DOI: https://doi.org/10.18485/philologia.2021.19.19.6
UDC: 811.111'34
81'243

■ ACOUSTIC CHARACTERISTICS OF AMERICAN ENGLISH MONOPHTHONGS IN SERBIAN EFL SPEAKERS – A CASE STUDY

ANA BJEKIĆ[1]
BILJANA ČUBROVIĆ[2]
University of Belgrade
Faculty of Philology
Belgrade, Serbia

Cilj ovog rada je da ispita kvalitet engleskih vokala u realizaciji izvornih govornika srpskog jezika i da dobijene rezultate uporedi sa referentnim vrednostima izvornih govornika američkog engleskog (Čubrović 2016). Uzorak je činilo 9 muških govornika engleskog kao stranog jezika iz jedne urbane gradske sredine u centralnoj Srbiji. Istraživanje je izvršeno na korpusu 9 vokala američkog engleskog /i ɪ ɛ æ ʌ u ʊ ɔ ɑ/ u rečima strukture konsonant+vokal+konsonant. Inicijalni konsonant bio je /p/ ili /t/, a finalni /t/. Rezultati istraživanja pokazuju da postoji uticaj maternjeg jezika na realizaciju engleskih vokala /i ɪ/, /ɛ æ/ i /u ʊ/ budući da ih govornici ne razlikuju u dovoljnoj meri u engleskom jeziku. Uočeno je da su kod zadnjih vokala /ɔ ɑ/ vrednosti F1 približne što ukazuje na proces stapanja ovih vokala u jedan, što se poklapa sa govornim navikama izvornih govornika američkog engleskog. Statističkom analizom mešovitih efekata utvrđeno je da postoji statistički značajna razlika u vrednostima F1 i F2 između izvornih i neizvornih govornika kod 8 vokala američkog engleskog /ɪ ɛ æ ʌ u ʊ ɔ ɑ/, a izuzetak je vokal /i/.

Ključne reči: vokali, američki engleski, srpski govornici engleskog kao stranog jezika, vokalski kvalitet.

1. INTRODUCTION

Several vowel studies focused on the acquisition of British English vowels by Serbian EFL speakers (Paunović 2002, 2011; Marković 2007, 2009; Dančetović 2017; Bjelaković

1 Kontakt podaci (Email): anabjekic@gmail.com
2 Kontakt podaci (Email): biljana.cubrovic@fil.bg.ac.rs

2018). The reasoning behind this is a widespread belief that Serbian EFL learners are taught standard British English. This is in line with Cruttenden's (2014: 327) claim that European EFL learners commonly perceive British English as a preferred pronunciation model. On the other hand, some recent studies have repeatedly shown that American English (AE) has become the "preferred pronunciation model in recent years, because speakers are overwhelmingly exposed to American English (more than 90% of them) [...]" (Čubrović/Bjelaković 2020: 149).

The goal of this study is to investigate spectral characteristics of American English vowel inventory (9 monophthongs) produced by 9 Serbian EFL speakers in order to expand the study carried out by Čubrović (2016) and look at the vowels of American English as realized by a group of respondents who live in the city of Čačak. Another aim of the study is to provide additional information about nonnative participants' production of qualitative differences between AE vowels and the existence of the low back merger in their speech.

American English (AE) has eleven distinct monophthongs /i ɪ e ɛ æ ʌ u ʊ o ɔ ɑ/ (Yavaş 2011: 77-78), but this experimental study investigates only nine - /i ɪ ɛ æ ʌ u ʊ ɔ ɑ/. Since the AE vowels /e/ and /o/ are often diphthongized, they are excluded from this study. Vowels /i ɪ ɛ æ/ are described as front, /ʌ/ as a central vowel, and /u ʊ ɔ ɑ/ as back vowels (Yavaş 2011: 78). AE vowel inventory differentiates between tense and lax vowels (Ladefoged/Maddieson 1996). Tense vowels /i e o u/ tend to be longer and higher than their lax counterparts /ɪ ɛ ʊ ɔ/ (Zsiga 2013: 28). However, the English vowel /æ/ "has equal duration with, or even greater duration than typically long and tense vowels such as /ɑ e o/" (Yavaş 2011: 80). Another salient characteristic of AE is the occurrence of a merger of its two back vowels /ɔ/ and /ɑ/. This results in some pairs of words being pronounced as homophones, e.g. *cot* vs. *caught*, *stock* vs. *stalk, bought* vs. *pot*. There is clear evidence that the low back merger is nearing completion in many regions across the United States[3] (Boberg 2015). On the other hand, some notable vowel studies (Peterson and Barney 1952; Hillenbrand *et al.* 1995[4]) did not take the merger into account.

In contrast to the complex vowel system of English, Serbian vowel inventory contains five vowels: /a/, /e/, /i/, /o/ and /u/. Five-vowel inventories are among the most frequently used vowel inventories in the languages of the world (Crothers 1978; Maddieson 1984). The vowels /i/ and /e/ are described as front (Miletić 1952), /a/ is a mid vowel, and /o/ and /u/ are referred to as back vowels (Simić/Ostojić 1996). Concerning height of the tongue, the high-vowel category is formed by /i u/, mid vowels are /e o/, and /a/ is typically defined as a low vowel. Traditional accounts of Serbian vowel system hold that each short vowel is matched by a long counterpart. In addition to this, Serbian is characterized by a complex pitch accent system. The quantity of Serbian

3 "This merger is now complete in northern New England, the West and Canada, as well as in parts of the Midland and South, [...]. It is in progress in the remaining parts of the Midland and South – more advanced in some communities and social groups than in others – and may even be making inroads among younger, upwardly mobile speakers in the areas that have historically resisted it, as the pronunciation features that prevented it in the past become socially stigmatized" (Boberg 2015: 233).

4 Hillenbrand *et al.* (1995) explicitly used the presence of the merger as a disqualifying criterion, i.e. they made sure they only used speakers who did not have the merger.

vowels is treated as part of the pitch accent system. Ivić and Lehiste (2002) found that vowel length affects the quality of some of the Serbian vowels, where vowels /e/ and /o/ exhibit the most conspicuous effects of this influence.

2. EXPERIMENT

2.1. PARTICIPANTS

The participants of this study are 9 Serbian speakers of American English, all students of Technical College of Vocational Studies, Čačak, at the time of recordings. Before the recordings were made, participants were asked to fill in a questionnaire, since a uniform group of participants was needed, with a preference for the American English pronunciation model. All participants spoke the same dialect of Serbian at the time of recordings. All 9 participants were born and lived in urban areas of Čačak and they had been learning English for 11 to 12 years at school. Their age ranged from 19 to 21 and none of them had lived in an English speaking country. Unlike some previous studies (Paunović 2002; Marković; 2007; Dančetović 2017; Bjelaković 2018), participants of this study were not students of an English Department and none of them had attended a course in phonetics or pronunciation.

2.2. MATERIALS AND RECORDING PROCEDURES

Recordings were made of participants reading words (*beat, bit, bet bat, but, boot, put, bought, pot*) containing 9 target vowels /i ɪ ɛ æ ʌ u ʊ ɔ ɑ/. The words have the CVC phonological structure, with an initial labial consonant – /p/ or /b/. The final consonant in all words is coronal /t/ so as to avoid possible effects of different manners and places of articulation on the preceding vowel. The words were embedded in the frame sentence "Say _____again", and repeated 3 times. The total number of tokens was 243 tokens (9 speakers x 3 repetitions x 9 words, one for each vowel).

Participants were given sentences in a Power Point Presentation, and only one sentence was presented at a time. Before the recordings were made, the participants were instructed to familiarize themselves with the sentences and read them "as naturally as possible" so as to avoid list intonation.

The recordings were made in a quiet room at Technical College of Vocational Studies in *Praat* speech software, Version 5.4 (Boersma/Weenink 2014) and digitized at 5000 Hz as recommended for male speakers (Boersma 2013: 395). The values of the first and second formants were extracted manually using the function "Get formant" in *Praat* (Boersma/Weenink 2014). Formant measurements were taken from the steady-state portion of the vowel. The mean values of F1 and F2 for every token were used to create F1-F2 graphs that display the vowel space. F1 is plotted on the vertical axis and it is related to the vowel height. F1 is inversely related to the vowel height, so low vowels have a high F1. F2 is plotted on the horizontal axis in graphs, and it is related to the degree of backness and lip rounding, so front vowels have high F2 and back vowels are characterized by low F2. Each point in graphs represents one vowel repetition by one participant. Speakers are marked as M1-M9.

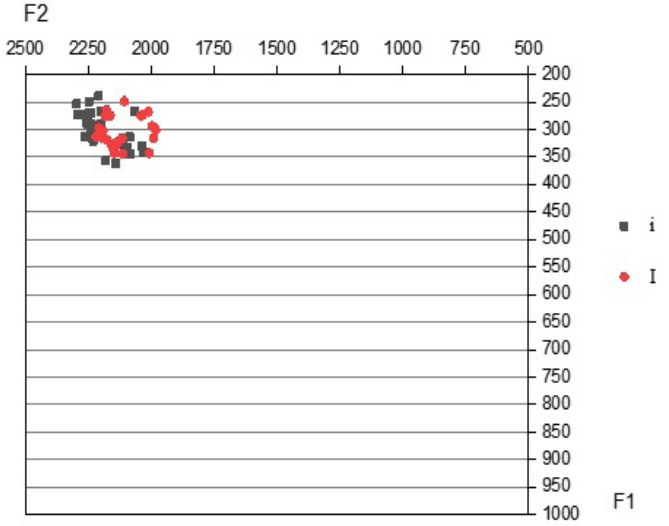
3. RESULTS

The results of the spectral analysis of the vowels follow. In order to establish differences between native and nonnative vowel production, the results of our Serbian EFL speakers will be compared to a group of native speakers of AE who predominantly come from the American Northeast or had spent a large portion of their adult life in this area (Čubrović 2016[5]). Acoustic analysis is followed by a mixed-effects statistical model analysis.

3.1. HIGH FRONT VOWELS /i/ AND /ɪ/

AE high front vowels /i/ and /ɪ/ are described as tense and lax, respectively. Graph 1 shows that high front vowels do not occupy distinct, non-overlapping positions in the acoustic vowel space since the nonnative speakers in this study tend to merge the English vowels /i/ and /ɪ/. This implies that the participants of the Serbian EFL speaker group have not fully acquired the qualitative difference between the vowels /i/ and /ɪ/. A similar result was reported in Čubrović (2016: 114) for a highly proficient group of English speakers of Serbian background with a long-term residence in the US.

Speaker M8 has higher realizations of F1 in all 3 repetitions (357 Hz, 329 Hz, 361 Hz) of the vowel /i/, compared to other respondents. This implies that vowel /i/ is more open in the realization of speaker M8.



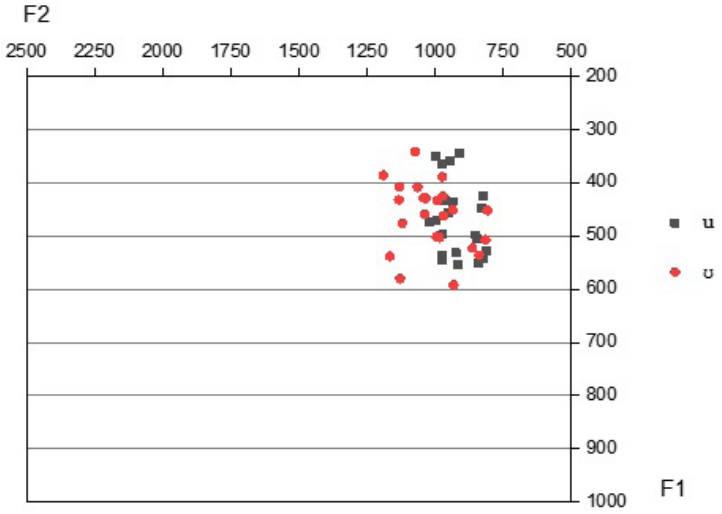
Graph 1. F1 and F2 of /i ɪ/ for NNSs

5 For more information on speakers' regional background, see Čubrović (2016: 98–99).

The mean values of the first two formants for tense /i/ are 298 Hz (F1) and 2191 Hz (F2), while for lax /ɪ/ they are 304 Hz (F1) and 2109 Hz (F2). Compared to reference acoustic data of native speakers in Čubrović (2016), our Serbian EFL participants tend to pronounce vowels /i/ and /ɪ/ with a generally lower F1, implying that their AE tense /i/ and lax /ɪ/ are higher in the vowel space compared to native speakers.

3.2. HIGH BACK VOWELS /u/ AND /ʊ/

Spectral features of high back vowels /u/ and /ʊ/ are analysed in the words *boot* and *put*. In AE, these two vowels are realized as tense and lax, respectively. The vowel /u/ was mispronounced[6] 5 times in this study (1 repetition of speakers M1, M6, M9 and 2 repetitions of speaker M2). These tokens were not included in the analysis, therefore Graph 2 includes 45 repetitions instead of 54 (9 speakers x 3 repetitions for each vowel).



Graph 2. F1 and F2 of /u ʊ/ for NNSs

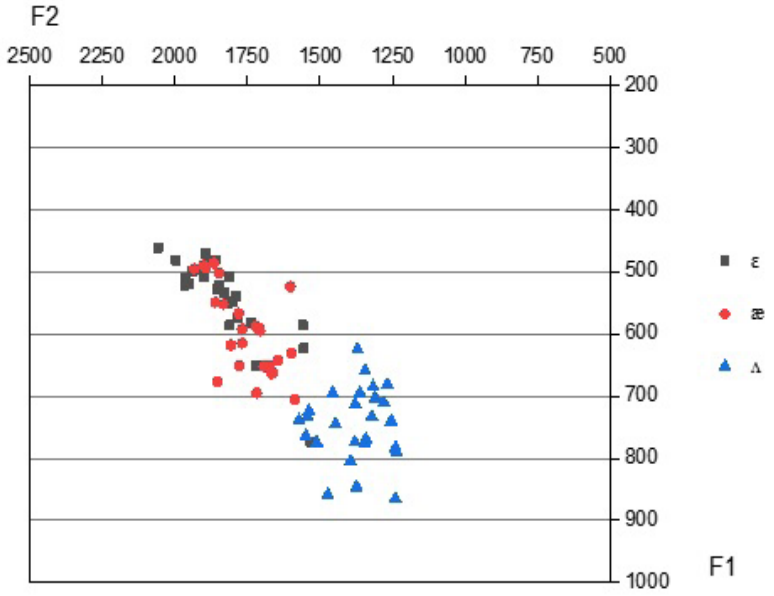
Unlike native speakers who clearly differentiate between these two high back vowels (Čubrović 2016: 103), our Serbian EFL speaker group has a tendency to merge the tense /u/ with the lax /ʊ/, as can be seen in Graph 2. The average formant measurements for /u/ in this study are 467 Hz (F1) and 920 Hz (F2), whereas for /ʊ/ they are 463 Hz (F1) and 1008 Hz (F2). Judging by a negligible difference in the values of F1 for /u/ and /ʊ/, this vowel distinction is not fully acquired in our Serbian EFL group.

6 It was noticed that the vowel /u/ was diphthongized in mispronounced tokens.

3.3. VOWELS /ɛ/, /æ/ AND /ʌ/

Graph 3 illustrates F1 and F2 of vowels /ɛ/, /æ/ and /ʌ/ in the words *bet*, *bat* and *but*. Two mispronounced repetitions of the vowel /æ/ (M2) are not included in Graph 3. Since the majority of our Serbian EFL participants tend to merge the AE vowels /ɛ/ and /æ/, it may be safely assumed that this pair of vowels is an obstacle to a native-like pronunciation of English. Marković (2009: 261) reached a similar conclusion claiming that the "vocalic contrast between the English vowels /ɛ/ and /æ/ poses a serious problem to the native speakers of Serbian."

The vowel /ɛ/ has an average value of 543 Hz for F1 and 1829 Hz for F2, whereas the vowel /æ/ has 595 Hz for F1 and 1754 Hz for F2. Three participants (M3, M4 and M5) had almost identical values of F1 for vowels /ɛ/ and /æ/, while 4 speakers (M1, M6, M7 and M9) had a tendency to pronounce /æ/ with a higher F1 than in /ɛ/. These 4 speakers (M1, M6, M7 and M9) have approximated native-like pronunciation of the vowels /ɛ/ and /æ/ [7].



Graph 3. F1 and F2 of /ɛ/, /æ/ and /ʌ/ for NNSs

Graph 3 does not show overlapping spectral values between /ʌ/ and /æ/. The mean values of the first two formants for /ʌ/ are 745 Hz (F1) and 1374 Hz (F2). A large span in the values of F1 and F2 in /ʌ/ is evident (F1: 625 Hz (M9) to 865 Hz (M5); F2: 1235 Hz (M5) to 1549 Hz (M4)). Both F1 and F2 values representative of the Serbian EFL speakers

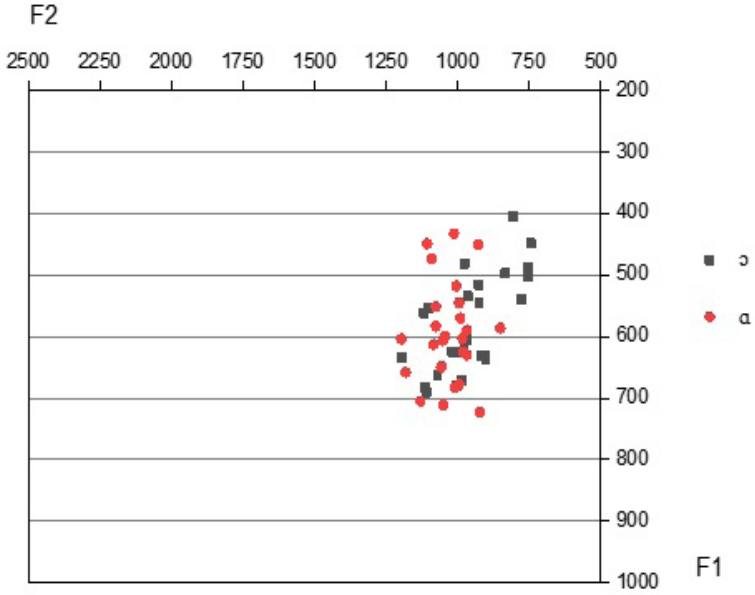
7 Mean values of NSs for /ɛ/ are 632 Hz (F1) and 1593 Hz (F2), whereas for /æ/ they are 733 Hz (F1) and 1568 Hz (F2) (Čubrović 2016: 105).

in this study for the vowel /ʌ/ are higher compared to those of native speakers'[8]. The Serbian EFL speaker group tends to pronounce the vowel /ʌ/ as a more open and less centralized vowel compared to native speakers of AE in Čubrović (2016).

3.4. LOW BACK VOWELS /ɔ/ AND /ɑ/

Even though it is assumed that the low back merger of /ɔ/ and /ɑ/ is well under way and almost completed in many parts of the US (Boberg 2015), Bradlow[9] (1995) and Čubrović[10] (2016) report that some native speakers of AE differentiate the vowels /ɔ/ and /ɑ/ and do not merge them.

In this study, the back vowels /ɔ/ and /ɑ/ are analyzed in words *bought* and *pot*. Both vowels are described as back and tense, but /ɔ/ is realized as a mid-vowel (or more precisely a low-mid vowel), and /ɑ/ is defined as a low vowel (Yavaş 2011: 78–79). Graph 4 shows 52 repetitions instead of 54 (9 participants x 3 repetitions for each vowel), since speaker M5 mispronounced 1 repetition each of vowels *bought* and *pot*. These were not included in the analysis.



Graph 4. F1 and F2 of /ɔ/ and /ɑ/ for NNSs

8 Mean values of NSs for /ʌ/ are 660 Hz (F1) and 1288 Hz (F2) (Čubrović 2016: 108).

9 Bradlow (1995) did not report on the merger of vowels /ɔ/ and /ɑ/, but only the slight increase in both formants (F1 and F2) compared to Peterson and Barney (1952).

10 "We observe that /ɔ/ i /ɑ/ in our data exhibit some degree of neutralization, as dialectal phonetic literature suggests. However, the merger is not completed in the speech of experimental subjects of this study" (Čubrović 2016: 106).

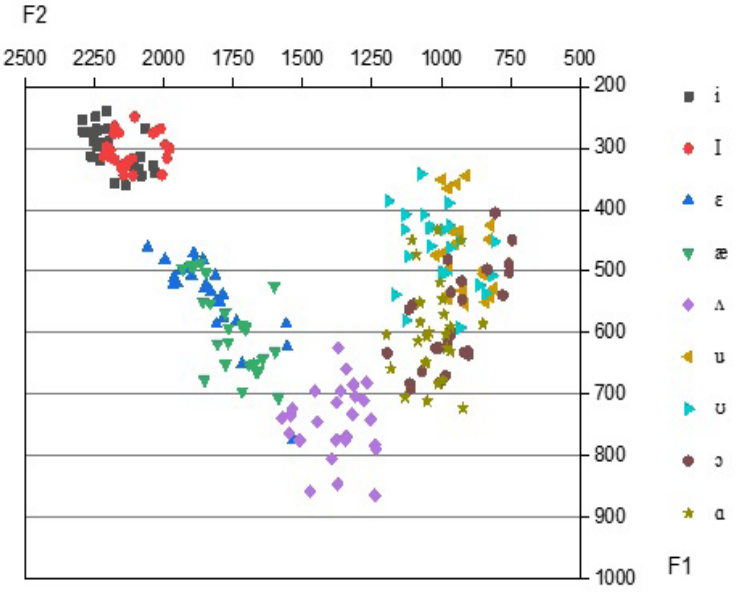
Graph 4 shows that majority of NNSs in this study tend to merge low back vowels, since /ɔ/ and /ɑ/ overlap in the vowel space. The mean formant values of /ɔ/ are 579 Hz (F1) and 954 Hz (F2) and the values of both F1 and F2 are lower than F1 and F2 of /ɑ/. F1 of /ɔ/ spans from 405 Hz (M7) to 690 Hz (M3), while F2 ranges from 744 Hz (M9) to 1195 Hz (M5).

The vowel /ɑ/ has an average value of 595 Hz for F1 and 1031 Hz for F2. F1 of /ɑ/ spans from 433 Hz (M1) and reaches 723 Hz (M6), whereas F2 starts from 851 Hz (M7) and amounts to 1197 Hz (M3).

Compared to reference acoustic data of NS[11] (Čubrović 2016), F1 and F2 of both back vowels /ɔ/ and /ɑ/ are consistently lower when realized by our Serbian EFL participants. This leads us to a conclusion that /ɔ/ and /ɑ/ are higher and more peripheral (back) compared to native speakers.

3.5. ACOUSTIC ANALYSIS SUMMARIZED

Graph 5 illustrates F1 and F2 of all 9 vowels /i ɪ ɛ æ ʌ u ʊ ɔ ɑ/ of AE as realized by 9 Serbian speaking experiment participants.



Graph 5. F1 and F2 of /i ɪ ɛ æ ʌ u ʊ ɔ ɑ/ for NNSs

Table 1 summarizes the results of the spectral analysis of the participants' vowel production with standard deviations given in brackets.

11 Mean values of F1 and F2 of the vowel /ɔ/ are 699 Hz (F1) and 1122 Hz (F2), whereas for /ɑ/ they are 753 Hz (F1) and 1182 Hz (F2) when produced by NS (Čubrović 2016: 107).

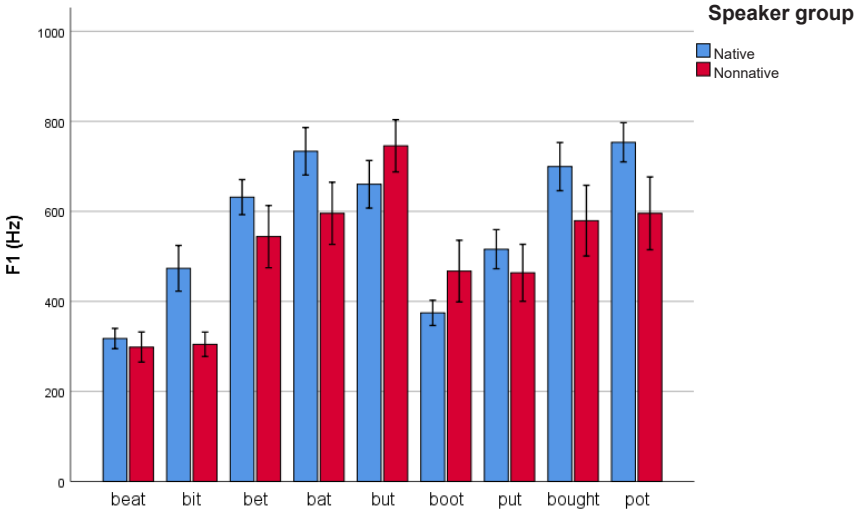
	/i/	/ɪ/	/ɛ/	/æ/	/ʌ/	/u/	/ʊ/	/ɔ/	/ɑ/
F1	298 (33)	304 (27)	543 (69)	595 (69)	745 (58)	467 (68)	463 (63)	579 (78)	595 (80)
F2	2191 (79)	2109 (78)	1829 (133)	1754 (101)	1374 (104)	920 (67)	1008 (110)	954 (123)	1031 (77)

Table 1. Mean values of F1 and F2 (with standard deviation) of 9 AE monophthongs /i ɪ ɛ æ ʌ u ʊ ɔ ɑ/ produced by NNS

4. STATISTICAL ANALYSIS OF NS AND NNS VOWEL PRODUCTION

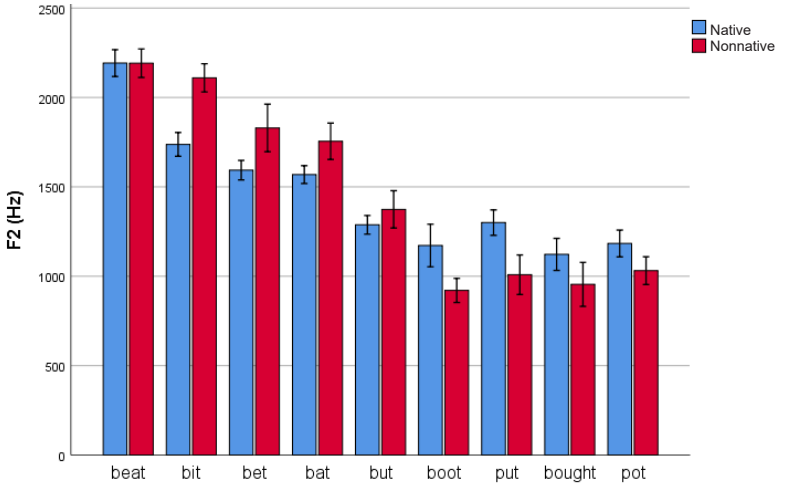
A mixed-effects statistical model was used to analyze and compare the acoustic data (F1 and F2) for native and nonnative speakers of AE. The analysis was run in R statistical software (2013), using Ime4 package (Bates *et al.* 2015). A separate model was run for each formant of each of the 9 monophthongs of English /i ɪ ɛ æ ʌ u ʊ ɔ ɑ/. A fixed effect was *speaker group* (native or nonnative), and *speaker* was a random effect.

The mixed-efffects model found statistically significant differences of speaker group for F1 for all vowels (see Stats tables 2 to 9) except for the AE vowel /i/ (see Stats table 1). Graph 6 shows the influence for speaker group on F1.



Graph 6. F1, Native speakers vs. Nonnative speakers

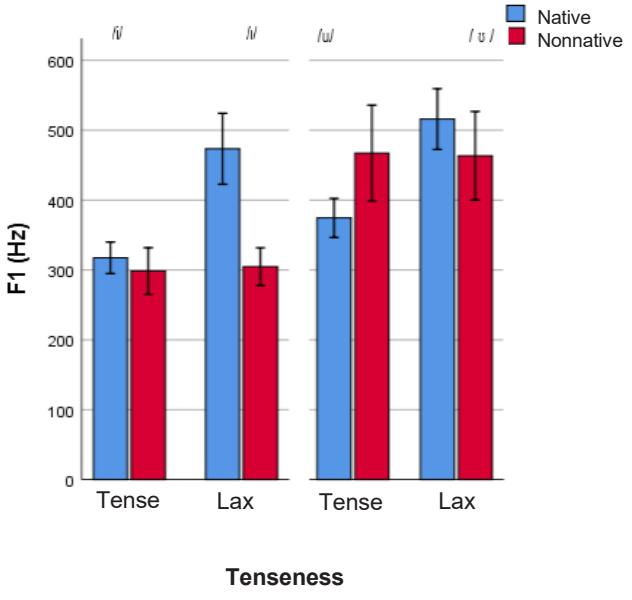
Statistically significant differences of F2 between native and nonnative groups were found for all English vowels (see Stats tables 11 to 18) except /i/ (see Stats table 10). Graph 7 shows the influence of speaker group on the values of F2.



Graph 7. F2, Native speakers vs. Nonnative speakers

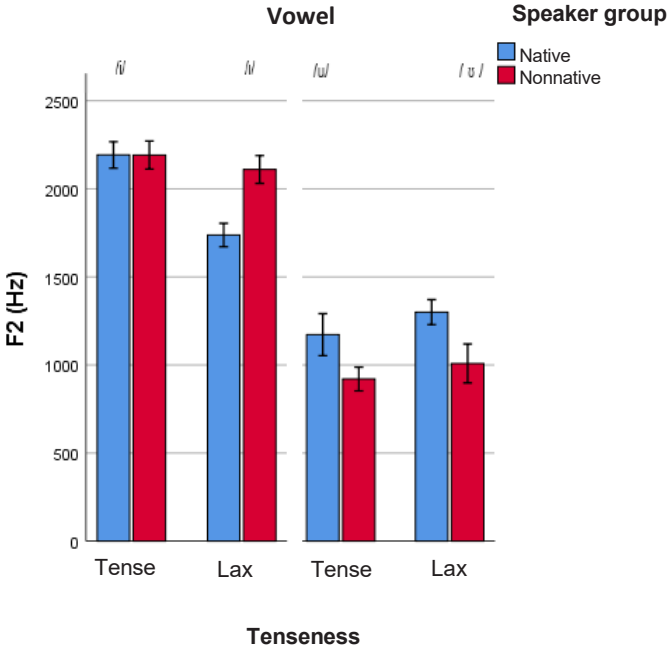
In order to find out if there is a significant difference in the production of tense and lax pairs of vowels - high /i ɪ/ and back /u ʊ/ - between native and nonnative speaker groups, the same mixed-effects model was run in R statistical software (2013), using lme4package (Bates *et al.* 2015). A separate model was run for each formant of each vowel. The fixed effect was *tenseness* (tense/lax), while the random effect was *speaker*.

The model found that there is a significant difference in the values of F1 for both sets of tense/lax vowels (see Stats tables 19 to 22). This is displayed in Graph 8.



Graph 8. F1 of /i ɪ / and /u ʊ/

The model found that there is a significant difference in the values of F2 for both sets of tense/lax vowels (see Stats tables 19 to 22), which is shown in Graph 9.



Graph 9. F2 of /i ɪ / and /u ʊ/

5. DISCUSSION AND CONCLUSIONS

This study examined the production of AE vowels (9 monophthongs) by Serbian EFL speakers in order to investigate their spectral characteristics (F1 and F2) and compare them to those of native speakers. Nine speakers participated in the study and they produced three repetitions of words containing the target vowels. The results of the experiment indicate that vowels /i ɪ/, /ɛ æ/, and /u ʊ/ largely overlap in the vowel space of Serbian EFL learners. This leads to a conclusion that the participants of this study have not acquired the qualitative vowel distinction typical of native speakers' vowels. Some previous studies (Paunović 2011; Čubrović 2016) reached similar conclusions. Studying vowel characteristics of a group of EFL speakers from Niš, Paunović (2011) claims that distinction between /i/ and /ɪ/ is "phonetically rather small" since "these two vowel areas are grouped close together" (Paunović 2011: 364). Similarly, Čubrović (2016) reported that Serbian-American English bilinguals tend to merge vowels /i ɪ/ and /u ʊ/. However, Bjelaković (2018) reached a different conclusion claiming that vowels /i/ and /ɪ/ are well acquired by an EFL group from Belgrade (Bjelaković 2018: 160).

Our experimental study confirmed that the participants struggle with the acquisition of the vowel contrast /ɛ/ - /æ/, as expected. Marković (2009: 261) reached a similar finding

in her study of the Novi Sad nonnative speakers of English. In addition to this, Bjelaković (2018) noticed that young and highly proficient nonnative speakers of English tend to substitute /ɛ/ and /æ/ with the Serbian short /e/. Similarly, Paunović (2011) noted that /æ/ is assimilated into a Serbian vowel category /a/ (Paunović 2011: 364).

Since F2 of front vowels /i ɪ ɛ æ/ is generally higher in the nonnative speakers' production compared to native speakers in the acoustic study done by Čubrović (2016), this implies that the former articulate the vowels /i ɪ ɛ æ/ as fronter.

Our acoustic study shows that Serbian EFL learners tend to merge two back vowels /ɔ ɑ/ in the way that F1 values for these two vowels approximate one another (579 and 595 Hz). Even though there is an acoustic overlap of the vowels /ɔ ɑ/, we cannot safely assume that our participants have acquired the low back merger of /ɔ/ and /ɑ/, which is well under way and almost completed in many parts of the US (Boberg 2015). Additional research with more participants is needed in order to establish whether EFL learners from Čačak have acquired the low back merger of /ɔ/ and /ɑ/.

The mixed-effects statistical analysis found that there is a significant difference in F1 and F2 between native and nonnative speakers for all English vowels except /i/. This means that the acoustic difference in native and nonnative speech is negligible for /i/, while the biggest difference is established in the quality of three vowels /ɪ ʊ ɑ/ (see Stats tables 2-18). Statistical analysis confirms that the quality of the AE vowel /i/ is acquired by our study participants.

On the other hand, Čubrović (2016) reached a slightly different result, since mixed-effects statistical analysis in her study revealed significant differences in F1 between native and nonnative speakers for all English vowels except /ʌ/, whereas significant differences in F2 between native and nonnative speakers were found for all English vowels except /ʌ/ and /i/ (Čubrović 2016: 125–126). It implies that Serbian-American English bilinguals have acquired the quality of vowels /i/ and /ʌ/ unlike our EFL learners who have acquired the quality of /i/ only.

Additionally, in order to compare the production of tense/lax pair of high front /i ɪ/ and high back vowels /u ʊ/ between native and nonnative speakers, the same statistical model was run. This analysis was expected to shed some light on the correlation between vowel length and vowel quality. The results show a statistically significant difference in the values of F1 and F2 for both sets of tense/lax pairs. The most striking statistical difference is observed in the quality of the back vowel pair /u ʊ/ (see Stats tables 19–22). This implies that our participants did not produce the appropriate quality of two sets of tense/lax pairs. The participants of this study seem to rely more on vowel length than on vowel quality.

With respect to the tense/lax pairs of high front /i ɪ/ and high back vowels /u ʊ/, a previous study (Čubrović 2016) reached a conclusion that there were significant differences across the tense/lax high vowel pairs in both native and nonnative speaker (Serbian-American English bilinguals) groups.

The findings reached may be a good starting point for future experimental vowel studies dealing with different aspects of vowel acquisition in the Serbian language context, but also in speakers of other language backgrounds. It could also prove beneficial to analyze vowel quantity and the vowel accommodation mechanisms that speakers use so as to get closer to the vowels of L1.

REFERENCES

Bjelaković, A. 2018. *Vokali savremenog standardnog britanskog izgovora i njihovo usvajanje kod izvornih govornika srpskog jezika*. Filološki fakultet Univerziteta u Beogradu. Unpublished doctoral dissertation. [Internet]. Available at: https://uvidok.rcub.bg.ac.rs/handle/123456789/41/discover [11.05.2019].

Boberg, C. 2015. North American English. In M. Reed and J. M. Levis (eds.) *The Handbook of English Pronunciation*. Malden, MA: Wiley Blackwell, 229–251.

Boersma, P. 2013. Acoustic Analysis. In R. J. Podesva and D. Sharma (eds.) *Research methods in linguistics*. Cambridge: Cambridge University Press, 375–397.

Boersma, P. and D. Weenink. 2013. *Praat: Doing Phonetics by Computer*. Version 5.4. [Internet]. Available at: http://www.fon.hum.uva.nl/praat/ [22.12.2017].

Bradlow, A. R. 1995. A comparative acoustic study of English and Spanish vowels. *Journal of the Acoustical Society of America 97*, 1916–1924.

Crothers, J. 1978. Typology and Universals of Vowel Systems. In J. H. Greenberg *et al.* (eds.) *Universals of Human Languages. Phonology* 2. Stanford, CA: Stanford University Press, 95–192.

Cruttenden, A. 2014. *Gimson's pronunciation of English*. 8th Edition. London/New York: Routledge.

Čubrović, B. 2016. *Acoustic Investigations of Serbian and American English Vowel Inventories*. Belgrade: Faculty of Philology.

Čubrović, B. 2017. Low back merger in native and non-native speakers of American English. *The Linguistics Journal 11(1)*, 222–231.

Čubrović, B. and A. Bjelaković. 2020. Pronunciation model selection, or do you speak American? In B. Čubrović (ed.) *Belgrade English Language and Literature Studies: BELLS 90*. Vol. 1. Belgrade: Faculty of Philology, 139–151.

Dančetović, N. 2017. *Fonetsko-fonološka percepcija i produkcija engleskih monoftonga na tercijarnom nivou*. Filološko-umetnički fakultet Univerziteta u Kragujevcu. Unpublished doctoral dissertation. [Internet]. Available at: https://phaidrakg.kg.ac.rs/detail_object/o:854?tab=0#mda [10.03.2020].

Hillebrand, J. M. *et al.* 1995. Acoustic Characteristics of American English Vowels. *Journal of the Acoustical Society of America 97*, 3099–3111.

Ivić, P. and I. Lehiste. 2002. *O srpskohrvatskim akcentima*. Sremski Karlovci/Novi Sad: Izdavačka knjižarnica Zorana Stojanovića.

Ladefoged, P. and I. Maddieson. 1996. *The Sounds of the World's Languages*. Cambridge, MA/Oxford: Blackwell Publishing.

Maddieson, I. 1984. *Patterns of Sounds, with a chapter contributed by Sandra Ferrari Disner*. Cambridge/New York: Cambridge University Press.

Marković, M. 2007. *Kontrastivna analiza akustičkih i artikulacionih karakteristika vokalskog sistema engleskog i srpskog jezika*. Univerzitet u Novom Sadu. Unpublished doctoral dissertation.

Marković, M. 2009. The perception and production of the English vowels /e/ and /æ/ by native speakers of Serbian. In A. Tsangalidis (ed.) *Selected papers from the 18th ISTAL*. [Internet]. Available at: http://ejournals.lib.auth.gr/thal/article/view/5444/5338 [25.03.2018].

Miletić, B. 1952. *Osnovi fonetike srpskog jezika*. Beograd: Znanje.

Paunović, T. 2002. *Fonetsko-fonološka interferencija srpskog jezika u percepciji i produkciji engleskih vokala*. Univerzitet u Nišu. Unpublished doctoral dissertation.

Paunović, T. 2011. Sounds Serbian? Acoustic properties of Serbian EFL students' speech. In E. Kitis *et al.* (ed.) *Selected Papers from the 19th International Symposium on Theoretical and Applied Linguistics* (ISTAL 19), April 3–5 2009. Thessaloniki: Aristotle University of Thessaloniki, School of English, Department of Theoretical and Applied Linguistics, 357–369.

Peterson, G. E. and H. L. Barney. 1952. Control methods used in a study of the vowels. *Journal of the Acoustical Society of America 24*, 175–184.

Simić, R. and B. Ostojić. 1996. *Osnovi fonologije srpskog književnog jezika*. Beograd: Univerzitet u Beogradu.

Yavaş, M. 2011. *Applied English Phonology*. 2nd Edition. Chichester, UK: Wiley-Blackwell.

Zsiga, E. C. 2013. *The Sounds of Language. An Introduction to Phonetics and Phonology*. Chichester, UK: Wiley–Blackwell.

APPENDIX

Statistical tables: NS and NNS

Random effects:					
			Variance	St. deviation	
Participants (Intercept)			531.1	23.05	
Residuals			324.0	18.00	
Fixed effects:					
	Estimate	Std. Error	df	t value	Pr(>|t|)
Intercept	317.446	8.427	16.000	37.672	<2e-16 ***
Speaker group (Nonnative)	-18.891	11.917	16.000	-1.585	0.132

Stats table 1. F1, /i/

Random effects:					
			Variance	St. deviation	
Participants (Intercept)			1299.5	36.05	
Residuals			489.5	22.12	
Fixed effect:					
	Estimate	Std. Error	df	t value	Pr(>|t|)
Intercept	473.51	12.75	15.73	37.143	< 2e-16 ***
Speaker group (Nonnative)	-166.58	18.12	16.00	-9.192	8.76e-08 ***

Stats table 2. F1, /ɪ/

Random effects:					
			Variance	St. deviation	
Participants (Intercept)			2225	47.17	
Residuals			1097	33.12	
Fixed effect:					
	Estimate	Std. Error	df	t value	Pr(>|t|)
Intercept	631.61	16.97	16.00	37.225	< 2e-16 ***
Speaker group (Nonnative)	-87.79	24.00	16.00	-3.659	0.00212 **

Stats table 3. F1, /ɛ/

Random effects:					
			Variance	St. deviation	
Participants (Intercept)			3140	56.04	
Residuals			1016	31.88	
Fixed effects:					
	Estimate	Std. Error	df	t value	Pr(>|t|)
Intercept	733.42	19.66	15.24	37.303	< 2e-16 ***
Speaker group (Nonnative)	-131.53	27.93	15.49	-4.709	0.000257 ***

Stats table 4. F1, /æ/

Random effects:					
			Variance	St. deviation	
Participants (Intercept)			2350.4	48.48	
Residuals			913.8	30.23	
Fixed effects:					
	Estimate	Std. Error	df	t value	Pr(>|t|)
Intercept	660.07	17.18	16.00	38.430	< 2e-16 ***
Speaker group (Nonnative)	85.45	24.29	16.00	3.518	0.00285 **

Stats table 5. F1, /ʌ/

Random effects:					
			Variance	St. deviation	
Participants (Intercept)			1632	40.40	
Residuals			1019	31.92	
Fixed effects:					
	Estimate	Std. Error	df	t value	Pr(>|t|)
Intercept	374.44	14.80	15.33	25.298	6.38e-14 ***
Speaker group (Nonnative)	93.74	21.23	16.12	4.415	0.000426 ***

Stats table 6. F1, /u/

Random effects:					
			Variance	St. deviation	
Participants (Intercept)			1063	32.61	
Residuals			1869	43.23	
Fixed effects:					
	Estimate	Std. Error	df	t value	Pr(>|t|)
Intercept	515.95	13.69	15.14	37.696	<2e-16 ***
Speaker group (Nonnative)	-54.88	19.71	16.18	-2.784	0.0132 *

Stats table 7. F1, /ʊ/

Random Effects:					
			Variance	St. deviation	
Participants (Intercept)			2607	51.05	
Residuals			2067	45.47	
Fixed effects:					
	Estimate	Std. Error	df	t value	Pr(>|t|)
Intercept	699.39	19.14	15.90	36.55	< 2e-16 ***
Speaker group (Nonnative)	-118.84	27.13	16.05	-4.38	0.000463 ***

Stats table 8. F1, /ɔ/

Random effects:					
			Variance		St. deviation
Participants (Intercept)			1952		44.18
Residuals			2364		48.62
Fixed effects:					
	Estimate	Std. Error	df	t value	Pr(>|t|)
Intercept	753.38	17.45	15.85	43.181	< 2e-16 ***
Speaker group (Nonnative)	-156.69	24.76	16.05	-6.328	9.88e-06 ***

Stats table 9. F1, /ɑ/

Random effects:					
			Variance		St. deviation
Participants (Intercept)			4952		70.37
Residuals			1423		37.72
Fixed effects:					
	Estimate	Std. Error	df	t value	Pr(>|t|)
Intercept	2192.1369	24.5549	16.0000	89.275	<2e-16 ***
Speaker group (Nonnative)	-0.5073	34.7258	16.0000	-0.015	0.989

Stats table 10. F2, /i/

Random effects:					
			Variance		St. deviation
Participants (Intercept)			3697		60.80
Residuals			1785		42.25
Fixed effects:					
	Estimate	Std. Error	df	t value	Pr(>|t|)
Intercept	1737.37	21.84	15.96	79.56	< 2e-16 ***
Speaker group (Nonnative)	372.50	31.07	16.27	11.99	1.71e-09 ***

Stats table 11. F2, /ɪ/

Random effects:					
			Variance		St. deviation
Participants (Intercept)			9971		99.85
Residuals			1133		33.66
Fixed effects:					
	Estimate	Std. Error	df	t value	Pr(>|t|)
Intercept	1593.16	33.91	16.00	46.983	< 2e-16 ***
Speaker group (Nonnative)	236.25	47.95	16.00	4.926	0.000152 ***

Stats table 12. F2, /ɛ/

Random effects:					
			Variance		St. deviation
Participants (Intercept)			5760		75.90
Residuals			1441		37.96
Fixed effects:					
	Estimate	Std. Error	df	t value	Pr(>|t|)
Intercept	1568.62	26.33	15.16	59.570	< 2e-16 ***
Speaker group (Nonnative)	176.23	37.38	15.36	4.715	0.00026 ***

Stats table 13. F2, /æ/

Random effects:					
			Variance		St. deviation
Participants (Intercept)			5578		74.69
Residuals			1657		40.71
Fixed effects:					
	Estimate	Std. Error	df	t value	Pr(>|t|)
Intercept	1287.64	26.10	16.00	49.336	<2e-16 ***
Speaker group (Nonnative)	86.51	36.91	16.00	2.344	0.0323 *

Stats table 14. F2, /ʌ/

Random effects:					
			Variance		St. deviation
Participants (Intercept)			7008		83.72
Residuals			3094		55.63
Fixed effects:					
	Estimate	Std. Error	df	t value	Pr(>|t|)
Intercept	1171.97	29.89	15.75	39.212	< 2e-16 ***
Speaker group (Nonnative)	-251.86	42.74	16.39	-5.893	2.06e-05 ***

Stats table 15. F2, /u/

Random effects:					
			Variance		St. deviation
Participants (Intercept)			4601		67.83
Residuals			4023		63.43
Fixed effects:					
	Estimate	Std. Error	df	t value	Pr(>|t|)
Intercept	1300.09	25.70	15.47	50.597	< 2e-16 ***
Speaker group (Nonnative)	-291.45	36.76	16.16	-7.927	5.82e-07 ***

Stats table 16. F2, /ʊ/

Random effects:					
			Variance	St. deviation	
Participants (Intercept)			7033	83.86	
Residuals			5117	71.53	
Fixed effects:					
	Estimate	Std. Error	df	t value	Pr(>|t|)
Intercept	1122.16	31.16	15.71	36.012	< 2e-16 ***
Speaker group (Nonnative)	-163.87	44.18	15.86	-3.709	0.00193 **

Stats table 17. F2, /ɔ/

Random effects:					
			Variance	St. deviation	
Participants (Intercept)			1701	41.24	
Residuals			4267	65.33	
Fixed effects:					
	Estimate	Std. Error	df	t value	Pr(>|t|)
Intercept	1183.23	18.63	15.95	63.514	< 2e-16 ***
Speaker group (Nonnative)	-152.34	26.48	16.22	-5.752	2.81e-05 ***

Stats table 18. F2, /ɑ/

Statistical tables: NNS English production of /i ɪ/ and /u ʊ/

Random effects:					
			Variance	St. deviation	
Participants (Intercept)			2107	45.90	
Residuals			2245	47.38	
Fixed effects:					
	Estimate	Std. Error	df	t value	Pr(>|t|)
Intercept	308.37	14.28	24.64	21.595	< 2e-16 ***
Tenseness	81.69	13.05	33.46	6.261	4.23e-07 ***

Stats tables 19. F1, /i ɪ/, nonnative speakers

Random effects:					
			Variance	St. deviation	
Participants (Intercept)			0	0.00	
Residuals			3554	59.62	
Fixed effects:					
	Estimate	Std. Error	df	t value	Pr(>|t|)
Intercept	402.51	12.17	47.00	33.076	< 2e-16 ***
Tenseness	86.10	17.04	47.00	5.053	7.02e-06 ***

Stats table 20. F1, /u ʊ/, nonnative speakers

Random effects:					
			Variance	St. deviation	
Participants (Intercept)			4071	63.8	
Residuals			18629	136.5	
Fixed effects:					
	Estimate	Std. Error	df	t value	Pr(>|t|)
Intercept	2184.33	30.53	34.80	71.55	< 2e-16 ***
Tenseness	-261.31	37.55	35.04	-6.96	4.28e-08 ***

Stats table 21. F2, /i ɪ/, nonnative speakers

Random effects:					
			Variance	St. deviation	
Participants (Intercept)			24982	158.06	
Residuals			5872	76.63	
Fixed effects:					
	Estimate	Std. Error	df	t value	Pr(>|t|)
Intercept	1047.21	41.57	18.62	25.194	7.65e-16 ***
Tenseness	106.45	22.10	31.15	4.816	3.60e-05 ***

Stats table 22. F2, /u ʊ/, nonnative speakers

SUMMARY

ACOUSTIC CHARACTERISTICS OF AMERICAN ENGLISH MONOPHTHONGS IN SERBIAN EFL SPEAKERS – A CASE STUDY

The purpose of this study is to investigate the acoustic realization of American English vowels in a group of Serbian EFL learners from an urban area in Central Serbia and compare them to the native speakers' vowel production. Recordings were made of 9 male speakers producing nine monophthongs of American English /i ɪ ɛ æ ʌ u ʊ ɔ ɑ/ in the CVC phonological structure, flanked by a labial consonant /p/ or /b/ and coronal /t/. These vowel realizations were compared to those of 9 male native speakers of American English who participated in a similar experiment (Čubrović 2016). The analysis of the spectral properties of nonnative English vowels shows that the influence of Serbian is evident in the production of some vowel pairs: /i ɪ/, /ɛ æ/ and /u ʊ/. Our Serbian EFL speaker group shows a tendency to merge two back vowels, /ɔ ɑ/, since F1 values for these two vowels are close. The mixed-effects statistical analysis found that there is a significant difference for F1 and F2 between native and nonnative speakers for all vowels except /i/ for the tense /i/.

KEYWORDS: vowels, American English, Serbian EFL speakers, vowel quality.

ARTICLE INFO:
Original research article
Received: August 1, 2021
Revised: November 13, 2021
Accepted: November 22, 2021