Sounds Serbian? Acoustic properties of Serbian EFL students' speech

Tatjana Paunović
University of Niš
tpaunovic@gmail.com

Abstract: This paper presents the findings of a study of the acoustic properties of Serbian EFL students' vowels. The participants were 12 junior-year students of the English Department. Their vowel production was recorded in three different contexts, i.e. speaking tasks: reading words in citation form, reading a text aloud, and speaking. The acoustic measurements included vowel duration and F1 and F2 formant frequency values. The results showed that neither the production of vowel qualities nor the duration differences used by the students were without problems.
Key words: EFL students' vowels, acoustic properties, formant frequency, duration

1. Introduction

The relatively rich vowel system of English poses a challenge to many EFL learners of different L1 backgrounds, so the acquisition of English vowels has been researched profusely. Nevertheless, many issues remain unexplained. For instance, it is still not clear whether English vowels are more challenging for students with smaller L1 vowel inventories, such as Serbian, or for those with more elaborate L1 vowel systems. Similarly, it is not clear whether vowel duration is more easily acquired in L2 if it plays a phonological role in L1, or not (Bohn & Flege 1990; Flege et al. 1997; Cebrian 2006). Some research results indicate that even in 'pure' vowels the dynamic patterns of formant movement may be part of category representations (Pisoni 1997; Kewley-Port & Goodman 2005; Iverson & Evans 2007; Neel 2008). Furthermore, it has been confirmed (Hillenbrand et al. 2000, 2001) that not only the immediate phonetic context but also the broader context can largely influence vowels' acoustic shape, particularly in back vowels (Strange et al. 2007: 1126).

Bearing all this in mind, it is not surprising that acquiring adequate vowel contrasts in L2 should be such a challenging task. At the same time, it is a very important one, considering the role played by vowels' spectral and temporal properties in overall speech intelligibility (Kewley-Port et al. 1996) and the perceived degree of 'accentedness' in L2 speech (Flege 1988; Fernandez Guerra 2001; Major 2002; Munro et al. 2006; Munro & Derwing 2001, 2008). Some findings show that segmental properties, particularly vowel quality and quantity, have even more influence than suprasegmentals in the perception of 'accented speech' (Fernandez Guerra 2001:154; Neel 2008). Still, as pointed out by MacKay and Fullana (2007:332), the results of intelligibility and 'accentedness' investigations are particularly inconclusive for L2 learners in formal settings, so much more research is needed in this area.

A substantial body of research has explored the role played by the L1 phonological system in the perception and production of L2 vowels, highlighted in particular by the Speech Learning Model (Flege 1995, 1997) and the Perceptual Assimilation Model (Best 1994, 1995). As both these models point out, L2 learners fail to perceive and produce the appropriate vowel identities because L2 speech is 'filtered' through the prism of their L1 categories, that is, learners perceive the "similarities to, and discrepancies with, the native segmental constellations that are in the closest proximity

to them in native phonological space" (Best 1995: 193), which can explain "perception and production problems in differentiating these categories" (Strange et al. 2007: 1127). However, most studies referring to these two models, as well as to the somewhat broader Environmental approach (Jia & Aaronson, 2003; Jia et. al. 2006), target bilinguals or L2 learners in naturalistic settings (e.g. Piske et al. 2002; Guion 2003), investigating various factors that affect their perception and production of L2 vowels and their quality and quantity. For instance, Chen (2006) investigated the contrasts produced between tense and lax vowels by Mandarin speakers of English, and showed that they relied much more on the temporal contrast (Chen 2006: 248). Munro & Derwing (2008) conducted a longitudinal study of the segmental production of adult ESL learners, L1 speakers of Mandarin and Slavic languages, again in a naturalistic setting, in Canada. The study showed that all the participants made improvement in vowel intelligibility over one year, though the improvement was not linear (Munro & Derwing 2008:448).

The research with EFL students in formal settings is less abundant, and has mainly focused on a limited number or segmental contrasts. For instance, Tsukada (2008) observed the production of four monophthongs and two diphthongs by Thai speakers, while Barboza (2007) investigated the properties of front vowels produced by Brazilian learners of English.

With Serbian EFL students, this kind of research is even more limited. Apart from several early studies (Mihajlović 1957, 1969; Djokić 1981, 1984), there are few recent studies of English vowel production or perception by Serbian learners. The extensive and detailed research of Krebs-Lazendic (Krebs-Lazendic 2008, Krebs-Lazendic & Best 2007) investigates the acquisition of English vowels by Serbian L1 speakers in a naturalistic setting in Australia, while Serbian EFL students' vowels in formal settings were investigated by Paunović (2003), Marković (2007, 2009), and Čubrović (2007), and, with young learners, Savić (2007, 2009). Some of the most relevant findings of these few studies indicate that vowel distinctions are very difficult to acquire by Serbian L2 students in formal settings, even after years of studying English as a foreign language, and even by students of otherwise rather high proficiency levels, for instance, English department students (e.g. in Marković 2007).

2. Present study

The purpose of this study was to investigate the acoustic properties of Serbian EFL students' vowels highlighted as problematic by both classroom experience and previous research. Specifically, we focused on the duration and formant structure of the English vowel contrasts involving the 'short' vowels /ɪ, ʌ, ɒ, ʊ/ and their 'long' pairs /iː, ɑː, ɔː, uː/, as well as the /e/-/æ/ and /ɜː/ - /ə/ distinctions. Our aim was to investigate (1) whether Serbian EFL students produced phonetic contrasts (F1, F2) between the relevant vowel categories in English, and (2) whether they used the phonetic cue of duration to support relevant vowel contrasts.

However, in difference to some of the studies mentioned above, this study did not aim to compare the properties of the vowels produced by our participants with English native-speakers' or L1 Serbian vowels, nor to examine the source or degree of mother tongue interference. Rather, in accordance with views proposed by Lindblom et al. (1992) or Neel (2008), we focused on the participants' vowel *system* as such, and on the relationships between the vowel categories *within* the system, that is, on observing whether the students' categories unambiguously stood as clearly differentiated areas of vowel space, distinct from all the others, and whether they were stable across different contexts and speaking styles. We opted for this approach because, as pointed out by

Neel (2008:584), "distinctiveness among neighboring vowels" and "acoustic comparison of confused vowels" are of crucial importance in determining vowel *intelligibility*. Therefore, although in the discussion below we do list the data offered in literature on English and Serbian vowels, we did not aim to make explicit comparisons of our participants' vowel qualities with these 'reference' formant values, but, rather, to observe the differences between 'neighbouring' vowels in our participants' vowel space i.e. vowel system.

2.1. English and Serbian vowels

English vowels have been investigated in numerous studies that present data about their formant frequency values. *Table 1* summarizes those offered by Wells (1962), Deterding (1990, in Marković 2007), Hawkins & Midgley (2005) and Deterding (2006) for British English, as well as by Peterson & Barney (1952), Denes and Pinson (1993) and Ladefoged (2001) for American English. In these studies, the participants were male speakers, apart from Deterding's, where female participants were included, too. The data from the Hawkins and Midgley's study (2005) are from the speakers aged 20 to 25, the age group closest to our participants.

Table 1. Mean F1 and F2 values of English vowels cited in 1) Wells 1962, 2) Deterding 1990, 3) Hawkins & Midgley 2005, 4) Deterding 2006, 5) Peterson & Barney 1952, 6) Denes and Pinson 1993 and 7) Ladefoged 2001

	1)male		2) male		2) female		3) male		4) male		5) male		6) male		7) male	
	F1	**F2**	**F1**	**F2**	**F1**	**F2**	**F1**	**F2**	**F1**	**F2**	**F1**	**F2**	**F1**	**F2**	**F1**	**F2**
iː	285	2237	275	2221	319	2723	276	2338	296	2241	270	2290	270	2290	280	2250
ɪ	356	2098	382	1958	432	2296	393	2174	396	1839	390	1990	390	1990	400	1920
e	569	1965	560	1797	645	2287	600	1914	532	1656	530	1840	500	1880	550	1770
æ	748	1746	737	1527	1011	1759	917	1473	667	1565	660	1720	660	1720	690	1660
ʌ	722	1236	695	1224	813	1422	658	1208	661	1296	520	1190	/	/	/	/
ɑː	677	1083	687	1077	779	1181	604	1040	680	1193	/	/	730	1100	710	1100
ɒ	599	891	593	866	602	994	484	865	643	1019	730	1090	450	850	/	/
ɔː	449	737	453	642	431	799	392	630	480	857	570	840	390	1050	590	880
ʊ	376	950	414	1051	414	1203	413	1285	395	1408	440	1020	300	850	450	1030
uː	309	939	302	1131	339	1396	289	1616	386	1587	300	870	490	1350	310	870
ɜː	581	1381	513	1377	650	1593	494	1373	519	1408	490	1350	/	/	/	/

The Serbian vowel inventory includes 5 vowels: two high vowels, the front and unrounded /i/ and the back and rounded /u/; two mid-vowels, the front unrounded /e/ and the back rounded /o/, and an unrounded open vowel, most often described as 'central', but 'inclined towards a back articulation' (Simić & Ostojić 1989:179). Each short vowel is matched by a long counterpart in the context of the system of *prosodic accents* (Lehiste & Ivić 1986; Simić & Ostojić 1989; Ivić1994). It has been noted that although small vowel systems are expected to have their vowels evenly distributed and maximally dispersed over the vowel space, research shows that it may not always be so (Butcher 1994:28, 32), since many three- and five-vowel systems are organized by a diametrically opposite principle, the one of *minimal distinctiveness*, so that vowels occupy a smaller and more compact part of the available phonetic vowel space. This, however, doesn't seem to be typical of the Serbian vowel system, judging by the average F1 and F2 values of the 5 Serbian vowels offered by Simić and Ostojić (1989) and by Lehiste and Ivić (1996), shown in *Table 2*.

Table 2. Mean F1 and F2 values of Serbian vowels (by male speakers) cited in 1) Simić & Ostojić 1989 and 2) Lehiste and Ivić 1996

	1)		2) in short syllables		2) in long syllables	
	F1	F2	F1	F2	F1	F2
i/	250	2350	410	2045	385	2155
/e/	550	1900	550	1760	480	1985
/a/	775	1250	780	1445	900	1525
/o/	550	900	560	1035	495	960
/u/	375	750	455	820	455	775

2.2. Participants

The participants in this research were twelve students of the English Department at the Faculty of Philosophy, University of Niš. They were 1st year students (average age 19.3, proficiency level B2+ CEF), just starting the introductory course in *English Phonetics and Phonology*. The group consisted of eight female and four male students, all Serbian L1 speakers. It was not possible to estimate the degree of their exposure to particular varieties of English (British, American), given that in the formal educational context in Serbia standard British still seems to be the predominant variety for in-class work, whereas students are largely exposed to other English varieties outside the classroom.

2.3. Methodology and procedures

For data gathering, we devised three different tasks of varying degrees of control, constraint, and structure, in order to observe how participants used phonetic cues across different speech styles and contexts.

Task one focused on vowels used in individual words, in citation form. Thirty-two words were put on flashcards (*beach, teacher, seat, sit, pick, sister, desk, bed, cat, Pat, butter, cut, father, bath, hard, dog, spot, top, saw, bought, thought, good, book, put, school, shoes, bird, herb, girls, under, about, Abbott*). All the words but four were monosyllabic, all with the target vowel in the stressed syllable, except for *schwa*. Each vowel was illustrated at least twice, and all the tokens were used in the carrier sentence *'I see ___ there'*. The purpose of this task was to provide clear examples of students' vowels in a maximally controlled context.

Task two consisted in reading aloud *The Story of Arthur the Rat*, frequently used for phonetic studies because it offers diverse examples of all the English speech sounds. The story was taken from Markham & Hazan (2002:16) and only slightly adapted. The participants were instructed to relax and read the text naturally. The purpose of this task was to provide examples of students' vowels produced in a strictly structured context, but in stretches longer than one utterance. The tokens selected for measurement were in the focus position of their tone units (TU) wherever possible (except for *schwa*); some of them were also TU-final.

In *Task three,* the participants were asked to briefly retell the story from memory. The purpose was to provide examples of students' vowels used in the least structured context, as similar as possible to spontaneous speech ('semi-spontaneous' speech, according to Markham & Hazan 2002).

The recordings were made in the computer room of the Faculty of Philosophy, University of Niš. Each participant's performance was recorded digitally, using a Phillips SBC MD650 microphone, directly into the *Speech Filing System* 4.7/Windows 2008 (© M. Huckvalle, UCL). The measurements included vowel formant frequency values (F1, F2), and vowel duration. Vowel duration measures were based on manual segmentation, relying on both the waveform and the wideband spectrogram. Formant frequencies were taken from the steady-state portion of the vowel. All the formant

values were first automatically estimated by the program and then re-checked manually. Since our aim was to investigate the interlanguage vowel system of EFL *learners*, which we expected to be imperfect in many ways and different from the native-speaker systems, we did not exclude from the analysis those tokens that were obviously different from the expected vowel qualities based on native-speaker data. In task three, taking into consideration the possible acoustic effects of coarticulation in connected speech, we tried to control for this, to the extent to which it was possible, by choosing the tokens that provided a similar phonetic environment to the tokens in the citation and reading tasks (Fabricius 2002:219).

3. Results and discussion

3.1. Vowel quality – formant frequencies

The quality measurements of the participants' vowels are summarized in *Table 3*. It shows the mean F1 and F2 values and standard deviation for male and female speakers in the three different tasks: words in citation form in the carrier sentence, reading the text, and retelling the story in semi-spontaneous speech.

Table 3. Mean F1 & F2 values and standard deviation for pure vowels for male and female participants, a) citation, b) reading, c) speaking

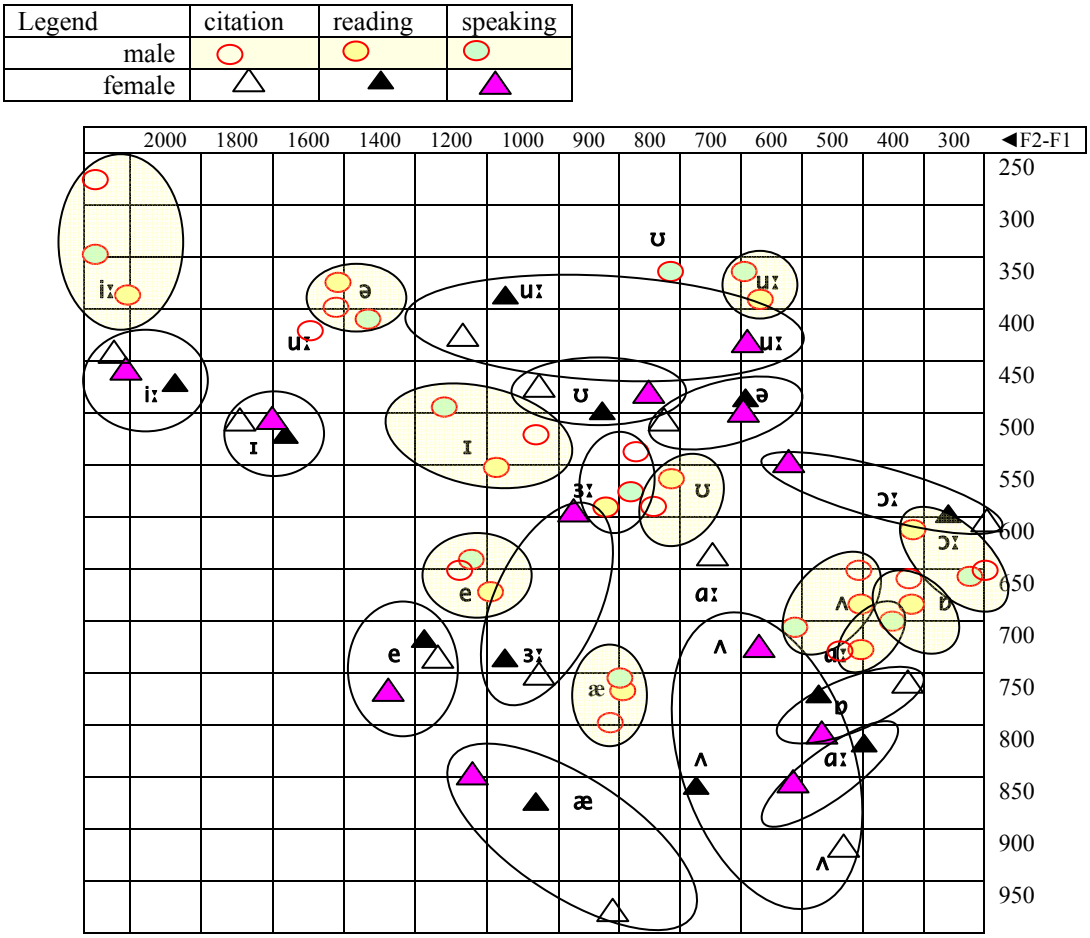
male	iː	ɪ	e	æ	ʌ	ɑː	ə	ɜː	ɒ	ɔː	ʊ	uː
a) F1	269	534	647	810	649	742	384	559	640	664	590	405
F2	2389	1473	1708	1620	1049	1181	1855	1327	989	942	1367	1924
b) F1	388	556	678	786	679	726	384	597	672	608	560	388
F2	2464	1549	1723	1601	1082	1130	1855	1466	1026	962	1326	1047
c) F1	339	487	627	753	706	757	401	586	695	667	369	369
F2	2415	1666	1684	1589	1233	1129	1787	1382	1078	987	1076	1002
Std F1	59.80	35.25	25.69	28.62	28.51	15.50	9.82	19.55	27.62	33.23	119.87	18.00
Std F2	38.08	97.22	19.67	15.63	98.10	29.74	39.26	70.00	44.71	22.55	157.51	519.81

female	iː	ɪ	e	æ	ʌ	ɑː	ə	ɜː	ɒ	ɔː	ʊ	uː
a) F1	457	542	745	1051	932	644	512	745	762	593	459	440
F2	2475	2169	1932	1881	1373	1288	1252	1729	1084	844	1440	1459
b) F1	484	534	725	885	856	808	488	747	785	598	499	471
F2	2417	2125	2016	1846	1448	1214	1094	1740	1284	898	1362	1527
c) F1	445	534	780	845	735	847	498	579	847	543	483	436
F2	2453	2122	2038	1923	1291	1359	1100	1484	1295	1010	1263	1034
Std F1	19.97	4.61	27.84	109.23	99.35	107.72	12.06	96.42	43.97	30.41	20.13	19.16
Std F2	29.28	26.31	55.94	38.55	78.53	72.51	89.54	144.73	118.77	84.67	88.71	267.17

Figure 1 shows formant values plotted onto the vowel space graph, where the vertical axis represents F1 values and vowel differences in the degree of openness/ height, while the horizontal axis represents the differences in the front/back dimension, that is, the difference between F1 and F2, which, according to Ladefoged (1993:199), yields a more precise representation of back vowels, since "the so-called front-back dimension has a more complex relationship to the formant frequencies", and some of the effects of the lip rounding accompanying the articulation of back vowels can be eliminated "by considering the second formant in relation to the first. The degree of backness is best related to the difference between the first and the second formant frequencies" (Ladefoged 2001:177). As suggested by Fabricius (2002:220), too, the F2-F1 value is conveniently used "as a representation of the peripherality of the vowel tokens," since "more peripheral vowels have a higher value for F2-F1", while,

conversely, "more central vowels have a lower value, since a higher F1 value is subtracted from a lower F2 one". She points out that this representation has the advantage over the traditional one since this latter one "would not take relative vowel height into account".

Figure 1. Mean formant values plotted for male participants (circles) and female participants (triangles), in the three tasks / contexts



Since the study included a relatively small number of participants, and a relatively small number of vowel tokens for each of the three speaking styles, in data analysis we relied mostly on descriptive statistics and simple comparisons of the raw formant data to observe inter- and intra-speaker variation in vowel qualities across the three speech styles, rather than on more complex statistical procedures. For instance, one-way ANOVA comparisons showed no statistically significant differences across the three speaking styles for either male or female speakers, and the only point that even roughly approximated statistical significance was the difference between the F2 values of the citation-style and semi-spontaneous speech with 4 male speakers (/ʊ/ p= .067, /uː/ p= .090).

Standard deviation values, though, did show some obvious variation in the quality of the vowels produced in citation, reading and speaking, both for individual speakers and for the mean values for male and female speakers (presented in *Table 1*). First, with respect to the dimension of vowel openness, more variability was observed in female than in male speakers. Whereas male speakers produced only one vowel, /ʊ/, as remarkably more open in the two structured contexts and as much more close i.e.

similar to the category of /u:/ in semi-spontaneous speech, female participants produced observable variation with respect to the degree of openness in four vowels /æ, ʌ, ɑː, ɜː/. Namely, both /æ/ and /ʌ/ were notably more open in citation form of words than in either reading or speaking, while /ɜː/ was obviously more open in citation and reading, and much more close in spontaneous speech. The degree of openness of /ɑː/ with female speakers showed a tendency to centralize the vowel in citation form, while it was much more open in both reading and speaking.

With respect to the front/back dimension (F2) male speakers generally showed less variation than female speakers. Still, with male speakers /ɪ/ was produced as more front and less centralized in the speaking task, as compared to the two more controlled contexts, that is, as closer to /i:/. Similarly, /ʊ/ was more 'carefully' centralized in citation and reading than in speaking, where it was produced in the general area of /u:/ With /u:/, surprisingly, three out of four male speakers showed a tendency to front the vowel conspicuously, in different vowel tokens and particularly in the citation style. Hence the highest standard deviation value for the front/back dimension with this vowel in the male participants. We cannot explain this finding without further investigation, but the fact that five out of eight female speakers showed the same tendency (in citation and reading), only to a somewhat lesser degree, indicates that this should be treated as a feature of our participants' interlanguage vowel systems rather than as a random mistake or an error in the data. Although the centralization of /u:/ has long been acknowledged as a phonetic change in some British varieties (Gimson 1970:120; Marković 2009:11), the inconsistencies in the realization of this vowel in different contexts must be interpreted as a sign that this vowel category has not been fully acquired and stabilized in our participants' interlanguage systems..

In addition to this, the female participants also showed variability in the dimension of peripherality with respect to /ʌ, ɑː, ɒ, ɔː, ʊ, ə/, as indicated by standard deviation values across the three speaking styles. Some of this variability was in accordance with our expectations, based on previous research with Serbian EFL students (Marković 2007, 2009, Paunović 2003), that the participants would assimilate the categories which are not yet acquired or stabilized in their vowel systems to the 'nearest' L1 vowel categories. For instance, /ɔː/ was produced as much more back in citation and reading than in speaking, where it is remarkably centralized. The same tendency was observed for /ɒ/, produced in the citation form as more peripheral (back) than in reading and speaking, where it is more central and closer to the qualities of our participants' /ɑː/. However, some vowel qualities produced here cannot be explained in this way. Namely, the quality of /ɑː/ produced in the citation form was much more close than any variety of a-vowels in Serbian or English.

From the point of view of L2 pronunciation teaching, if intelligibility is set as an attainable goal for students, most attention should be paid to distinguishing the relevant pairs of English vowel categories, that is, the 'neighbouring' categories in the vowel space, since a sufficient difference from the 'neighbouring' vowel categories is essential for intelligibility, as pointed out by Neel (2008:584). In this respect, we could say that our participants did make an attempt to categorize their English vowel space into twelve distinct vowel qualities. Especially in the front region, judging by the overall organization of the vowel space and by the insignificant standard deviations between the speaking styles, we could conclude that the participants have vowels /iː/, /ɪ/, /e/ and /æ/ as separate, distinct, and relatively stable categories in their interlanguage vowel systems.

Yet, not all of our participants' vowel categories can be said to be distinct and well-defined, nor are they stable across different speaking styles and contexts. First, there are big differences in the quality of several vowels produced in the three different speaking styles. For instance, both inter—and intra- speaker variations for /ʌ/ produced by the female participants were so great that this vowel cannot even be said to occupy a compact area in the vowel space. Second, especially in the back and central regions of the vowel space, there is some significant overlapping of categories. For instance, with male speakers the areas of /ʌ, ɑː, ɒ/ and even /ɔː/ are clustered closely together and minimally dispersed in the vowel space. With female speakers these categories are more dispersed, but there is a significant overlapping between the categories /ʌ, ɑː, ɒ/. Third, although the categories in the front area of the vowel space do not overlap, they are minimally distinct for our female participants. For instance, although maintained *phonologically* in all the three contexts, the distinction between /iː/ and /ɪ/ is phonetically rather small, and these two vowel areas are grouped close together. In this sense, in our participants' vowel space several vowel categories are not clearly delimited, and do not comply with the principle of 'sufficient separation' (Butcher 1994:28), even where the formant values produced indicate at least a partial acquisition of vowel categories.

Finally, the last problem that can be observed in our data is the phonetic identity of the phonological categories. For instance, similar to the findings in some previous research (Marković 2009:9), our participants' vowels were characterized by notably steady formant values throughout the vowel. Even /iː/ and /uː/ were rarely – if ever – characterized by formant movement and the diphthongization typical of native speakers' pronunciation (cf. Neel 2008). Similarly, although Neel (2008:583) points out that /æ/ is characterized by "a considerable change in F2 over the course of the vowel", no trace of this tendency was observed in our participants' vowels.

Much more importantly, some of our participants' categories were phonetically very different from typical native-speaker categories. Specifically, although the category of /æ/ produced by both male and female participants was clearly delimited from the neighbouring areas and rather stable, the fact remains that its quality was that of a central and open vowel (the most open one), very different from the target category in English, which indicates that the participants have probably assimilated it into the L1 category of /a/. In this sense, our participants are like "early learners" in the research by Krebs-Lazendic (2008:154, Krebs-Lazendic & Best 2007) – they have "assimilated two members of the contrast to two different Serbian categories (/e/ and /a/, respectively)", unlike "late learners" who tend to assimilate both these English vowels into a single Serbian category, /e/. Therefore, this poses the question of whether we can consider this category to have been *acquired* by our participants, even though it is distinct, delimited and stable.

Similarly, when /ɜː/ and /ə/ are concerned, with both male and female speakers these two stand out in the vowel space as distinct and relatively stable categories, except in the spontaneous speech of the female participants, where /ɜː/ is different from the values produced in citation and reading. However, the position of these two vowels in the vowel space, especially of the unstressed schwa and especially with male speakers, is not typical of the acoustic 'character of schwa' as described in literature. Fabricius (2002:217) points out that surveys of average formant values for English vowels usually do not consider schwa, since the "acoustic profile of schwa has generally been reported as varying widely according to phonetic context". However, in our data the participants' values for unstressed schwa varied minimally, probably because even words in citation

form were used in a carrier sentence and the rest of the tokens were from connected speech/reading. Still, rather than being central, *schwa* occupies a rather close position in the vowel space, and for some reason is much more front with male speakers than with female speakers, who produced it as even more retracted than /ʊ/ in all the three speaking styles. Therefore, although from the point of view of their interlanguage vowel systems we could say that these categories are distinct and stable in our participants' speech, they are notably different from the target vowel qualities.

3.2. Vowel duration

It has been observed that EFL students tend to rely on vowel length rather than on vowel quality to distinguish between relevant vowel categories in L2, as pointed out by Chen (2006:248) or Cebrian (2006:372), whose participants showed "overreliance on duration" in making vowel distinctions. This tendency was not documented by the statistical analysis of our data, although the participants did produce consistent duration differences between short and long English vowels. Generally, /ɪ, ʌ, ɒ, ʊ/ were produced as shorter than the corresponding long vowels within any given speaking style. *Table 4* shows mean vowel duration values (in milliseconds) for vowel tokens and the standard deviation values for the three tasks – citation form, reading and speaking. It also shows the duration of vowels in the tokens that were tone-unit final in the speaking task.

Table 4. Mean values of vowel duration in the reading task, speaking task, and when the vowel occurred in the focus which was the final stressed syllable in the tone unit

	iː	ɪ	e	æ	ʌ	ɑː	ə	ɜː	ɒ	ɔː	ʊ	uː
citation	155	88	121	178	91	157	32	179	105	184	86	169
reading	121	73	99	120	86	127	25	128	99	134	72	139
speaking	152	83	105	234	85	153	21	156	110	138	81	158
TU end	167	129	151	240	150	206	/	156	147	153	91	185
Std.	18.82	7.64	11.37	115.45	3.21	16.29	5.57	25.54	5.51	27.78	7.09	15.18

Specifically, in the reading task the mean values for short vowels (except for /ə/ in unstressed syllables) ranged from 72ms to 120ms. If we disregard /æ/ and /e/, which seem to have an 'ambiguous' length identity, the short vowels were consistently produced as shorter than 100ms (72-99ms, mean=83ms), and long vowels as perceptibly longer (120-139ms, mean=130ms). In relevant vowel pairs, /iː/-/ɪ/, /ɑː/ -/ʌ/, /ɔː/-/ɒ/, /uː/-/ʊ/, the differences were 47ms on the average (means=45, 41, 35, and 67ms respectively). In the speaking task, the short vowels ranged from 81 to 110ms (mean=90ms), long vowels ranged from 138 to 158ms (mean=149ms). In the relevant vowel pairs, as above, the average difference was 56ms (means=69, 68, 28, and 77ms respectively).

Although all the vowels but /ɜː/ were actually longer in the speaking task than in reading but not citation, the standard deviation values across the three tasks were not found statistically significant. This indicates a generally slower tempo of delivery in the reading task, but what we find important here is the fact that the participants maintained a consistent use of duration to support vowel distinctions *within* each of the speaking styles, even in semi-spontaneous speech, and not only in task one that focused their attention on vowel differences.

With respect to /æ/ and /e/ the participants showed a curious tendency to produce /æ/ as a long vowel. In the reading task /æ/ was produced with a duration typical of long vowels (mean=120ms), and in the speaking task, this vowel was actually *the longest*

vowel produced, about 80ms longer than other long vowels; even in the citation task it was produced as longer than three other long vowels: /iː, ɑː, uː/. It seems that our participants tried to 'improve' the symmetry of the English vowel system in terms of length oppositions, imposing the length distinction on /æ/-/e/ by analogy with the other 'paired' vowels.

Finally, it could be observed that duration was used by our participants at the prosodic level, too, as a tone-unit boundary signal. Namely, in the speaking task, when the token occurred in the tone-unit final position, both long and short vowels tended to be prolonged, so that the duration values of short vowels that occurred in this position were much higher than those of the same vowel occurring in a stressed syllable but not tone-unit finally. Moreover, some short vowels in TU-final position were as long or even longer than the long vowels in a non-final position. Compare, for instance, the average TU-final value for /ɒ/=147ms with the non-TU-final average duration of /ɔː/=138ms, both measured in the same speaking task. This might indicate that in the participants phonological system the phonetic cue of duration is used at both the segmental and at the prosodic level, with the primacy at the prosodic level, when it is used as a discourse and illocutionary signal.

4. Conclusion

Abandoning the native-speaker model as a 'yardstick' (Jenkins 2006:175) in favour of intelligibility as a more realistic goal in L2 teaching has not erased the many problems EFL students are facing when struggling to acquire the English vowel system. Though we may not aim at sounding like native speakers any longer, there still remains the need for EFL students to re-structure their perceptual and articulatory vowel space to accommodate all the relevant vowel categories and all the relevant distinctions in their English interlanguage system. Even if not compared to the 'typical' native-speaker qualities, EFL students' vowel categories steel need to be clearly delimited, 'distinctive' in Neel's sense (2008:584), and stable in different contexts and different speaking styles, since this seems to be crucial for vowel intelligibility. Without explicit comparison with native-speaker data, this study showed clearly enough that our participants did not produce appropriate quality differences between vowel categories, at least not consistently, in all the three contexts, but that they did use duration a phonological signal to support vowel distinctions.

Therefore, the findings of research studies such as this one – notwithstanding the common limitations such as a small number of participants, and the inability to control the participant's use of vowel tokens in semi-spontaneous speech – can be a very useful starting point in addressing intelligibility problems in EFL students' speech, without having to resort to 'native-speaker' measures for comparison. Comparative studies of students' L1 and English vowels are undoubtedly important, because they show L1 interference in acquiring English vowel categories, but a close observation of the structure of the students' interlanguage vowel systems, and of the *relative* position and distinctness of specific vowel categories in it, is also invaluable, because it offers a dynamic picture of students' vowel system development, showing which vowel categories need to be further differentiated or phonetically modified, and which need to be stabilized in different contexts and speech styles.

References

Barboza, C.L.F. (2007). "Production of English front vowels by Brazilian EFL teachers in Western Rio Grande do Norte". In Rauber, A. S., Watkins, M.A. and B.O. Baptista (eds), *New Sounds 2007:*

Proceedings of the Fifth International Symposium on the Acquisition of Second Language Speech. Florianópolis, Brazil: Federal University of Santa Catarina, 28-37.

Best, C. T. (1994). "The emergence of native-language phonological influences on infants: A perceptual assimilation hypothesis". In J. C. Goodman and H. C. Nusbaum (eds), *The development of speech perception*. Cambridge: MIT Press, 167-224.

Best, C. T. (1995). "A direct-realist view of cross-language perception". In W. Strange (ed), *Speech Perception and Linguistic Experience: Issues in Cross-Language Research*. Balitmore, MD:York Press, 171– 204.

Bohn, O-S. and J.E. Flege (1990). "Interlingual identification and the role of foreign language experience in L2 vowel perception". *Applied Psycholinguistics* 11: 303–328.

Butcher, A.R. (1994). On the phonetics of small vowel systems: evidence from Australian languages. In R. Togneri (ed): *Proceedings of the 5th Australian International Conference on Speech Science and Technology.* Canberra: Australian Speech Science and Technology Association, Vol I, 28-33.

Cebrian, J. (2006). "Experience and the use of non-native duration in L2 vowel categorization". *Journal of Phonetics*, Vol. 34, Issue 3: 372-387.

Chen, Y. (2006). "Production of Tense-Lax Contrast by Mandarin Speakers of English". *Folia Phoniatrica et Logopedica,* vol. 58: 240–249.

Čubrović, B. (2007). "What's in a symbol? The case of happY and thank YOU vowels and the Serbian EFL learner". PTLC Phonetics Teaching & Learning Conference August 24-26, 2007,UCL, at <http://www.phon.ucl.ac.uk/ptlc/proceedings/ptlcpaper_05e.pdf>

Denes, P.B. & E.N. Pinson (1993). *The Speech Chain; the Physics and Biology of Spoken Language*, 2nd ed. New York: Freeman.

Deterding, David H. (1990). *Speaker normalization for automatic speech recognition.* PhD thesis, University of Cambridge.

Deterding, David H. (2006). "The North Wind versus a Wolf: short texts for the description and measurement of English pronunciation". *Journal of the International Phonetic Association*, 36/2: 187-196.

Djokić, D. (1981). "Uticaj fonološke strukture srpskohrvatskog jezika na učenje izgovora engleskih diftonga". *Godišnjak Saveza društava za primenjenu lingvistiku Jugoslavije*, br 4-5, Zagreb, 69-73.

Djokić, D. (1984). "Sposobnost auditivne diskriminacije i produkcije vokala na stranom jeziku". *Godišnjak SDPLJ, zbornik radova III kongresa SDPLJ*, 1983, Sarajevo, 687-693.

Fabricius, Anne (2002). "Weak vowels in modern RP: An acoustic study of happy-tensing and kit/schwa shift". *Language Variation and Change*, 14:, 211–237.

Fernandez Guerra, A. (2001). "Intelligibility vs. accentedness. Some issues concerning foreign language speech, learning and teaching". In F. Fernandez (ed), *Los estudio ingleses, en el umbral del tercer milenio*, Miscellaneous Vol 3, *Studies in English Language and Literature, Lingua Inglesa.* Universitat de Valencia, 143-156.

Flege, J. E. (1988). "Factors affecting the degree of perceived foreign accent in English sentences". *Journal of the Acoustical Society of America*, 84*:* 70–79.

Flege, J. E. (1995). "Second-language speech learning: Theory, findings, and problems". In W. Strange (ed), *Speech perception and linguistic experience: Theoretical and methodological issues*. Timonium, MD: York Press, 233–277.

Flege, J. E. (1997). "English vowel production by Dutch talkers: More evidence for the 'similar' vs 'new' distinction". In A. James & J. Leather (eds), *Second- Language Speech, Structure and Process*. New York: Mouton de Gruyter, 11-52.

Flege, J. E., Bohn, O-S. & S. Jang (1997). "The effect of experience on nonnative subjects' production and perception of English vowels*". Journal of Phonetics* 25: 437–470.

Gimson, A.C. (1970). *An Introduction to the pronunciation of English.* Bristol: Edward Arnold.

Guion, S. G. (2003). "The vowel systems of Quichua-Spanish bilinguals". *Phonetica*, 60: 98–128.

Hawkins, Sarah & Jonathan Midgley (2005). "Formant frequencies of RP monophthongs in four age groups of speakers". *Journal of the International Phonetic Association* 35/2: 183-199.

Hillenbrand, J. M., Clark, M. J., & T.M. Nearey (2000)."Some effects of duration on vowel recognition". *Journal of the Acoustical Society of America*, 108:3013–3022.

Hillenbrand, J. M., Clark, M. J., & T.M. Nearey (2001). "Effects of consonant environment on vowel formant patterns". *Journal of the Acoustical Society of America*, 109:748–763.

Iverson, P. & B.G. Evans (2007). "Learning English vowels with different first-language vowel systems: Perception of formant targets, formant movement, and duration*". JASA*, 122,5: 2842-2854.

Ivić, P. (1994). *Prozodija reči i rečenice*. Sremski Karlovc: Izdavačka knjižarnica Zorana Stojanovića.

Jenkins, Jennifer (2006). "Current Perspectives on Teaching World Englishes and English as a Lingua Franca". *TESOL Quarterly*,Vol. 40, No. 1:157-181.

Jia, G., & D. Aaronson (2003). "A longitudinal study of Chinese children and adolescents learning English in the US". *Applied Psycholinguistics*, 24: 131– 161.

Jia, G., W. Strange, Y. Wu, J. Collado & Q. Guan (2006). "Perception and production of English vowels by Mandarin speakers: Age-related differences vary with amount of L2 exposure". *Journal of the Acoustical Society of America*, 119, 2:1118–1130.

Kewley-Port, D. & S. S. Goodman (2005). "Thresholds for second formant transitions in front vowels". *Journal of the Acoustical Society of America*, 118, 5: 3252–3260.

Kewley-Port, D., Akahane-Yamada, R. & K. Aikawa (1996). "Intelligibility and acoustical correlates of Japanese accented English vowels". *Proceedings of International Conference on Spoken Language Processing,* Vol. 96, 450–453.

Krebs-Lazendic, L. (2008). *Early vs. late Serbian-English bilinguals' responses to two Australian English vowel contrasts.* Unpublished PhD thesis, University of Western Sydney, available at <http://arrow.uws.edu.au:8080/vital/access/manager/Repository/uws:5890>

Krebs-Lazendic, L. and C.T. Best (2007). "Early and Late Bilinguals' Vowel Perception and Production: English Vowel Contrasts that Give Serbian-English Bilinguals a H(E)AD-ache". In Rauber, A. S., Watkins, M.A. & B.O. Baptista (eds): *New Sounds 2007: Proceedings of the Fifth International Symposium on the Acquisition of Second Language Speech*, 282-292. Retrieved from <http://www.nupffale.ufsc.br/newsounds/Papers/24.Krebs-Lazendic_Best.pdf>

Ladefoged, P. (1993). *A Course in Phonetics*. 3rd ed. San Diego: Harcourt Brace College Publishers.

Ladefoged, P. (2001). *A Course in Phonetics*. 4th ed. Boston: Heinle & Heinle, Thomson Learning.

Lehiste, I. & P. Ivić (1986). *Word and Sentence prosody in Serbocroatian*. Cambridge,MA: MIT Press.

Lehiste, I. & P. Ivić (1996). *Prozodija reči i rečenice u srpskohrvatskom jeziku* (prev. Lj. Subotić). Sremski Karlovci, Novi Sad: Izdavačka knjiarnica Zorana Sojanovića.

Lindblom B., Brownlee S., Davis B. & S-J. Moon (1992). "Speech transforms". *Speech Communication,* 11: 357-368.

MacKay, I.R.A. & N. Fullana (2007). "Starting Age and Exposure Effects on EFL Learners' Sound Production in a Formal Learning Context". *New Sounds 2007: Proceedings of the Fifth International Symposium on the Acquisition of Second Language Speech*, pp 324- 335. Retrieved from < http://www.nupffale.ufsc.br/newsounds/Papers/28.Mackay_Fullana.pdf>

Major, R. C. (2002). "The phonology of the L2 user". In V. Cook (ed), *Portraits of the L2 user.* Clevedon, UK: Multilingual Matters, 67–92.

Markham, D. & V. Hazan (2002). UCL Speaker database. In *Speech, Hearing and Language: Work in progress,*Volume 14. Department of Phonetics and Linguistics, UCL, 1-17.

Marković, M. (2007). *Kontrastivna analiza akustičkih i artikulacionih karakteristika vokalskih sistema engleskog i srpskog jezika*. Unpublished PhD thesis, University of Novi Sad, Serbia.

Marković, M. (2009). "Different Strategies in Acquiring L2 Vowels: The Production of High English Vowels /iː, ɪ, uː, ʊ/ by Native Speakers of Serbian". In Čubrović, B. and T. Paunović (eds), *Ta(l)king English Phonetics Across Frontiers*. Cambridge: Cambridge Scholars Publishing, 3-18.

Mihajlović, LJ. (1957). "Problem učenja glasova stranog jezika". *Živi jezici*, god I, 1-2, Beograd: Društvo za strane jezike i književnosti NR Srbije, 57-64.

Mihajlović, LJ. (1969). "Kontrastivna analiza fonoloških sitema". U Filipović, R (ed), *Prilozi i gradja*, Zagreb: Institut za lingvistiku i YScecp Zagreb, 30-34.

Munro, M. J. & T.M. Derwing (2001)."Modeling perceptions of accentedness and comprehensibility of L2 speech". *Studies in Second Language Acquisition*, 23: 451–468.

Munro, M. J. & T.M. Derwing (2008). "Segmental Acquisition in Adult ESL Learners: A Longitudinal Study of Vowel Production". *Language Learning,* 58, 3: 479–502.

Munro, M. J., Derwing, T. M. & S.L. Morton (2006). "The mutual intelligibility of L2 speech". *Studies in Second Language Acquisition*, 28: 111–131.

Neel, A. T. (2008). "Vowel Space Characteristics and Vowel Identification Accuracy". *Journal of Speech, Language, and Hearing Research* , Vol. 51: 574–585.

Paunović, T. (2003). *Fonetsko-fonološka interferencija srpskog jezika u percepciji i produkciji engleskih vokala.* Unpublished PhD thesis, University of Nis, Serbia.

Peterson, G. E. & H.L. Barney (1952). "Control methods used in a vowel study". *Journal of the Acoustical Society of America*, 24: 175–184.

Piske, T., J.E. Flege, I.R.A. Mackay & D. Meador (2002). "The Production of English Vowels by Fluent Early and Late Italian-English Bilinguals". *Phonetica,*Vol. 59, No 1: 98-128.

Pisoni, D. B. (1997). "Some thoughts on 'normalization' in speech perception". In K. Johnson & J. W. Mullenix (eds), *Talker Variability in Speech Processing.* San Diego, CA: Academic Press, 9–32.

Savić, M. (2007). *Izgovor u nastavi stranog jezika – efikasnost različitih pristupa i tehnika na mladjem školskom uzrastu.* Unpublished MA thesis, University of Novi Sad, Serbia.

Savić, M. (2009). "Pronunciation instruction with young learners – Does it make a difference?" In Čubrović, B. and T. Paunović (eds), *Ta(l)king English Phonetics Across Frontiers*. Cambridge: Cambridge Scholars Publishing, 165-180.

Simić, R. & B. Ostojić (1989). *Osnovi fonologije srpskohrvatskoga književnog jezika*. Nikšić: Univerzitetska riječ.

Strange, W., A. Weber, E.S. Levy,V. Shafiro, M. Hisagi & Kanae Nishif (2007). "Acoustic variability within and across German, French, and American English vowels: Phonetic context effects". *Journal of the Acoustical Society of America*, 122, 2: 1111–1129.

Tsukada, K. (2008). "An acoustic comparison of English monophthongs and diphthongs produced by Australian and Thai speakers". *English world-wide*, Vol. 29, Issue 2: 194-211. Retrieved from <http://hdl.handle.net/1959.14/21317>

Wells, John (1962). *A study of the formants of the pure vowels of British English*. PhD Thesis. Formant data retrieved from <http://www.phon.ucl.ac.uk/home/wells/formants/table-1.htm>