Original research article
811.111’342
811.111’344
https://doi.org/10.18485/bells.2020.12.3

Tatjana Paunović*
University of Niš
Faculty of Philosophy
Niš, Serbia

EFL STUDENTS’ SPONTANEOUS SPEECH: TONALITY, TONICITY, AND TONE**

Abstract
The present study offers a quantitative and acoustic analysis of a corpus of spontaneous speech produced by Serbian EFL students, in terms of intonation unit organization (tonality), nucleus placement (tonicity), and the realization of nuclear pitch contours (tone). The results show that, regarding tonality, the participants’ spontaneous speech was characterized by numerous interruptions and hesitations, with comparatively few complete IUs. Regarding tonicity, the nucleus was almost invariably placed in the default position (the last stressed syllable). And in terms of tone, quite different nuclear pitch contours were produced in sentence-internal (continuing) and sentence-final intonation units, although in both structural positions the participants used both falling and rising contours. The phonetic parameters of pitch height, span, slope, alignment, duration, and intensity all proved to be relevant for these distinctions.

Key words: Serbian EFL, nuclear pitch contour, spontaneous speech, tonicity, tonality, tone

---

* E-mail address: tatjana.paunovic@filfak.ni.ac.rs

** Prepared as a part of the project *Innovations in the Teaching Process and Research in the Domains of English Linguistics and Anglo-American Literature and Culture*, conducted at the University of Niš – Faculty of Philosophy (No. 360/1-16-1-01).

1. Introduction

In foreign-language (L2) speech, intonation can be a source of difficulty, leading to possible miscommunications, or even to evoking unwanted interpretations, attributions and attitudes in the interlocutor. Problems can occur with *tonality* – the organization of utterances into intonation units, *tonicity* – expressing prominence within units, and *tone* – the realization of pitch movement on the nuclear tone (Halliday 1967, 1970; Wells 2006). With speakers of English as a foreign language (EFL), problems have been documented in prominence placement, in the realization of pitch movement, as well as in the overall phonetic properties of intonation contours (Mennen 2007: 55), including particularly the pitch range (Mennen et al. 2012), but also other gradient properties (Grice and Bauman 2007), e.g. pitch peak alignment (Graham and Post 2018).

However, despite a growing interest in investigating L2 prosody, and the mounting body of evidence about the properties of EFL students' interlanguage intonation, there are still very few studies that would include learners' use of prosody in relatively spontaneous oral production and not only in carefully controlled tasks, such as reading. On the other hand, there is some evidence that EFL speakers' intonation may be quite different in these two situations, especially in educational contexts. Therefore, the present study aimed to investigate how Serbian EFL learners organize their spontaneous speech into intonation units (IU), how they locate nuclear prominence within IUs, and how the nuclear accent is realized in complete IUs. The study focused particularly on the phonetic properties of pitch movement, pitch range, pitch slope, tone alignment, and intensity in the realization of nuclear accent contours.

2. Previous research

Empirical research has found similar problematic areas with EFL speakers of different first-language (L1) backgrounds. For instance, Jensen (2009) found that Danish EFL students did not distinguish the nucleus position from the first pre-nuclear accented syllable. Ramirez-Verdugo found that Spanish EFL speakers, in addition to commonly placing the nuclear prominence on the last stressed word in the intonation unit in all contexts, also produced inappropriate pitch movement in the nuclear position

(Ramirez Vergugo 2002, 2006; Ramirez Verdugo and Trillo 2005). Busà and Urbani (2011) found that Northern German EFL speakers produced a remarkably narrower pitch range than L1 (British) English speakers, and notably less pitch variation (Busà and Urbani 2011: 381). Similarly, investigating the types of tones, the pitch range of individual tones, and the pitch range of whole tunes in the production of 11 Slovene EFL speakers, Komar (2005) found that the participants used a narrower pitch range on the falling tones and, compared to L1 English speakers, a considerably smaller step up in pitch from the end of the pre-tonic segment to the beginning of the fall (Komar 2005: 3).

Whether EFL speakers' prosodic properties can be attributed to L1 transfer has also been a matter of concern, since EFL learners' interlanguage often shows properties that cannot be straightforwardly related to either L2 or L1. For instance, Toivanen (2003) investigated the intonation of 12 Finnish EFL students, acting out a pre-written conversational dialogue. Focusing on the phonetic properties of pitch movement, the study showed that the participants did make a distinction between falling and rising contours, but had problems using rising tones for "informational and/or pragmatic 'openness'", e.g. continuation, and used mostly falling tones in statements (Toivanen 2003: 168). Toivanen and Waaramaa (2005) observed a similar tendency of Finnish EFL speakers to use predominantly falling tones, often with a 'breathy voice quality' and a creak (Toivanen and Waaramaa 2005:181). Toivanen links such findings to the fact that "standard Finnish intonation is almost exclusively characterized by falling tones" (Toivanen 2003: 167).

However, investigating the influence of L1 Greek in EFL learners' prosody, Kainada and Lengeris (2015) reach a somewhat different conclusion. Observing the performance of 8 Greek EFL learners (4 male, 4 female), in reading tasks with polar questions and a written text, compared to corresponding samples of L1 (British) English from Grabe and colleagues' IViE corpus (Grabe et al.2001, in Kainada and Lengeris 2015: 274), the authors point out that EFL learners' prosodic properties differed from L1 English with respect to speech rate (slower in EFL), pitch span (narrower in EFL) and pitch level (lower in EFL). Moreover, not only was the Greek EFL speakers' pitch span in L2 English narrower than L1 English speakers', but it was also narrower than their pitch span in L1 Greek (Kainada and Lengeris 2015: 279), so this property could not be ascribed to L1 influence.

With Serbian EFL learners, several relevant research studies offer a number of comparable findings, despite their different aims, different participant groups, and methodologies applied. For instance, Marković (2011) compared Serbian EFL speaker's prosodic properties of pitch range, main stress, and tunes to L1 (British) English speech. In a reading task with declarative sentences from a narrative text, 15 university EFL students exhibited a remarkably narrower pitch range, different prosodic cues used for the main stress, and significantly different tunes than in L1 English (Marković 2011: 244). Also, while L1 English speakers used either a falling tone or a fall-rise as a signal of incompleteness or continuation, the participants mostly used a rising tone at the right IU boundary for continuation, which the author ascribes to L1 transfer (Marković 2011: 247). Lastly, the findings showed that the main stress had a lower pitch, a shorter duration and lower intensity in EFL speaker's production (Marković 2011: 248).

In Nikolić's study (2019), 6 Serbian EFL students' discourse-intonation features used in a dialogue-reading task were compared to those produced by 2 L1 (American) English speakers. The findings showed that EFL speakers produced a significantly narrower pitch range in all the investigated contexts, although they used finality and continuation signals appropriately, i.e. slight rises or falls for continuation, and falling tones (often followed by a laryngeal creak) for finality. Similarly, Paunović (2013) found that, in a discourse reading task, 4 Serbian EFL speakers and 4 L1 (British) English speakers signaled finality (of discourse topic) by a falling pitch, commonly followed by a laryngeal creak and a notable drop in intensity.

Paunović and Savić (2008) examined how 15 Serbian EFL students (10 female, 5 male) used intonation in a reading task, focusing on nuclear accent placement, direction of pitch change, and key or pitch range at transition points. The participants used appropriate falling tones or flat low tones for finality, and moderate rises or mid-level tones and slight falls for continuation (Paunović and Savić 2008: 71). However, the overall pitch range for different kinds of discourse (narrative/dialogue) was a problem for many participants, and all the signals were used notably less consistently in the dialogue than in the narrative.

Paunović (2019) investigated the prosodic signals of utterance-level information structure, i.e. the distinctions between broad, narrow and contrastive focus in L1 English, L1 Serbian and Serbian EFL. The participants were 4 male speakers of L1 (British) English and 4 male Serbian EFL students. In reading tasks, the EFL participants' intonation showed hybrid properties between their L2 and the specific variety of their L1 (the Prizren-Južna Morava dialect). Notably, the nuclear pitch range was significantly narrower in EFL students' production compared to L1 English, not due to low F0 maximum, but rather to higher F0 minimum values, transferred from L1 Serbian (Paunović 2019: 230).

Lastly, Paunović (2015) used both reading and speaking (retelling) tasks (semi-spontaneous speech, Markham and Hazan 2002) to investigate Serbian EFL students' pitch height and pitch range used to signal phrasing, information structure, and interactive functions. The findings showed that the 6 participants (3 male, 3 female) used pitch-related cues appropriately to signal unit boundaries and prosodic prominence, but that their pitch range was rather narrow, especially in reading dialogues. Overall, the participants used a narrower, and slightly higher pitch range in speaking than in reading (Paunović 2015: 89). An important finding was that in the speaking task the participants frequently used final rising tones instead of the expected falls. This could be, as suggested by Hirschberg (2002), interpreted as a signal of the participants' insecurity and lack of confidence (Paunović 2015: 90). However, in the light of more recent findings (Paunović, forthcoming), this increasing tendency to use a rising tone as a finality signal may also be a novel pragmatic and sociolinguistic property of EFL students' speech, similar to tendencies observed in different varieties of L1 English.

As shown in this summary, Serbian EFL learners have demonstrated several problems in their interlanguage intonation, the most commonly observed one being a narrow pitch range, but also the realization of the nuclear and post-nuclear pitch movement. However, only two of the described studies included elicitation tasks with semi-spontaneous speech in addition to reading, so the investigation of EFL students' spontaneous speech may show intonational properties different than the ones described in previous research.

3. Present study

The aim of the study was to analyse a collected corpus of spontaneous speech by Serbian EFL speakers in order to observe: 1. how the participants structured or 'packed' their speech into IUs (tonality); 2. how they located nuclear pitch accents within IUs (tonicity); and 3. how they melodically realised the nuclear accents (tones), in terms of the acoustic cues of pitch

(F0) movement, pitch range, pitch slope, tone alignment, intensity, and duration.

3.1. Methodology and procedures

The speech corpus was produced by 3 female and 2 male Serbian EFL speakers (mean age 21), 3rd year students at the English Department of the Faculty of Philosophy, University of Niš (CEFR C1 level). The recordings were made in a 'natural environment' (Llisterri 1992: 2), during an oral session in the course *Intercultural communicative competence,* in January 2018. The students were required to deliver a short (5-7 min.) expository talk on a topic they have researched for their final essays in the course. In this sense, the students' talks, though not interactive and conversational but rather careful and consultative in style (Joos 1968; Labov 1972), can still be considered spontaneous and not pre-rehearsed speech, since they were 'unscripted' (Llisterri 1992:19). In addition, each talk was followed by a short discussion among the students, who could ask for clarification, express disagreement, or comment on any aspect of the talk. The discussion parts were also recorded and included in the corpus for analysis.

Source	**Gender**	**Min.**		**Words**	
Speaker 1	F	9.20		1,532	
Speaker 2	F	6.50		912	
Speaker 3	M	12.30		1,838	
Speaker 4	F	12.00		1,576	
Speaker 5	M	6.25	**47.1**	756	**6,614**
Discussion 1	F, M	4.30		470	
Discussion 2	F, M	3.20		370	
Discussion 3	F, M	1.20		178	
Discussion 4	F, M	2.00	**11.17**	105	**1,123**
			58.27		**7,737**

Table 1. The structure of the spontaneous speech corpus

The recorded materials consisted of 58.27 minutes of students' speech, 47.10 min. of individual students' talks and 11.17 min. of discussions. Transcribed verbatim, with marked pauses, interruptions, and repetitions, the corpus consisted of 7,737 words, 6,614 in students' individual talks, and 1,123 in group discussions (Table 1).

Since students' talks were expository, they can be regarded as primarily exemplifying the textual metafunctions (Halliday 1994: 35-36), related to expressing the relevance status of information "as perceived by the speaker" (Ramirez Verdugo 2002: 119), and the speech function of giving information (Halliday 1970: 51). In the discussion parts of the corpus, the speech functions of demanding information were also relevant, as well as suggestions.

3.2. Analyses

The corpus was analysed in three steps. Firstly, using the transcriptions and the auditory and visual inspection of the recordings, IUs were identified in each participant's speech. IU boundaries were identified by relying on the 'external' and 'internal' phonetic criteria (pauses, anacrusis, pitch level and direction of unaccented syllables, final lengthening – Cruttenden 1997: 32–35), as well as on the syntactic, semantic and contextual clues. In cases of ambiguity ('intonational sandhi'), we opted for the more complex analysis into two IUs (Cruttenden 1997: 36). All the repetitions, and all the IUs that were incomplete – interrupted, unfinished, or in any way 'broken' – were excluded from further analysis. The complete IUs were classified and coded with respect to their syntactic position, as sentence-final or non-final (continuing), and for the type of syntactic structure they belonged to. Errors of grammar and use that did not affect the structure of the IU were ignored (e.g. **three dimension; *reflect to the whole family*).

Secondly, complete IUs were analysed for their internal structure, in the traditional British school model (Cruttenden 1997; Halliday 1967; Wells 2006), and the location of the nuclear prominence identified. Following this theoretical description, in addition to the nuclear syllable, the syllables following it (the tail) and stretching up to the right edge of the IU were also treated as potentially relevant for the realisation of the nuclear tone and coded for analysis. The identified nuclear accents were classified with respect to their overall pitch direction into two broad groups – falls and rises (Wells 2006). Complex tones, fall-rises and rise-falls, were treated as varieties of rises and falls respectively, but were coded for this information in a separate variable, and their component pitch movements measured separately in the acoustic analysis.

Thirdly, the acoustic analysis was performed in *Praat* (6.0.31, Boersma & Weenink, 1992-2010) with the settings of 400Hz pitch ceiling, 50Hz pitch floor, which proved appropriate for both female and male speakers. The total of 127 IUs showing a clear pitch contour throughout were selected for the acoustic analysis. F0 contours were visually inspected for octave errors and pitch halvings and the beginning of each nuclear tone movement was located manually. The measurements were taken for: F0/ pitch maximum, minimum, mean (in Hz); intensity maximum, minimum, and mean (in dB); and the duration of the pitch movement from its beginning to end (in seconds). If an overall falling contour was preceded by an identifiable rise within the nucleus, or the rising contour was preceded by a fall, the duration of that tone component was measured separately. From these measurements, the pitch range (span, excursion) of the nuclear tone was calculated (in Hz and semitones, ST), and the pitch slope was calculated as the pitch span (in ST) divided by the duration of the pitch movement. Tone alignment (the onset of the rising or falling movement) was also coded as a categorical variable – as early if it started during the nuclear syllable, immediate if it started on the post-nuclear syllable, and as late if it occurred further away from the nucleus towards the right edge of the IU.

The obtained data were statistically analysed by the SPSS software package (IBM, v.20). Upon testing the normality of data distribution, non-parametric tests were applied: Spearman's rank correlations for the relatedness of variables, Kruskal-Wallis tests to compare the means of the phonetic parameters (tested variables) in different nuclear tones (the grouping variable), and Mann-Whitney U tests to compare variable means in pairs. Standard deviations (SD) were reported where relevant.

4. Results

With respect to *tonality*, i.e. how the participants structured their speech into IUs, the analysis showed that a relatively big corpus of spontaneous speech yielded a notably small number of complete and analysable IUs, confirming Cruttenden's observation (1997: 29) that in natural conditions spontaneous speech is commonly characterized by abundant hesitations, interruptions, false beginnings, self-corrections, repetitions, and, in fact, a very small proportion of complete and uninterrupted sentences. Of the complete IUs, 75% came from students' talks, and 25% from the discussions; all the complete IUs were part of declarative statements, 80 (63 %) IUs occurring in sentence-final positions, and 47 (37%) as sentence-internal (continuing). Male and female participants contributed relatively equally, with 54% IUs produced by female and 46% by male participants. The complete IUs contained from 2 to 8 words (average 4.26).

Regarding *tonicity*, in the analysed IUs, the nuclear accent was placed on the last stressed syllable in a vast majority of cases, in 97% of IUs, while only in 4 (3%) IUs the nucleus was on the penultimate stressed word (*gender <u>pay-gap</u> in Hollywood; and he managed to make <u>82</u> million; and general adaptation for the <u>long</u> term; carry a <u>positive</u> connotation*). The words bearing the nuclear accent varied in their syllable structure: in 24% of the IUs there were no post-nuclear syllables, in 43% the nucleus was followed by one syllable, in 30% by two syllables, in 3% there were three or four post-nuclear syllables, and only one IU (cited above) had 6 syllables following the nuclear one.

Finally, with respect to *tone*, the overall falling pitch contour was found in 78% and rising movement in 22% of the IUs. Of the falling nuclear accents, 61% were found in sentence-final IUs, but as many as 39 % in sentence-internal IUs. Of the rising nuclear accents, only 30% were found sentence-internally, and as many as 70% in undoubtedly sentence-final positions. To further explore the possible relatedness of nuclear accent realization as a rise or a fall and the structural context in which the nucleus occurred, the IUs were re-coded for further statistical analysis into four types of nuclear tone realizations (cf, the examples in the Appendix):

1. nuclear falls in sentence-final IUs,
2. nuclear falls in continuative IUs,
3. nuclear rises in continuative IUs, and
4. nuclear rises in sentence-final IUs.

Then the measured acoustic properties were compared for these four nuclear types to explore possible differences in their phonetic realization. Also, the variables of post-nuclear syllable number, tone alignment, and type of discourse (talk or discussion) were included in further statistical analysis as potentially relevant for the realization of these four types of nuclear pitch contours.

4.1. Nuclear accent realizations: Final and continuative falls and rises

Correlations were found between the four nuclear accent types and almost all of the measured parameters (all $p<.000$) – strong for the F0 maximum ($r_s=.658^{**}$), minimum ($r_s=.618^{**}$), and mean ($r_s=.674^{**}$), moderate for the pitch range in ST ($r_s=-.352^{**}$), and the pitch slope ($r_s=-.245^{**}$), and weaker for intensity maximum ($r_s=.188^{*}$) and mean ($r_s=.275^{**}$), indicating that all these parameters were different in the four nuclear realizations. A strong correlation was also found for tone alignment ($r_s=.507^{**}$, $p<.000$), and a somewhat weaker one for the duration of the pitch movement ($r_s=-.208^{*,}$ $p<.019$), but not for the type of discourse or the number of post-nuclear syllables. Since the four types of nuclear realizations were assigned nominal values that increased from final falls, via continuation falls and continuation rises to sentence-final rises, the positive correlations suggested, for instance, that rising nuclear tones could have higher F0 values, and a later alignment. Conversely, the negative correlation indicated that the duration of the pitch movement could be shorter in rises than in falls.

Therefore, the four nuclear accent types were compared (Kruskal-Wallis test) for their mean values of the measured acoustic properties. Statistically significant differences were found for the F0 maximum, minimum, and mean, the pitch range and pitch slope, and for the intensity maximum, minimum and mean, while pitch movement duration only approached statistical significance. This is summed up in Table 2, which shows the Chi-square for these variables. The means and standard deviations (with confidence intervals) for all the measured parameters in the four different nuclear accent realizations are shown in Table 3.

	F0			Range		Slope		Intensity		
	Max	Min	Mean	Hz	ST	Dur.	ST/s	Max	Min	Mean
Chi-Square	56.56	63.46	57.87	54.49	61.06	7.50	65.12	14.70	10.74	20.93
df	3	3	3	3	3	3	3	3	3	3
Asymp. Sig.	.000	.000	.000	.000	.000	.058	.000	.002	.013	.000
a. Kruskal Wallis Test										
b. Grouping Variable: NUC type										

Table 2. Means comparison (Kruskal-Wallis) for the pitch and intensity variables in the four types of nuclear accent realizations

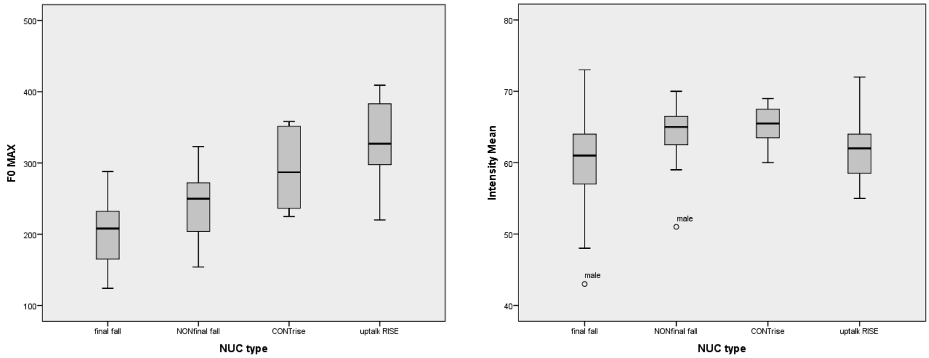
The comparison of mean values showed two different patterns. On the one hand, the mean values of F0 maximum and mean were higher in the rising than in the falling tones irrespective of the structural position they occurred in. This is illustrated in Graph 1 – Left, which shows boxplots of normality for F0 maximum. On the other hand, both continuation contours had higher mean values for F0 minimum than both final contours, irrespective of the pitch movement direction. Intensity means showed a similar pattern – a noticeable drop in intensity accompanied both final contours, irrespective of the movement direction, while both continuation contours kept higher intensity parameters. This is illustrated in Graph 1 – Right, which shows boxplots of normality distribution for Intensity maximum.

		Mean	**Std. Dev.**	**95% Confidence Interval for Mean**	
				Lower Bound	**Upper Bound**
F0 Max	Final fall	199.89	44.44	188.50	211.27
	Cont. fall	241.03	41.87	227.45	254.60
	Cont. rise	291.63	56.61	244.29	338.96
	Final rise	331.00	56.50	303.77	358.23

F0 Min	Final fall	85.33	35.08	76.34	94.31
	Cont. fall	175.62	43.92	161.38	189.85
	Cont. rise	163.38	26.09	141.57	185.18
	Final rise	155.68	54.19	129.57	181.80
F0 Mean	Final fall	149.48	36.76	140.06	158.89
	Cont. fall	206.38	40.15	193.37	219.40
	Cont. rise	227.25	41.69	192.40	262.10
	Final rise	248.00	54.76	221.61	274.39
Range Hz	Final fall	115.70	43.34	104.60	126.81
	Cont. fall	65.41	27.44	56.51	74.31
	Cont. rise	128.25	45.93	89.85	166.65
	Final rise	175.32	63.23	144.84	205.79
Range ST	Final fall	15.50	5.50	14.09	16.91
	Cont. fall	5.85	3.33	4.77	6.93
	Cont. rise	9.95	2.93	7.50	12.40
	Final rise	13.96	6.52	10.82	17.10
Duration of pitch movement	Final fall	.399	.148	.361	.437
	Cont. fall	.380	.114	.343	.417
	Cont. rise	.345	.074	.283	.406
	Final rise	.316	.127	.255	.378

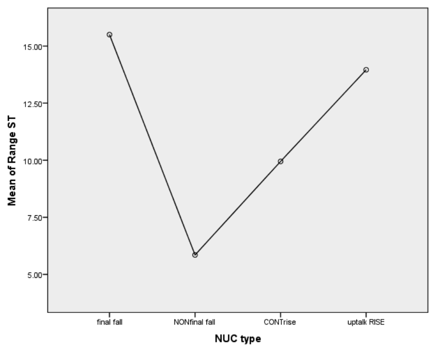
Slope ST/sec.	Final fall	42.53	19.84	37.45	47.61
	Cont. fall	15.56	8.12	12.93	18.19
	Cont. rise	29.19	7.94	22.55	35.83
	Final rise	46.34	18.93	37.22	55.47
Intensity Max	Final fall	64.61	6.75	62.88	66.33
	Cont. fall	68.90	3.58	67.74	70.06
	Cont. rise	68.50	2.67	66.27	70.73
	Final rise	67.11	9.24	62.65	71.56
Intensity Min	Final fall	48.92	7.59	46.98	50.86
	Cont. fall	51.92	6.26	49.89	53.95
	Cont. rise	55.75	7.59	49.40	62.10
	Final rise	48.21	4.67	45.96	50.46
Intensity Mean	Final fall	59.97	6.05	58.42	61.52
	Cont. fall	64.23	3.45	63.11	65.35
	Cont. rise	65.25	3.01	62.73	67.77
	Final rise	62.00	4.40	59.88	64.12

Table 3. Means of the measured phonetic parameters in the four types of nuclear accent realizations

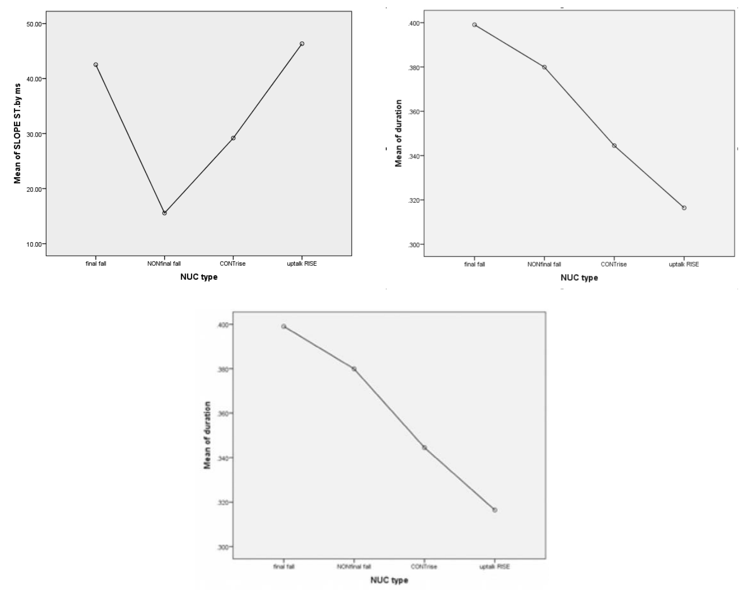


Graph 1. Left: Normality boxplots of the parameters F0 maximum for the four types of nuclear accent realizations; Right – Normality boxplots of Intensity mean

The mean pitch range in ST was also notably wider in both final contours, whether falling or rising, and remarkably narrower in both continuation contours, particularly in continuation falls. This is illustrated by Graph 2, which shows the plotted means of the pitch range in ST. It should also be noted that the pitch range difference between the means of final falls and final rises was much smaller than 3 ST, and thus irrelevant (Nooteboom 1997: 645). The pitch slope was also notably steeper in both final contours and shallower in both continuation contours, with continuation falls showing the 'mildest' and the shallowest pitch contour. This is illustrated in Graph 3 – Left, which shows the plotted means of the pitch slope in the four types of nuclear accent realizations.



Graph 2. Plotted means of the nuclear pitch range in ST in four nuclear accent realizations



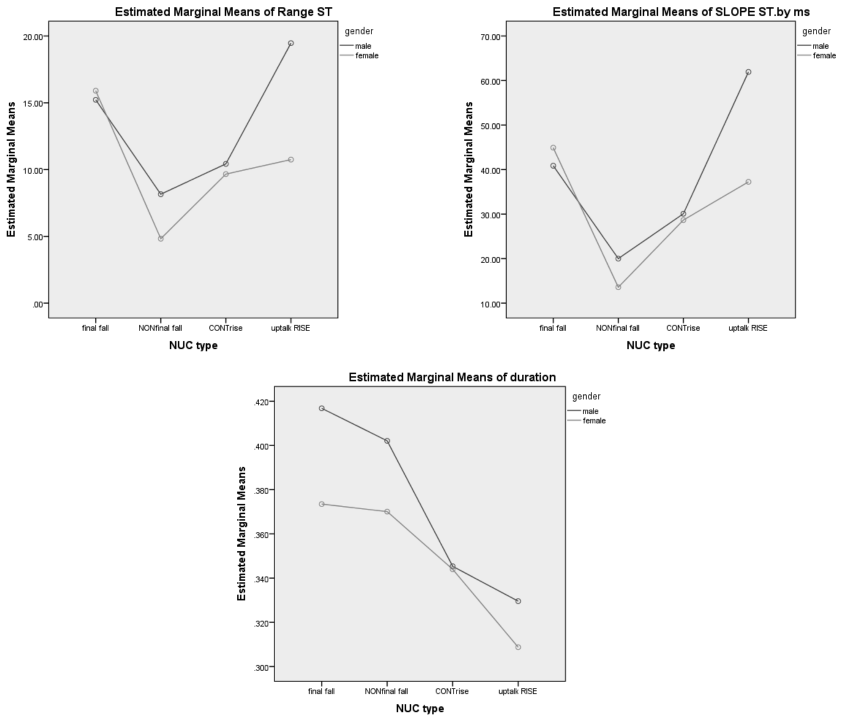
Graph 3. <u>Left</u>: Plotted means of the pitch slope in the four types of nuclear accent realizations; <u>Right</u>: Plotted means of the nuclear pitch movement duration in the four types of nuclear accent realizations

However, the parameter of pitch movement duration, like F0 maximum, was more sensitive to the nuclear pitch movement direction than to the structural position. Although this parameter did not show a clear statistical significance, means and mean ranks comparisons suggested that falling contours – whether final or not – had a longer duration, while the pitch movement in the rises was shorter. Still, the structural position did show some influence, since continuation falls were shorter than final falls, and continuation rises were not as short as final rises, which were the shortest. This is illustrated in Graph 2 – Right, which shows the means of the nuclear pitch movement duration in the four types of nuclear accent realizations.

This observation is further supported by the finding that final rises were more commonly preceded by a falling portion of the nuclear tone before the rise. Of the IUs with the post-nuclear (immediate or late) onset of the primary tone movement (18% of all the IUs analysed), 97% were found in rising contours, and 70% of these were actually in sentence-final IUs. Additionally, the duration of the falling component preceding the nuclear rise was longer in final rises (mean=.138sec, StD=.107) than in continuation rises (mean=.111sec, StD=.088). This was partially supported by the Mann-Whitney comparison of continuing and final nuclear rises. Although statistically significant differences were found between them only for the intensity parameters ($p<.050$) and the pitch slope ($p<.019$), the mean rank comparisons did indicate that in addition to having higher F0 maximum and mean, a lower F0 minimum, a wider pitch range, a shorter duration and a steeper slope, the final nuclear rises also had a longer duration of the falling portion of the nuclear contour before the rise, compared to continuation rises.

4.2. Gender differences and similarities

Overall, female participants produced more rising nuclear tones than male participants. While the distribution of falling tones was balanced (52% by female and 48% by male speakers), of the IUs with rising nuclear accents 63% were produced by female and 37% by male participants. In addition, a significant correlation was found between the type of nucleus and gender ($r_s=.226^*$, $p<.011$), suggesting that the four nuclear contours were not realized in the same way by female and male participants. Indeed, mean rank comparisons (Mann-Whitney U test) for the variable of gender showed significant differences (all $p=.000$) between female and male participants in the values of F0 maximum, minimum and mean (higher with females), the nuclear pitch range in ST (surprisingly, wider in male participants), and the steepness of the pitch slope (steeper with males).



Graph 4. Estimated marginal means compared in female and male participants for: Upper Left – Pitch range in ST; Upper Right – Pitch slope; Lower Central – The duration of the nuclear pitch movement

However, despite these differences, the comparison of correlation coefficients across genders showed that within either group the same variables correlated with the four nuclear categories, suggesting that both male and female participants produced the four different nuclear tone realizations with specific phonetic properties.

This interpretation was also supported by the Two-way between-groups (Univariate) analysis of variance, which indeed showed some interaction of the variables of gender and nucleus type, but still showed the main influence of the type of nuclear pitch realization, i.e. that both male and female participants distinguished between the four different types of nuclear pitch contours. This is illustrated in Graph 4, which shows the estimated marginal means for female and male participants in the four types of nuclear contours for the pitch range in ST (upper left), pitch slope (upper right), and the duration of pitch movement (lower central). The interaction of the two variables is obvious, but the patterns of distinguishing between the four nuclear pitch contours are still generally similar in male and female participants.

5. Discussion

The presented results show that this corpus of spontaneous speech shared some characteristics with EFL students' oral performance described in previous research, but it also showed some peculiar properties.

Concerning *tonality*, the main characteristic of the corpus was a disproportionally small number of complete IUs, and a great number of interruptions, hesitations, repetitions and incomplete structures, even though the participants talked on well-researched and familiar topics. Similar findings were reported in Paunović (2015) for the task involving semi-spontaneous speech: long pauses, hesitations, filled pauses, and broken pitch contours. This is described as a common property of spontaneous speech in general (Cruttenden 1997), but in the present study it might have been enhanced by the stress and anxiety that commonly accompany students' oral performance in formal educational contexts. Therefore, further research should also focus on EFL students' spontaneous production in more relaxed communicative situations, where their interlanguage intonation may show different tonality properties.

With respect to *tonicity*, in line with the findings of some previous research (e.g. Ramirez-Vergugo 2002, 2006), an important property of the participants' spontaneous speech was that nuclear pitch accents were almost invariably located in the default nuclear position, with just a few exceptions. This could be ascribed to the expository and informative nature of the discourse produced, so further research should focus on more interactive and conversational communicative situations that would elicit other types of discourse, too.

Concerning *tone*, i.e. the melodic realisation of the nuclear accents, our EFL participants did not comply with the traditional expectation of associating falling contours with 'closed' ('non-continuative') and rising contours with 'open' ('continuative') meanings (Cruttenden 1997: 163). They produced both falling and rising nuclear tones in both types of structural contexts, final and non-final, but with different pitch contours.

These findings differ from Marković (2011), where Serbian EFL speakers signalled continuation mainly by rising tones at the right IU boundary, and also from Paunović (2015), where continuation was signalled mostly by rising tones, too. However, they are partly similar to the findings in Paunović and Savić (2008: 71), where EFL speakers used falling tones for finality, or flat tones with a low key, and used moderate rises or slight falls for continuation.

An important finding in this study was that some of the phonetic parameters were more closely related to the direction of pitch movement, while others were more sensitive to the structural context in which the nuclear tone was realized. The parameters more closely linked to pitch direction included: the *pitch height*, since both continuing and final rises reached higher F0 maxima than both falls; the *duration of pitch movement*, as both rises were shorter than falls; and *tone alignment*, since it was later for rises than for falls. The parameters that showed more context sensitivity included: *intensity*, since final contours had lower intensity irrespective of pitch direction; the *F0 minimum*, since both final contours involved lower pitch minima than either of the continuing contours; the *pitch range in ST*, since it was wider in both final nuclear contours; and the *pitch slope*, which in both final rises and falls was steeper than in either of the continuing contours.

Another important finding was the fact that the nuclear pitch range, especially in sentence-final IUs, was not nearly as narrow as reported in previous research. For instance, Busà and Urbani (2011: 381), Komar (2005), Mennen et al (2012), and Kainada and Lengeris (2015) all point out that their participants' pitch range in L2 English was notably narrower than L1 English speakers'. With Serbian EFL speakers, a narrow pitch range was frequently reported, too (Marković 2011; Nikolić 2019; Paunović and Savić 2008; Paunović 2019).

However, a comparison of the nuclear pitch range (in ST) observed in the present study to some previous findings with Serbian EFL speakers shows important differences. For instance, Paunović (2013) found that the highest nuclear pitch range mean was 8.0 ST (in discourse-topic initial IUs), followed by 7.4 ST (in topic-final IUs). Similarly, in Paunović (2019), the widest mean pitch span (realized on the nuclear *and* post-nuclear syllable) was 9.9 ST (in contrastive-focus IUs), followed by 8.9 ST (in narrow-focus IUs) and 6.2 ST (in broad-focus IUs). In both these studies, the nuclear tones investigated were falls, and both studies used reading tasks for elicitation. In Paunović (forthcoming), rising nuclear contours were investigated, with an elicitation task that involved pre-rehearsed (semi-spontaneous) speech, and the widest mean pitch range observed was 8.3 ST. Compared to all these findings, the mean values of the nuclear pitch range in the present study were indeed remarkably higher, especially in the sentence-final IUs – 15.50 ST for final falls and 13.96 ST for final rises. Even the mean value of the pitch range in continuation rises (9.95 ST) in the present study was, if not higher, then at least on a par with the means reported by previous studies, for instance, in discourse-topic initial IUs (Paunović 2013) and even under contrastive stress Paunović (2019).

This finding is important because it suggests that the narrower pitch range observed in the oral production of EFL speakers in much previous research might have resulted from the type of elicitation task – it might have reflected the properties of the participants' *reading skill* rather than those of their interlanguage intonation. If the EFL speakers in the present study produced notably wider nuclear pitch ranges even in expository and informative spontaneous speech, it could be expected that they may produce even wider pitch ranges in more conversational and interactional communicative situations, so this is an important direction for further research.

Summing up the phonetic properties of the four types of nuclear contour realizations, it can be said that the final falls were characterized by the lowest pitch – maximum, mean, and, particularly minimum – and that they often ended in a creak (in over 30% of occurrences), as also observed in several previous studies (Nikolić 2019; Paunović 2015). Continuative falls had the narrowest pitch range and the mildest or shallowest pitch slope.

Continuative rises started with a moderate onset (F0 minimum), not as low as the onset of final rises, and reached a notably narrower pitch range than final rises. Their slope was moderate, too, about 60% of the final rise slope, and about twice as steep as that of continuation falls.

Finally, the rises that were used in sentence-final IUs exhibited very peculiar properties – they reached the highest F0 maximum, had a wide pitch range, and – due to their shortest duration – also had the steepest slope of all the nuclear contours. The onset of the final rise was characterized by late alignment much more often than any other nuclear contour, and was preceded by a fall more often than continuation rises, with a longer duration of this fall.

This kind of rising nuclear contour, found in sentence-final structural contexts, was not reported in previous studies in which reading elicitation tasks were used, which indicates that this rising contour is typical of spontaneous speech rather than other kinds of oral production. Only the studies that involved semi-spontaneous speech (e.g. Paunović 2015) found that rises were used 'instead of falls' in the sentence-final contexts, but described them as 'inappropriate' (Paunović 2015:82), echoing the frequently expressed view that final falls indicate certainty and authority while rises voice uncertainty and deference (cf. Ramirez-Verdugo 2002: 120; Hirschberg 2002). However, in a more recent study (Paunović, forthcoming), the rising pitch contours found in declarative-sentence final IUs (referred to as *uptalk*) are described as a distinct finality signal, deliberately used by the EFL students in pre-rehearsed semi-spontaneous speech, and different from continuation-signalling rises. Uptalk rises exhibited distinct phonetic properties – a large pitch excursion, a steep rise slope, and a high rise peak (Paunović, forthcoming), much like the nuclear rising contours found in the sentence-final IUs in the present study. This finding is very important, because it points to a possible development of a novel prosodic device in EFL students' intonation in spontaneous speech, which undoubtedly requires much further research.

6. Conclusion

The prosody of spontaneous speech is not widely researched, for several reasons. They include not only technical difficulties and ethical concerns, but also methodological problems with analysing spontaneous speech corpora. In the present study, the rather traditional concepts of tonality, tonicity and tone provided a very useful framework for analysis, facilitating the description and classification of the observed characteristics of EFL students' intonation in spontaneous speech. The findings showed that EFL students' speech exhibited specific properties at all the three levels. The most important ones included abundant disruptions at the level of tonality, a preference for the default nuclear position at the level of tonicity, and a number of specific properties at the level of tone realization. An important tone property was a wider nuclear pitch range than in other kinds of oral production; another was the use of four distinct nuclear pitch contours, resulting from the interaction of pitch movement direction – falling, rising – and the structural contexts in which the nuclear accents were realized – final or non-final. Each of these nuclear contours was characterized by a distinct combination of the phonetic properties of pitch height, range, slope, alignment, duration and intensity.

This study was limited to a specific type of discourse (expository, informative), a rather careful speech style, and a small number of participants, and these limitations clearly highlight the necessary lines of further research. It should include a broader variety of communicative situations and discourse types in which some different intonational properties of spontaneous speech may emerge. Lastly, a better insight into EFL students' spontaneous speech intonation would undoubtedly be offered by a comparative study of their L1 Serbian spontaneous speech.

References

Boersma, P. and D. Weenink (2013). Praat: Doing phonetics by computer (Version 5.2.04, 1992–2010) [Computer software]. (11 November 2017) <http://www.praat.org>.

Busà, G. M. and M. Urbani (2011). A cross-linguistic analysis of pitch range in English L1 and L2. In: W. S. Lee and E. Zee (eds.), *Proceedings of the 17th International Congress of Phonetic Sciences* (ICPhS XVII), Hong Kong: University of Hong Kong, 380–383.

Cruttenden, A. (1986[1], 1997[2]). *Intonation*. Cambridge: Cambridge University Press.

Grabe, E. et al. (2001). Modelling intonational variation in English: The IViE system. In: S. Puppel and G. Demenko (eds.), *Prosody 2000,* Poznan: Adam Mickiewicz University, 51–57.

Grice, M. and S. Bauman (2007). An introduction to intonation – functions and models. In: J. Trouvain and U. Gut (eds.), *Non-Native Prosody: Phonetic Description and Teaching Practice,* Berlin/New York: Mouton de Gruyter, 25–51.

Halliday, M. A. K. (1967). *Intonation and grammar in British English*. The Hague: Mouton.

Halliday, M. A. K. (1970). *A course in spoken English: Intonation*. London: Oxford University Press.

Halliday, M. A. K. (1994). *An introduction to functional grammar*. London: Arnold.

Hirschberg, J. (2002). The Pragmatics of Intonational Meaning. In: B. Bel and I. Marlien (eds.), *Proceedings of Speech Prosody 2002,* Aix-en-Provence Laboratoire Parole et Langage, Université de Provence, 65–68.

Jensen, C. (2009). Prominence relations in British English – A contrastive perspective. *English Studies*, 90(1), 78–99.

Joos, M. (1968). The isolation of styles. In: J. A. Fishman (ed.), *Readings in the Sociology of Language*. The Hague: Mouton, 185–191.

Kainada, E. and A. Lengeris (2015). Native language influences on the production of second-language prosody. *Journal of the International Phonetic Association*, 45(3), 269–287.

Komar, S. (2005). The impact of tones and pitch range on the expression of attitudes in Slovene speakers of English. In: *Proceedings of Phonetics Teaching and Learning Conference 2005* (PTLC), London: University College London, 1–4.

Labov, W. (1972). The Isolation of Contextual Styles. In: W. Labov, *Sociolinguistic Patterns,* Philadelphia: University of Pennsylvania Press, 70–109.

Llisterri, J. (1992). Speaking styles in speech research, *ELSNET/ESCA/SALT Workshop on Integrating Speech and Natural Language*. Dublin, Ireland, 15-17 July 1992. (3 October 2013) <http://liceu.uab.es/~joaquim/publicacions/SpeakingStyles_92.pdf>.

Markham, D. and V. Hazan (2002). Speech, Hearing and Language: Work in progress. In: D. Markham and V. Hazan (eds.), *UCL Speaker database* (Vol. 14), London: UCL, 1–17.

Marković, M. (2011). Acquiring second language prosody: Fundamental frequency. *English language and Anglophone literatures Today* (ELALT) (Novi Sad, 19 March 2011), 238–249. (14 August 2018) <http://elalt.info/elalt_I.pdf#page=238>.

Mennen, I. (2007). Phonological and phonetic influences in non-native Intonation. In: J. Trouvain and U. Gut (eds.), *Non-Native Prosody: Phonetic Description and Teaching Practice,* Berlin/New York: Mouton de Gruyter, 53–76.

Mennen, I. et al. (2012). Cross-language difference in f0 range: a comparative study of English and German. *Journal of the Acoustical Society of America*, 131(3), 2249–2260.

Nikolić, D. (2019). Tone and Break Index (TOBI) theoretical framework applications with EFL speakers. Presented at the Conference *Language, Literature, Theory,* April 2019, Faculty of Philosophy, University of Niš, Serbia. (21 September 2019) <https://www.researchgate.net/publication/332707667_Dusan_Nikolic-JKT2018-TONE_AND_

BREAK-INDEX_TOBI_THEORETICAL_FRAMEWORK_APPLICATIONS_WITH_EFL_SPEAKERS>.

Nooteboom, S. (1997). The prosody of speech: melody and rhythm. *The Handbook of Phonetic Sciences*, 5, 640–673. (22 January 2011) <https://www.phil.uu.nl/tst/2012/Werk/PROSTEM.pdf>.

Paunović, T. (2013). Beginnings, endings, and the in-betweens: Prosodic signals of discourse topic in English and Serbian. In: B. Čubrović and T. Paunović (eds.), *Focus on English Phonetics*, Newcastle: Cambridge Scholars, 191–213.

Paunović, T. (2015). Pitch height and pitch range in Serbian EFL students' reading and speaking tasks. *Nasledje,* 32, 73–95.

Paunović, T. (2019). Focus on focus: Prosodic signals of utterance-level information structure in L1 Serbian, L1 English, and Serbian L2 English. *Zbornik Matice srpske za filologiju i lingvistiku*, LXII/2, 213–238.

Paunović, T. (forthcoming). The elephant in the room: *Uptalk* in EFL Students' Classroom Speech. Under review.

Paunović, T. and M. Savić (2008). Discourse intonation-Making it work. *ELOPE: English Language Overseas Perspectives and Enquiries*, 5(1–2), 57–75.

Ramirez Verdugo, D. (2002). Non-native interlanguage intonation systems. A study based on a computerized corpus of Spanish learners of English. *ICAME Journal*, 26, 115–32.

Ramirez Verdugo, D. (2006). Prosodic realization of focus in the discourse of Spanish learners and English native speakers. *Estudios Ingleses de la Universidad Complutense,* Vol. 14, 9–32.

Ramirez Verdugo, D. and J. R. Trillo (2005). The pragmatic function of intonation in L2 discourse: English tag questions used by Spanish speakers. *Intercultural Pragmatics*, 2(2), 151–168.

Toivanen, J. (2003). Tone choice in the English intonation of proficient non-native speakers. *Phonum*, 9, 165–168.

Toivanen, J. and T. Waaramaa (2005). Tone choice and voice quality of dispreferred turns in the English of Finns. *Logopedics Phoniatrics Vocology,* 30(3–4), 181–184.

Wells, J. C. (2006). *English Intonation: An Introduction*. Cambridge: Cambridge University Press.

Received: 27 February 2020
Accepted for publication: 23 July 2020

Appendix

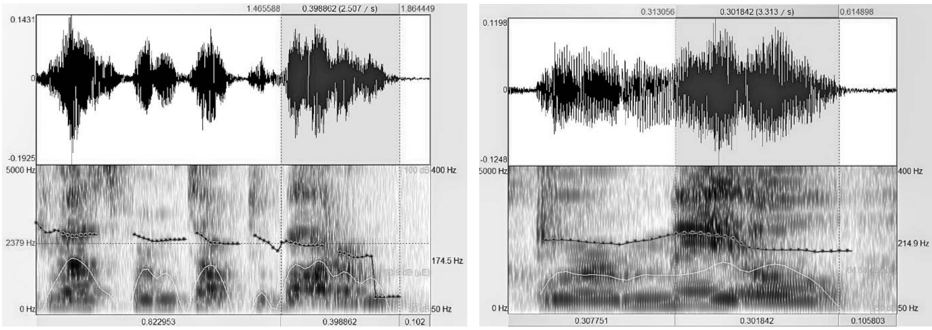


Figure 1. Left: Spectrogram of the IU [*dressed from top to bottom.*] illustrating a final fall by a female speaker. Right – Spectrogram of the IU [*you're male,*] illustrating a continuation fall by a female speaker. The words bearing the nuclear accent are highlighted.

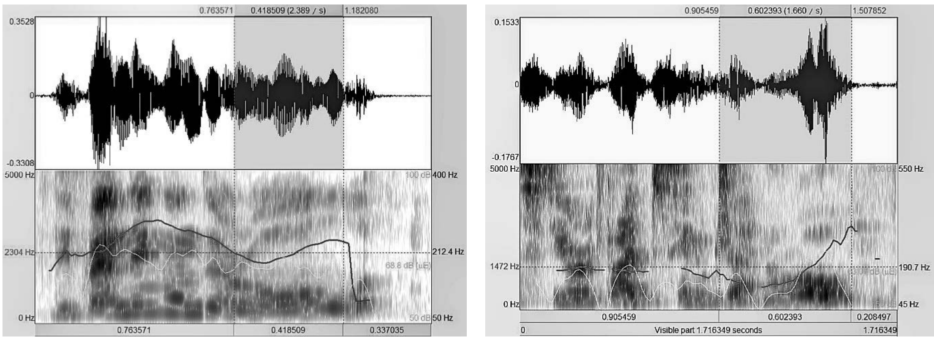


Figure 2. Left: Spectrogram of the IU [*depending on the role, <er>*] illustrating a continuation rise by a female speaker. Right: Spectrogram of the IU [*stereotypes in general.*] illustrating a final rise by a male speaker. The words bearing the nuclear accent are highlighted.

Татјана Пауновић

ТОНАЛНОСТ, ТОНИЧНОСТ, ТОН: ИНТОНАЦИЈА У СПОНТАНОМ ГОВОРУ СТУДЕНАТА ЕНГЛЕСКОГ КАО СТРАНОГ ЈЕЗИКА

Сажетак

У овом раду представљена је квантитативна и акустичка анализа корпуса спонтаног говора српских студената енглеског као страног језика, у погледу: организације интонацијских целина (тоналност), избора нуклеуса (тоничност), и реализације нуклеарних тонских контура (тон). Резултати су показали да, у домену тоналности, спонтани говор испитаника карактеришу испрекиданост, застоји и бројне паузе, као и релативно мали број комплетних и завршених интонацијских целина. У домену тоничности, нуклеус се, готово без изузетка, налазио у подразумеваној (дифолт) позицији, на последњем наглашеном слогу интонацијске целине. У домену реализације тона, различите тонске контуре реализоване су у две кључне позиције – у интонацијским целинама које се налазе на крају реченице, и онима унутар реченице. У обе ове структурне позиције испитаници су користили како силазне тако и узлазне тонске контуре, али су се оне битно разликовале по акустичким својствима. Сви кључни фонетски параметри – висина тона, тонски опсег, нагиб кретања тона, поравнање тона, као и трајање и интензитет – показали су се значајним за сигналиризирање интонацијских дистинкција.

Кључне речи: српски ученици енглеског као страног језика, нуклеарне тонске контуре, спонтани говор, тоналност, тоничност, тон