. Ann Arbor: Michigan Slavic Publications.

AKUSTIČKA ANALIZA TONSKOG AKCENTA KAO REGIONALNI FORENZIČKI MARKER U SRPSKOM

U ovom istraživanju upoređuju se akustički korelati tonskog akcenta dva urbana dijalekta u srpskom jeziku, i to urbanog govora Niša i Novog Sada. Devetnaest izvornih govornica srpskog snimljeno je u spontanom govoru i analizirani su trajanje vokala u naglašenom i nenaglašenom slogu, kao i odnosi frekvencija osnovnog tona. Pokazali smo da je ton između naglašenog i nenaglašenog vokala uvek u opadanju u govoru Niša, dok u govoru Novog Sada ovaj ton odslikava ton očekivanog akcenta u toj reči. Isto važi i za interval između kraja naglašenog i početka nenaglašenog vokala koji je uvek uzlazan kod govornica iz Niša. U Nišu se naglašeni vokali u rečima gde se očekuje silazni akcenat izgovaraju kao duži u odnosu na reči gde se očekuje uzlazni akcenat. U proseku, naglašeni vokali su duži u govoru Novog Sada. Imajući u vidu da regionalne varijacije u govoru mogu biti važni forenzički markeri (Kašić i Đorđević 2009), ovo istraživanje može da doprinese disciplini kao što je forenzička fonetika, naročito forenzičkom proflisanju govornika. Značaj ovog istraživanja ogleda se i u tome što ispituje akcenatski ton u srpskom na korpusu spontanog govora. Kao takvo, može biti od važnosti za istrage i pravne postupke gde je potrebna pomoć forenzičara lingviste.

Ključne reči: *forenzička fonetika, profilisanje govornika, regionalni markeri, tonski akcenat, frekvencija osnovnog tona*