

SPECTRAL AND TEMPORAL CHARACTERISTICS OF CZECH VOWELS IN SPONTANEOUS SPEECH

NIKOLA PAILLEREAU and KATEŘINA CHLÁDKOVÁ

ABSTRACT

This paper provides a comprehensive account of spectral and durational characteristics of Czech monophthongal vowels. It improves on the existing literature (that almost exclusively focused on read speech) in that it examines vowels in spontaneous speech recorded from 10 men and 10 women, who were recruited from the general population not restricted to students or media reporters (which were the populations used in previous studies). The present material thus represents a relatively naturalistic data set. The acoustical analyses of vowel spectral properties are not limited to only the first and the second formant (F1 and F2) but include also higher formants. Duration normalized for word length as well as long/short duration ratios are compared across all vowel qualities. In line with previous acoustic data on Czech high front vowels, the present results confirm that the phonologically short /ɪ/ is realized with a higher F1 than the phonologically long /i:/. The results further demonstrate that the mid front /ɛ/ and /ɛ:/ are realized with a relatively high F1 and are numerically even closer to the low /a/ and /a:/ than to the other mid vowel quality, the back /o/ and /o:/. A novel finding is that short back vowels /o/ and /u/ have a higher F2 than their long counterparts: this slight fronting is likely attributable to the spontaneous style of speech as well as to the mostly coronal context in which the vowels were embedded. In contrary to recent literature that reported extremely low long/short ratios in high vowels our findings show that duration marks the phonological length distinctions consistently across all five vowel pairs: long vowels are on average 1.76 times longer than short vowels. The study concludes with a discussion of the implications that the vowel acoustic properties may have on the way the Czech vocalic system is transcribed.

Key words: vowels, Czech, vowel formants, vowel duration, spontaneous speech, phonological transcription

1. Introduction

Each of the world's languages contrasts its vowels by their spectral quality, that is, by a set of frequency components called formants which are the resonant frequencies of the vocal tract (Fant 1960). Vowels are typically described in terms of the first and the second formant (F1 and F2), the former being roughly correlated to the vertical position of the

tongue and to jaw opening and the latter roughly to the horizontal position of the tongue and to lip settings (Crothers, 1978).

Although vowel descriptions commonly refer to a two-dimensional space with F1 plotted on the vertical and F2 on the horizontal axis, studies show that at least in some languages higher formants, especially the third and the fourth formant (F3 and F4), may serve as a main cue to vowel identity. F1 and F2 suffice to describe vowels whose dominant energies are located below 1000 Hz and whose higher formant frequencies are consequently weakened and become perceptually non-salient; these are typically back vowels such as /u/ or /o/ (Vaissière, 2011). However, when the energy is concentrated in higher frequencies, F3 and F4 can come to play a major role: this is especially the case of languages contrasting front rounded and unrounded vowels, such as French and Swedish, where F3 is roughly corelated to labiality (Fant, 1969; Vaissière, 2009). Higher formants alone might even differentiate vowel contrasts that had been traditionally understood as F1-based: in that respect, some native speakers of French do not distinguish their native /i/ and /e/ in terms of F1 and F2 but instead in terms of F3 and F4 (Kamiyama, 2011). Moreover, higher formants are pertinent in a cross-linguistic comparison of vowel spectra: while the acoustic target of French /i/ is to make F3 as high as possible such that it comes close to F4, thus making the F3/F4 zone perceptually most salient, the acoustic target of English /i/ is to make F2 and F3 come close together (Gendrot et al., 2008). In most languages, the realization of the phoneme /i/ indeed aims at maximal F2, but the “French” F3-F4 pattern is not uncommon and has been observed also in some speakers of English (Flemming, 2019). Since higher formants such as F3 and F4, and the distance between them, have been shown to cue vowel identity in at least some languages, it is desirable to include these higher formants in acoustic description of front vowels cross-linguistically.

1.1. The Czech vowel system

Czech vowel phonemes are distinguished by their spectral properties and by their duration. Czech has been described as contrasting 5 monophthongal vowel qualities, namely, [i]-like, [e]-like, [a]-like, [o]-like, and [u]-like, each of which occurs as short and long. To capture both the spectral and durational properties of the Czech vowels, the 10 monophthongal phonemes are by many recent authors transcribed as /i: ɪ ε: ε a: o: o u: u/ (Dankovičová, 1997; Podlipský et al., 2009; Chládková et al., 2009; Šimáčková et al., 2012; Paillereau, 2016; Skarnitzl et al., 2016; Chládková et al., 2019).

Although the monophthongal vowel inventory of Czech is symmetrical phonologically by differentiating the high front /i: ɪ/ from the high back /u: u/, and the mid front /ε: ε/ from the mid back /o: o/, phonetically the mid front vowels are consistently realized with much higher F1 values than the mid back vowels (Skarnitzl & Volín, 2012; Šimáčková et al., 2012; Paillereau, 2016; Chládková et al., 2019). Besides the phonetic ‘lowness’ of the front mid vowel, what is perhaps the most intriguing feature of the Czech vowel system is the realization of vowel quantity contrasts.

The short-long phoneme distinction within each of the five phonological vowel qualities has been typically realized primarily by duration (Chlumský, 1928). Yet, the phonological length contrast within the high front vowel pair is consistently realized through

spectral properties, with the short member having a higher F1 (and a lower F2) than the long one, as captured in the commonly employed transcription /i:/ versus /ɪ/. The spectral distinction in the high front short-long vowel pair was observed already by the early Czech phoneticians (e.g. Frinta, 1909; Hála, 1962) who, however, did not consider it significant enough to be captured in the transcription (Frinta, 1925). The spectral differentiation of /i:/-/ɪ/ has been objectively confirmed by a number of recent acoustic measurements (Skarnitzl & Volín, 2012; Šimáčková et al., 2012; Paillereau, 2016; Chládková et al., 2019). Spectral differentiation of a phonological length contrast, comparable to that attested in /i:/-/ɪ/, has not (yet) been found for the high back vowels, although some note a potential trend in that respect (either explicitly as Skarnitzl & Volín, 2012, or implicitly by transcribing the vowels as /u:/ /ʊ/ in Duběda, 2005). Czechs not only realize the phonological length contrast between /i:/ and /ɪ/ through spectral differences when speaking but they also rely on spectral cues when listening. Two recent speech perception experiments report a strongly spectrally-guided perceptual differentiation of the long-short /i:/ /ɪ/ contrast and, at the same time, show that the extent to which spectrum cues the long-short contrast in the high back /u:/ /ʊ/ is smaller (Podlipský et al., in press; Paillereau & Skarnitzl, 2019).

About a century ago, (stressed) phonologically long vowels were measured as being twice as long as the (stressed) short ones (Chlumský, 1928). Only ten years ago, then, an analysis of vowels produced by 6 speakers reported strikingly smaller durational ratios, especially for the high front and high back vowel pairs: the long phoneme being only 1.3 times longer than the short one for the high front vowels (originally reported in Podlipský et al., 2009, subsequently referred to in Skarnitzl, 2012; Skarnitzl & Volín, 2012; Skarnitzl et al., 2016). The comparison of the early 1928 and the later 2009 measurement might seem to indicate a diachronic trend whereby the declining durational difference come to be supplemented, or perhaps even overtaken, by a more pronounced spectral difference in order to maintain the contrast (see also a similar proposal by Šimáčková et al., 2012). This proposal remains a speculation, partially due to the limited number of speakers in the 2009-sample and the difference in speech style between Chlumský's study of spontaneous speech and the Podlipský's et al. study of read speech.

1.2. Aims of the present study

The aim of our study is to provide a thorough acoustic analysis of Czech monophthongal vowels from spontaneous speech. Spontaneous production may better represent natural speech realization than recordings of read material, the latter being the focus of most recent studies. Our population are non-students, which is another improvement on previous studies that recruited students or professional media presenters (both of which are rather specific populations unlikely representative of the average speaker of Czech).

Vowels are analysed here in terms of vowel formants and duration. Our objectives are as follows. Firstly, we assess and compare the spectral F1 and F2 properties of all 10 monophthongs to show whether and to what extent short-long contrasts are differentiated by spectrum (being specifically interested in the spectral distinction within high front and high back vowels), and whether the F1 of front mid vowels is more close to

that of the back mid vowels or to that of the low vowel /a/. Secondly, we aim to find out whether, in spontaneous speech, durational ratios of long to short vowels are comparable to those reported for read speech in Podlipský et al. (2009). The ratios in spontaneous speech could be smaller, which would indicate that the importance of duration in Czech speech is indeed declining (in line with what the divergent results between old and new studies suggest). On the contrary, the long/short ratios could as well be larger than previously reported which would indicate that in spontaneous speech (in which vowel spectral qualities are in general reduced as compared to read, careful speech) duration reliably cues vowel distinctions. Thirdly, we analyse and report the F3 and F4 and test whether the psychoacoustic distances between the higher formants help differentiate amongst the four front vowels (which is what has been found in e.g. French). Finally, in relation to the vowels' acoustic characteristics, we discuss the IPA symbols that had been and could be used in the phonemic transcription of Czech vowels.

2. Method

2.1. Speakers

Ten male and ten female speakers who have been living in the Prague region for at least 5 years and who did not have any noticeable regional accent were recruited for the purpose of the study. Male speakers were aged between 27 and 48 years (mean = 34.6, s.d. = 5) and female speakers between 25 and 34 years (mean = 29.6, s.d. = 2.1). They were healthy individuals with no hearing or speech impairments and were paid for their participation.

2.2. Recording procedure

Speakers were instructed to spontaneously comment on 20 objects that were given at their disposal. The 20 objects had been carefully chosen so that their names would contain all Czech monophthongal vowels /ɪ i: ε ε: a a: o o: u u:/ in a word-initial, i.e. stressed, syllable. The vowels were embedded in a controlled consonantal context (as far as this was possible with object names): preceding consonants were mainly bilabials and following consonants mainly alveolars. The speakers were instructed to mention the name of each object at least twice when talking about it. To ensure that the objects would be named consistently across participants, and in a non-diminutive form (which would alter the number of syllables in a word), all the objects had a sticker with their name written on it. The production task was mainly a monologue but when speakers were running out of ideas, the experimenter engaged in a conversation about the objects. The 20 words from which vowels were segmented and analysed are listed in Table 1. Recordings were made in a sound-treated booth using a head-mounted condenser microphone AKG C520 and an Edirol UA 25 sound card connected to a PC running the Audacity software (version 2.3.0. retrieved from <http://audacity.sourceforge.net>). The material was digitized at a 44.1-kHz sampling frequency and 16-bit quantization.

Table 1. Words (orthographic, phonemic, translation) containing the target vowels

vowel	words orthographic	words phonemic	English translation
ɪ	miska, pytel	/miska/, /pitɛl/	bowl, sack
i:	víla, sítko	/vi:la/, /si:tko/	fairy, sieve
ɛ	meloun, metla	/mɛloun/, /mɛtla/	melon, whisk
ɛ:	léto, pérko	/lɛ:to/, /pɛ:rko/	summer, (bail) spring
a	balón, maska	/balo:n/, /maska/	ball, mask
a:	šátek, páska	/ʃa:tɛk/, /pa:ska/	scarf, tape
o	bota, kostka	/bota/, /kostka/	shoe, cube
o:	tóny, glóbus	/to:nɪ/, /glo:bus/	tones, globe
u	dudlík, husa	/dudli:k/, /husa/	pacifier, goose
u:	hůlka, kůra	/hu:lka/, /ku:ra/	wand, crust

2.3. Acoustical analyses

Word and vowel onsets and offsets were marked and labelled using Praat (Boersma and Weenink, 2018). A vowel token was included in the analysis if the target word form did not change in the number of syllables (suffix alternations not resulting in syllable-count change were accepted), and if the word was not mispronounced. Word onsets and offsets were marked as the onsets and offsets of the first and last segment, respectively, aligned to zero crossings of the waveform. Vowel onsets and offsets were marked on the basis of both the spectrogram and the waveform: the vowel interval had to contain visible energy in a broad-band spectrogram and visible formants (especially F2), and its first and the last waveform-period had to have a similar shape as the token's medial periods.

Vowel formants were measured by the optimized ceiling method (Escudero et al., 2009; Chládková et al., 2011) which searched for such a formant ceiling that yielded minimal variation in the measured F1, F2, and F3 values, per vowel category and per speaker. With the optimal ceiling settings, values of the first four formants were measured over the entire vowel portion with a Gaussian-like window centered at vowel midpoint, using the Burg algorithm implemented in Praat (Boersma and Weenink, 2018). Tokens for which the analysis yielded unlikely values (e.g., /a/-tokens measured with /u/-like low F1 and low F2 values) were reanalysed manually. The final set contained 1386 vowel tokens (133 occurrences of /ɪ/, 153 of /i:/, 130 of /ɛ/, 143 of /ɛ:/, 136 of /a/, 149 of /a:/, 152 of /o/, 119 of /o:/, 135 of /u/, 136 of /u:/), of which 692 were uttered by women and 694 by men.

2.4. Statistical analyses

Formant values measured in Hz were transformed to ERB using the Praat *hertzToErb()* function that implements the formula:

$$y = 11.17 \ln \left(\frac{x + 312}{x + 14680} \right) + 43$$

where x is the formant value in Hertz.

Vowel duration measured in ms was normalized for total word duration using the formula:

$$y = a \frac{x_V}{x_W}$$

where x_V is a token's vowel duration in seconds, x_W is the same token's word duration in seconds, and $a = 0.5$ which is the rounded word duration average across all 1386 words in the data set.

Statistical analyses were performed in R (R Core team, 2008), using packages *lmerTest* (Kuznetsova et al., 2017) and *emmeans* (Lenth et al., 2018). The ERB-transformed F1 and F2 and the normalized duration were each submitted to a linear mixed-effects model with vowel length, vowel quality, and sex as fixed factors with orthogonal contrasts that were specified uniquely in each of the three models as follows. For F1 we tested *i* vs. *u*, *i* vs. *o*, *e* vs. *a*, and *e* vs. *o*; for F2 we tested *e* vs. *i*, *o* vs. *u*, *a* vs. *e*, and *a* vs. *o*; for duration we compared each of *i*, *a*, *o*, and *u* to *e* as the reference category (note that here and in the following sections, we use vowel orthographic symbols in italics to denote one of the five phonological vowel qualities collapsing across the short-long phonemes of that vowel quality). Speaker was entered as a random factor with per-vowel quality and per-vowel length random slopes.

Another two mixed-effects models were run to test the higher-formant characteristics of the four front vowels: one for the F3-F2 difference and one for the F4-F3 difference (in ERB). Vowel and sex were fixed factors (with the following orthogonal contrasts for vowel /i:/ vs. /ɪ/, /i:/ vs. /ɛ:/, and /ɪ/ vs. /ɛ/), including speaker as a random factor. A last model was run to test long/short duration ratios across the five vowel qualities. Long/short ratios were computed separately for each vowel quality per speaker from the normalized duration values. Sex and vowel quality were fixed factors, testing the following 4 orthogonal contrasts: *a* vs *iu*, *a* vs *eo*, *i* vs *u*, and *e* vs *o*; speaker was entered as a random factor.

3. Results

Figure 1 shows the 10 Czech monophthongs in an ERB-scaled F1-F2 space separately for women and men. Figure 2 visualizes the vowels' spectral characteristics from F1 through F4, pooled across sexes. Table 2 then lists F1 and F2 values in Hertz, and Table 3 gives the front vowels' F3 and F4 values, and their psychoacoustic distances from F2 and F3, respectively. Table 4 shows raw and normalized vowel durations and the long/short ratios.

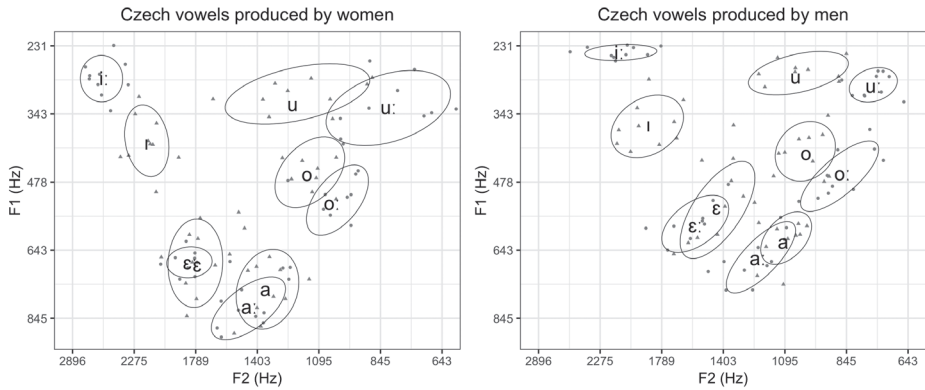


Figure 1. Czech vowels plotted in the F1-F2 plane, symbols show averages over 10 speakers per sex, ellipses cover one standard deviation from the mean, i.e. 68 % of the data. Grey symbols depict per-speaker median values (on which the group-averaging was applied); circles = long vowels, triangles = short vowels. The axes are scaled in ERB (with a 2-ERB distance between neighbouring axis labels), values are shown in Hz.

Table 2. First- and second-formant values (in Hz) and 95% confidence intervals (estimated with the *emmeans* R package) of the 10 Czech monophthongs, for 10 women and 10 men.

vowel	women				men			
	F1		F2		F1		F2	
	mean	95% c.i.	mean	95% c.i.	mean	95% c.i.	mean	95% c.i.
i:	287	268–308	2504	2382–2632	235	217–253	2157	2052–2267
ɪ	411	386–438	2177	2064–2296	347	324–371	1876	1778–1978
e:	671	636–707	1825	1731–1924	583	552–616	1571	1490–1657
ɛ	650	613–688	1726	1630–1827	564	531–598	1485	1402–1573
a:	784	745–825	1436	1362–1513	685	650–722	1232	1168–1300
a	733	694–773	1322	1249–1399	639	604–675	1133	1069–1200
o:	529	497–563	1024	966–1085	455	426–485	872	821–925
o	474	442–507	1161	1095–1230	404	376–434	992	934–1052
u:	341	319–364	851	787–919	283	263–304	720	663–780
u	330	307–355	1221	1134–1313	274	252–296	1044	969–1125

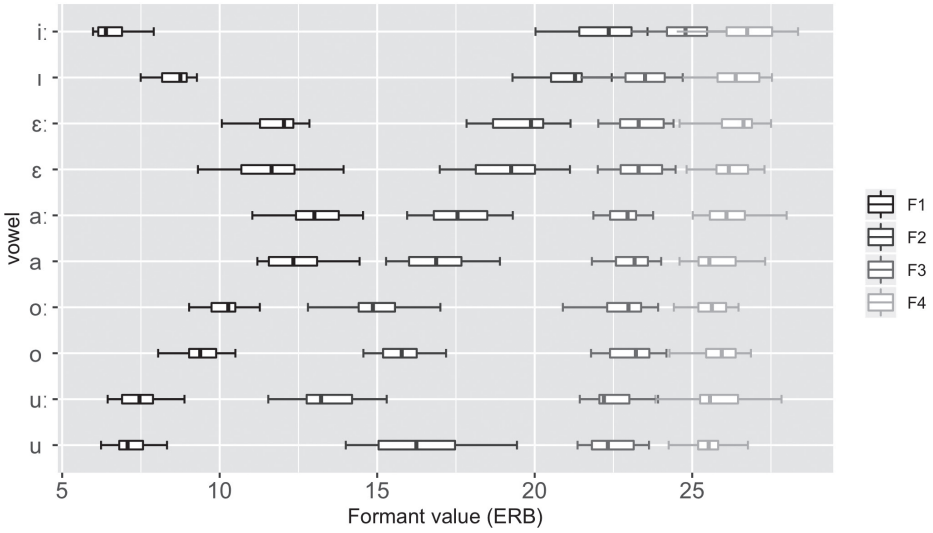


Figure 2. The F1–F4 values of the 10 Czech monophthongs, pooled across sexes. Boxes range from the 25% to the 75% percentile, vertical lines mark the 50% percentile, whiskers represent 1.5 times the interquartile range shown by the boxes.

Table 3. Third- and fourth-formant values (in Hz and ERB) of the 10 Czech monophthongs, the difference F3-F2 and F4-F3 (in ERB), averaged across sexes.

vowel	F3				F3-F2 (ERB)	
	Hz		ERB		mean	95% c.i.
	mean	95% c.i.	mean	95% c.i.		
i:	3202	3105–3301	24.82	24.57–25.08	2.65	2.37–2.92
ɪ	2688	2608–2770	23.38	23.13–23.63	2.37	2.10–2.65
ɛ:	2668	2588–2750	23.32	23.07–23.57	3.78	3.50–4.05
ɛ	2675	2595–2757	23.34	23.09–23.59	4.26	3.99–4.54
	F4				F4-F3 (ERB)	
	Hz		ERB		mean	95% c.i.
	mean	95% c.i.	mean	95% c.i.		
i:	4058	3940–4181	26.74	26.50–26.97	1.91	1.69–2.14
ɪ	3887	3775–4004	26.40	26.16–26.63	3.01	2.79–3.23
ɛ:	3893	3780–4009	26.41	26.17–26.64	3.08	2.86–3.31
ɛ	3785	3676–3898	26.18	25.95–26.42	2.84	2.62–3.06

3.1. Vowels formants

For F1, we found a main effect of sex confirming the anatomically conditioned sex difference in vowels having a larger F1 in women than in men (by on average 1 ERB, $t[22.6] = 6.277$, $p = 2 \times 10^{-6}$). As for the vowel quality contrasts, all of our 4 comparisons of vowel qualities turned out significant implying that *i* has an overall larger F1 than *u* and smaller F1 than *o* (by 4.9 and 9.4 ERB, respectively), and that *e* has a larger F1 than *o* and a smaller F1 than *a* (by 9.4 and 5.7 ERB, respectively; all $ps < .001$). Importantly, significant interactions with vowel length showed that the vowel quality comparisons are differentially modulated by vowel length: the *i* vs. *o* difference is 4 times larger for the long vowels than for the short vowels, being 3.6 and 0.9 ERB respectively, and the *i* vs. *u* difference is in different directions for short than for long vowels, short /*i*/ having larger F1 than short /*u*/ by 1.3 ERB and long /*i*/ having smaller F1 than long /*u*/ by 0.9 ERB. As for short-long comparisons within each vowel quality, the estimated means and confidence intervals (see also Figure 2) show that for three vowel qualities the short and long members differ significantly in their F1: /*i*/ has a larger F1 than /*i*:/ by 2 ERB, /*a*/ has a smaller F1 than /*a*:/ by 0.5 ERB, and /*o*/ has a smaller F1 than /*o*:/ by 0.7 ERB.

For F2, we again found a main effect of sex showing that vowels have overall larger F2 in women than in men (by on average 1.2 ERB, $t[21.2] = 6.443$, $p = 2 \times 10^{-6}$). Also, short vowels were found to have an overall larger F2 than long vowels (by on average 0.3 ERB, $t[58.6] = 3.947$, $p = 2 \times 10^{-4}$). All main effects of vowel quality as well as all interactions of vowel quality and vowel length came out as significant, we thus directly turn to the pairwise comparisons of estimated means. Comparisons of vowel qualities detected a significant between-vowel difference for all pairs except for short /*u*/ versus short /*o*/ (and short /*u*/ versus short /*a*/, which however was not a planned comparison in our design). The comparison of F2 between short and long members within each vowel quality reveal that for *i*, *e*, and *a* the long member has a higher F2 than the short member (by 1.2, 0.5, and 0.7 ERB), while for *u* and *o* it is the short member that has a higher F2 than the long one (by 2.8 and 1 ERB, respectively).

As for the higher formants, the analyses showed that the F3-F2 distance is larger in /*i*:/ than in /*ε*:/ by 1.02 ERB, and larger in /*i*/ than in /*ε*/ by 2 ERB ($t[60] = -5.8$, and -11.4 , respectively, both $ps < .001$). The F4-F3 distance is smaller in /*i*:/ than in /*i*/ by 0.85 ERB, and smaller in /*i*:/ than in /*ε*:/ by 0.74 ERB ($t[60] = 5.5$, and -5.5 , respectively, both $ps < .001$).

3.2. Vowel duration

For duration, the intercept was estimated as 0.096 norm s ($t[19.9] = 50.709$, $p < 2 \times 10^{-16}$), meaning that the average duration of vowels in our data set was 0.096 normalized seconds (that is, the mean vowel duration was 96 milliseconds in an average 500-ms-long word). There was a main effect of vowel length confirming that long vowels have overall larger duration than short vowels (by on average 0.051 norm s, $t[20.1] = -16.296$, $p = 5 \times 10^{-13}$). Furthermore, vowels produced by men were slightly longer than vowels produced by women (by on average 0.006 norm s, $t[28.4] = -2.342$, $p = .026$).

The main effect of vowel quality was significant for the *e-i* and for the *e-a* comparison suggesting that *e* is longer than *i* and shorter than *a* (by on average 0.030 and 0.008 norm s, respectively, both *ps* < 0.05). As vowel quality interacted with vowel length for three out of the four vowel contrasts (*e-i*, *e-o* and marginally for *e-u*), we turn to the inspection of the estimated means to unpack the interactions (involving both the planned and unplanned comparisons). Correcting alpha for all of the 20 individual comparisons, the data reveal that amongst long vowels, /a:/, /ɛ:/, /o:/, and /u:/ are significantly longer than /i:/ by about 0.030 norm s (a similar but nonsignificant trend is seen in /ɛ:/ and /o:/ tending to be longer than /u:/, by about 0.012 norm s). Amongst the short vowels, /a/, /ɛ/, and /u/ are trending towards being longer than /i/ by about 0.014 norm s, reaching significance only for the /ɛ/-/i/ comparison. As for the long-short comparisons within vowel qualities, all turned out significant implying that duration distinguishes a short and a long member in all 5 vowel pairs.

The model for duration ratios yields an intercept of 1.76, implying that long vowels are on average 1.76 times longer than short vowels ($t[80] = 20, p < 2 \times 10^{-16}$). The analysis further reveals that the long/short ratio in high vowels (*i* and *u*) is smaller than the ratio in the low vowel *a* (by 0.32, $t[80] = 2.978, p = .0038$) which in turn is smaller than the ratio in the mid vowels (*e* and *o*; by 0.35, $t[80] = -3.341, p = .0013$). The long/short ratio being the largest in mid vowels seems to be driven mainly by the large long/short ratio in *o* which significantly outweighs the long/short ratio in the other mid vowel quality *e* (by 0.19, $t[80] = 2.064, p = .042$); see also Table 4.

Table 4. Raw and word-length normalized duration of the 10 Czech monophthongs, and long/short duration ratios, averaged across sexes.

vowel	Raw duration (s) mean and 95% c.i.		Normalized duration (norm s) mean and 95% c.i.		Long/Short ratio
i	0.052	0.046–0.058	0.061	0.054–0.067	1.66
i:	0.090	0.082–0.097	0.098	0.090–0.106	(1.49–1.83)
ɛ	0.069	0.062–0.075	0.076	0.071–0.081	1.78
ɛ:	0.125	0.114–0.136	0.133	0.126–0.140	(1.61–1.95)
a	0.072	0.066–0.077	0.075	0.070–0.080	1.73
a:	0.138	0.128–0.147	0.126	0.119–0.133	(1.56–1.90)
o	0.059	0.053–0.065	0.067	0.060–0.075	1.97
o:	0.137	0.126–0.147	0.132	0.122–0.143	(1.80–2.14)
u	0.069	0.063–0.075	0.074	0.069–0.079	1.65
u:	0.104	0.093–0.114	0.119	0.111–0.128	(1.48–1.82)

4. Discussion

In this study we recorded the spontaneous speech of 20 speakers representative of the general Czech-speaking population (who use the standard variety of Czech spoken in the central Bohemian area) and analysed the vowels occurring in the initial, stressed, syllable of disyllabic content words (nouns). We performed acoustical and statistical analyses of the vowels' spectral properties, namely, F1 and F2 in all 10 monophthongs, and F3 and F4 in the four front vowels, and on duration, namely, vowel duration normalized for word duration, and long/short duration ratios.

4.1. Acoustic characteristics of Czech monophthongs

The results showed that the high front vowel pair is reliably distinguished by F1: the long /i:/ has a smaller F1 than the short /ɪ/, by 2 ERB, a difference which by far exceeds the just noticeable difference for formants (which is 0.2 ERB for [ɪ]-like vowels, Kewley-Port, 1995). The significant lowering of the short /ɪ/ in the vowel space is further documented by this vowel being, in terms of F1, four times closer to the short mid back /o/ than the long /i:/ is to the long mid /o:/. This F1 distinction between /ɪ/ and /i:/ is in line with previous acoustic measurements of vowels from read speech (Skarnitzl & Volín, 2012; Šimáčková et al., 2012; Paillereau, 2016) and matches the impressionistic observations of spontaneous speech from the 20th century (Frinta, 1909, 1924; Beneš, 1943; Chlumský, 1928; Hála, 1955; note that Hála, 1941, 1962 noticed an openness not only of the short but also of the long front high vowel).

The data further showed an asymmetry across the mid vowels. The front /ɛ/ and /ɛ:/ are realized with higher F1 than the back /o/ and /o:/. This disentanglement between front and back vowels is further strengthened by the front (phonologically) mid vowels being more similar in F1 to the *low* /a/ and /a:/ than to the other mid vowel quality, the back /o/ and /o:/. The front-back asymmetry could be explained in terms of Lindblom's Adaptive Dispersion Theory (Liljencrants & Lindblom, 1972) which argues that the (changes in) individual vowel qualities are determined by the entire system of vocalic contrasts. Thus, in order to maximize the perceptual contrast between short /ɪ/ (which is realized with much higher F1 than the long /i:/) and the front /ɛ/ and /ɛ:/, the F1 of the front mid vowels aims at high(er) F1 values. In the back part of the vowel system, no evidence is found for a lowering of the short high /u/ and there is thus no reason for the mid /o/ to be pushed towards higher F1 values.

In terms of F2, the long vowels had more peripheral values than their short counterparts. Interestingly, however, this effect for the back vowels was more than twice as large as that for the front vowels indicating a significant fronting of the short /o/ and /u/. The apparent fronting of the short back vowels possibly had two interrelated causes. Firstly, most of the post-vocalic consonants were coronals that notoriously cause rising of back vowels' F2 (Stevens & House, 1963), and due to the short vowels' inherent shortness the coarticulatory effects of flanking consonants affect a larger proportion of the vowel than is the case for inherently long vowels. Secondly, due to a generally less careful articulation in spontaneous (as compared to read) speech, the back vowels for which speakers aimed at only short duration underwent target undershoot not reaching the peripheral, low, F2

values representative of phonological backness. To what extent it was the consonantal context or the spontaneous speech style that lead to the fronting of the short back vowels remains a question open for future research.

Curiously, our data revealed that long low vowel /a:/ has a slightly higher F2 than the short low /a/. Although the perceptual reality of the 0.7-ERB difference is questionable, a fronting of the long /a:/ has been mentioned previously by Skarnitzl & Volín (2012) and reported by Paillereau (2016) for speakers of the regional Pilsner dialect of Czech.

Results on higher formants showed that F3 and F4 are converging in the long /i:/ more so that they do in the short /i/ (and in the short /ɛ/). This finding is interesting from a cross-linguistic perspective: the F4-F3 difference that we found for Czech /i:/ resembles that of the French (prepalatal) /i/ that had been thought to exhibit a cross-linguistically unique pattern of F3-F4 focalization (Gendrot et al. 2008, Vaissière 2011).

Table 5. Third- and fourth-formant values (ERB) of /i/ in 8 languages (Table 1 from Gendrot et al. 2008) and of the Czech /i:/ and /i/ (the present data, in bold), as well as the difference F4-F3 (in ERB), averaged across sexes.

	F3	F4	F4-F3
French	24.15	25.92	1.77
Czech /i:/	24.82	26.74	1.91
Arabic	23.62	25.59	1.97
Mandarin	24.12	26.29	2.17
Spanish	23.96	26.27	2.31
Italian	23.67	26.16	2.49
English	23.05	25.79	2.74
German	23.40	26.23	2.83
Czech /i/	23.38	26.40	3.01
Portuguese	23.05	26.13	3.08

Table 5 gives an overview of F3 and F4 values, and their psychoacoustic distance (in ERB), that had been previously reported for 8 languages by Gendrot et al. (2008) along with the currently measured values for Czech /i:/ and /i/. It is seen that while the focalization is numerically smallest in the French /i/, Czech /i:/ appears to be more focalized than the /i/ in the 7 remaining languages (and at the same time seems to have the highest F3 and F4 values of the entire sample). Investigation of higher formants may be beneficial not only from cross-linguistic perspective but also cross-dialectally. The F1-F2 difference between /i:/ and /i/ that we report here holds for Bohemian varieties of Czech and its extent is reportedly smaller in Moravian varieties (Šimáčková et al., 2012): future studies could investigate whether there are (also) any dialectal differences in the extent to which higher formants cue the distinction between the short and the long high front vowel.

We found that duration reliably distinguishes between the short and the long phoneme across all five vowel qualities. Amongst long vowels /i:/ was the shortest and since a similar trend was seen also in the short vowel set, the apparent shortness of /i:/ did

not lead to an exclusively smallest long/short ratio for the /i:/-/ɪ/ vowel pair. Long/short ratios of the high front and high back vowels were the smallest, followed by an intermediate long/short ratio for the low vowel quality and the largest ratio for the mid vowels. Crucially, however, the /i:/-/ɪ/ ratio measured here, i.e. 1.66, was much larger than the /i:/-/ɪ/ ratio of 1.29 reported by Podlipský et al. (2009) (and by comparing our lower confidence bound 1.49 to the mean of Podlipský et al., this difference was most likely significant). The methodological differences between ours and Podlipský's et al. study lying in the speech style (spontaneous vs read, respectively), population (general public vs news reporters, respectively), and in the number of participants (20 vs 6, respectively) suggest that the data from the current study may reflect Czech vowel durations more veridically than the data reported in the 2009 study.

Apart from the disparate finding for /i:/-/ɪ/, the long/short ratios for the remaining 4 vowel pairs resemble the ratios reported for these vowel pairs by Podlipský et al. The average long/short ratio in our spontaneous speech material was 1.76 which is smaller than the long/short ratio of 2 reported by Chlumský (1928), and except /o:/ none of the long vowels comes close to potentially being twice as long as the short one (with the highest upper confidence bounds of 1.9, 1.95, and 2.14 for /a:/-/a/, /ɛ:/-/ɛ/, and /o:/-/o/, respectively). It thus appears that duration ratios between long and short vowels may have become reduced over the past century. However, further research is needed that would assess and directly compare vowel durations across speech styles to resolve the conflict between ours and Podlipský et al. (2009) study with respect to the /i:/-/ɪ/ ratio.

As a final note on duration, we found that the long/short ratio was the largest for /o:/ vs. /o/, an effect which most likely stems from the fact that the long /o:/ is not a genuine Czech phoneme; it has come to the language with recent borrowings, and occurs only in a small set of relatively infrequently used words (Ludvíková & Kraus, 1966; Podlipský et al., 2009; Šimáčková et al., 2012). Because there is a link between item frequency and prototypicality of articulation (e.g. Aylett and Turk, 2006), the infrequent long vowel /o:/ may be realized as a hyperarticulated, unnaturally produced speech segment.

4.2. On the phonological notation

As noted in the Introduction, across authors and across studies there seems to be an inconsistency in how Czech vowels are transcribed phonemically. One, and nowadays probably the most frequently used, approach to transcribing Czech vowels is phonetically motivated and thus depicts both the length and the quality distinction in the high front vowels by transcribing them as /i:/ and /ɪ/ and also depicts the significant lowering of the mid front vowels – in contrast to the mid back vowels – by transcribing them as /ɛ(:)/ and /o(:)/, respectively. The phonetically motivated transcription has been used across acoustic vowel studies (including the present one) as well as in phonological descriptions of Czech (Dankovičová, 1997; Podlipský et al., 2009; Chládková et al., 2009; Šimáčková et al., 2012; Paillereau, 2016; Skarnitzl et al., 2016; Chládková et al., 2019).

The other approach to transcribing Czech vowels seems to be formally motivated such that it aims to capture the phonological symmetry of the system omitting some of the (relevant) phonetic information, which results in /i: i e: e a: o: o u: u/ and has been used

by Bičan (2013) and Palková (1997), both of whom make explicit notes on the phonetic deviations violating the symmetry. Yet other recent authors' symbol use seems to be motivated both phonetically and phonologically resulting in a somewhat inconsistent description. For instance, transcribing the monophthongal phonemes as /i: ɪ ε: ε a: a ɔ: ɔ u: ʊ/, Duběda (2005) captures the actual phonetic realization of the front vowels but, at the same time, attempts to instantiate a front-back symmetry by using distinct symbols for the short versus the back high back vowel, and by transcribing the mid back vowel as an open /ɔ(:)/. The rather ambiguous choice to realign the back vowels to conform to the phonetically-grounded realizations of the front vowels has not been, to the best of our knowledge, supported by any acoustic or perceptual studies (although early Czech phoneticians did note a lowering of /o/ in the contemporary speech, see below).

Most of the earlier authors were, too, aware of the vowels' unique phonetic realizations but purposefully referred to the system as symmetrical with their goal being to *prescribe* how Czech speakers should realize vowels wishing to prevent the actually observed, disfavored open realizations (mostly pertaining to lowering of the front mid vowel; e.g. Hála, 1941, 1962 and Beneš, 1943, but see also Borovičková & Maláč, 1967 who describe the realizations of /i/ and /i:/ as spectrally similar). Frinta (1909, 1924) was one of the few early authors using phonetically motivated symbols aiming to *describe* the Czech phonemes as they are realized by an average speaker of Czech (and not to prescribe how the vowels should be pronounced). On the basis of impressionistic observations, Frinta (1909, 1924) used /ɛ/ and /ɔ/ to capture the lowering of the mid vowels and used /i/ and /i:/ for the high front vowels noting a spectral difference between them but not considering it large enough to be captured in the transcription.

In the present study that is aimed as a description of the spectral and durational characteristics of Czech vowels, we employed the transcription /i: ɪ ε: ε a: a ɔ: o u: u/ capturing the significant spectral distinction within the high front vowel pair and the lowering of the mid front vowels. The present data do not support the use of /ʊ/ for the short high back vowel as we did not detect an F1 difference between the short and the long high back vowels (not detecting any F1 difference between /u/ and /u:/ of course does not mean that the difference may not exist but it does justify not introducing the use of two different symbols for those two vowels). We also keep transcribing the mid back vowels as /o(:)/ to depict the significant asymmetry in the F1 of front versus back mid vowels.

The variations in phonemic symbol use are apparent not only between authors but also between studies by the same authors who transcribe the Czech mid front vowel as /e/ in some cases (Skarnitzl & Volín, 2012) but as /ɛ/ in others (Podlipský, Skarnitzl & Volín, 2009; Skarnitzl, Šturm & Volín, 2016). Firstly, as Wells (2001) pointed out, the choice of IPA symbols can be adapted according to the audience that one and the same author may aim at with different studies. The above described inconsistency does not seem, however, to be due to different audiences that the authors aim at – all of them reporting on acoustic (and perceptual) properties of vowels. It rather demonstrates a general difficulty to transcribe mid vowels in a language that has only 3 degrees of vowel height with an IPA chart that was designed on the basis of French, English and German vowel inventories (Grammont, 1933), all of which contrast 4 degrees of height, and thus contrast also /e/ and /ɛ/. The mid front vowel of languages with 3 degrees of vowel height is then mostly transcribed as /e/ (to what extent that symbol reflects the true phonetic realization of

this vowel is not discussed here) which may support the occasional tendency to use that symbol also for Czech (e.g. by Nicolaidis, 2003; Lengeris & Hazan, 2010 in Greek; Fox et al., 1995; Cervera et al., 2001; Chládková et al., 2011 in Spanish; Hirata & Tsukada, 2009; Niimi et al., 1994; Kamiyama & Vaissière, 2009; Hirayama, 2003 in Japanese; Jones, 1953; Padgett, 2004; Lyakso et al., 2009 in Russian).

To conclude on the phonemic transcription motivated by acoustic results, it should be noted that even though the aim is to render the phonemic transcription as explicit as possible (i.e. truthfully reflecting the phonetic reality), different diacritics rendering any possible phonetic detail are still avoided. For instance, Šimáčková et al. (2012) employed two different length marks [:] and [.] to capture the different durations of the long high front vowel across two major dialects of Czech. Although we found here that the long high front vowel is shorter than the other long vowels, the durations of the long vowels are larger than the durations of the short vowel across all five vowel pairs; therefore, we represent the long phoneme by appending /:/ to the vowel symbol throughout for all the five Czech length contrasts. The long/short ratio reported here does not seem to be exceptionally small for the high front vowel pair, instead it seems to gradually decrease from mid to low and to high vowels. This could be understood as a physiologically conditioned duration-ratio phenomenon causing long vowels at the periphery of the vowel space to be sustained for a shorter amount of time than long vowels closer to the central part of the vowel space.

We should note here that other languages, too, lack a consensus on the phonological transcription of vowels. To name what is perhaps the most widely known instance, in order to transcribe lax/tense vowels in British English three main types of transcriptions have been used: quantitative transcription (using the same vowel symbol and appending a length mark, e.g. Palmer, 1920; Jones, 1932), qualitative transcription (using different symbols and no length mark, e.g. Ladefoged & Broadbent, 1957), or quantitative-qualitative transcription (using both, e.g. Cruttenden, 2014 and most contemporary authors). Another example is that of Japanese, in which the inconsistency concerns the phonemic notation of the back high vowel; many authors use /u/ (Hirata & Tsukada, 2009; Niimi et al., 1994; Kamiyama & Vaissière, 2009; Hirayama, 2003) but it is also possible to find the symbol /u/ (Lambacher et al., 2005), which reflects the unrounded phonetic realization of the vowel.

There have been debates on the correctness of the different notations. According to some authors, phonemic symbols should correspond to the most frequent allophones and only those differences which cannot be expressed in terms of phonological rules should be made explicit by using a specific phonemic symbol (Duchet, 1992). According to this point of view, marking vowel length in tense/lax English vowel pairs would be a redundant information, because it can be inferred. On the contrary, it is accepted that the different aforementioned notations were formed according to IPA principles and thus are all scientifically correct (Wells, 2011). It is then up to each author to decide which notation she or he will use. The choice should be determined by what is being the message and to what type of readers the study is addressed.

We adhere to the idea that different phonemic notations are acceptable and “right” in their own sense. This is why – despite having adopted a phonetically-motivated transcription in the current work – the authors of this paper are not strictly opposing a transcription of Czech vowels that would employ the symbols /i i: e e: a a: o o: u u:/ if the focus is

on a formal description of the system and if the study is not targeting an audience who might try to pronounce the vowels according to the notation (e.g. speakers with speech disorders undergoing formal training or learners of Czech as a second language). After all, for the language-learning child, the only important information that she extracts from the phonetic environment might be that there are ten different clusters or perhaps categories for vowels in the ambient language (and the scientist may arbitrarily choose to transcribe them in a non-IPA based alphabet as ♣△☹※⊙*✓☞☼◇) and perhaps the child sooner or later figures out that those ten discrete units are in fact a combination of, for instance, five times two category levels (such as ●○☆◀◄☞☼■□). While we still know little about how and when the developing child structures the phonetic vowel space in particular ways, the linguist has the knowledge, a particular aim, and the choice of how to appropriately convey their message. Crucially, whether an author's main aim is to reflect the phonetic reality, or whether it is to formalize and simplify, the approach she or he takes should be consistent and applied across all units of the system.

5. Conclusions

The present paper contributes a thorough spectral and durational characteristics of Czech vowels. Twenty speakers representative of the general, standard-Czech speaking population were recorded while spontaneously producing speech. Analyses of their vowels revealed that the mid front vowels are significantly lowered in the vowel space, appearing less distant in their F1 from the low vowels than from the mid back vowels. Confirming previous studies, the short high front vowel was found to be spectrally distinct from its long counterpart, namely, lowered along the F1 dimension. No such F1 differences were detected in the /u/-/u:/ vowel pair, which, instead revealed a significant difference in F2 with the short phoneme being fronter than the long one (and similarly for the /o/-/o:/ contrast). Whether this F2 distinction between short and long phonemes in back vowels is a feature of spontaneous speech or whether it is due to the consonantal context occurring in the present study remains to be shown in future work. Our data demonstrated that in spontaneous speech duration reliably distinguishes between short and long phonemes across all vowel pairs, including /i:/ vs /i/, which runs contrary to some recent speculations that the short-long contrast in high front vowels may no longer be (primarily) cued by duration (Šimáčková et al., 2012). The study concluded with a discussion of whether and how phonological transcription can best reflect an author's goal and help the reader understand the linguist's message.

ACKNOWLEDGMENTS

This work was funded by an internal grant from Charles University PRIMUS/17/HUM19 and by Czech Science Foundation grant no. 18-01799S. The authors thank Kristýna Hrdličková, Zuzana Oceláková, Martina Černá, and Radka Klimičková for help with data collection and annotation.

REFERENCES

- Borovičková, B. & Maláč, V. (1967). *The Spectral Analysis of Czech Sound Combinations*. Praha: Academia.
- Beneš, J. (1943). Porušování základní barvy českých samohlásek. *Náše řeč*, 27(3), 64–65.
- Bičan, A. (2013). *Phonotactics of Czech*. Frankfurt am Main: Peter Lang.
- Boersma, P. & Weenink, D. (2018). *Praat: doing phonetics by computer. Version 6.0.40*. Retrieved from www.praat.org.
- Cervera, T., Miralles, J. L. & Gonzalez-Alvarez, J. (2001). Acoustical analysis of Spanish vowels produced by laryngectomized subjects. *Journal of Speech, Language, and Hearing Research*, 44, 988–996.
- Chládková, K., Boersma, P. & Podlipský, V. J. (2009). Online formant shifting as a function of F0. *Proceedings of Interspeech 2009*, 464–467.
- Chládková, K., Escudero, P. & Boersma, P. (2011). Context-specific acoustic differences between Peruvian and Iberian Spanish vowels. *Journal of the Acoustical Society of America*, 130(1), 416–428.
- Chládková, K., Černá, M., Paillereau, N., Skarnitzl, R. & Oceláková, Z. (2019). Prenatal infant-directed speech: vowels and voice quality. In: *Proceedings of ICPhS 2019*, 1525–1529.
- Chlumský, J. (1928). *Česká kvantita, melodie a přízvuk*. Praha: Česká akademie věd a umění.
- Crothers, J. (1978). Typology and universals of vowel systems. In: J. Greenberg, C. A. Ferguson & E. A. Moravcsik (Eds.), *Universals of Human Language*, Vol. 2: *Phonology* (pp. 93–152). Stanford: Stanford University Press.
- Cruttenden, A. (2014). *Gimson's Pronunciation of English*. London and New York: Routledge.
- Dankovičová, J. (1997). Czech. *Journal of the International Phonetic Association*, 27, 77–80.
- Duběda, T. (2005). *Jazyky a jejich zvuky: univerzálie a typologie ve fonetice a fonologii*. Praha: Karolinum.
- Duchet, J. L. (1992). *La Phonologie*. Paris: P.U.F.
- Escudero, P., Boersma, P., Rauber, A. S. & Bion, R. A. H. (2009). A cross-dialect acoustic description of vowels: Brazilian and European Portuguese. *Journal of the Acoustical Society of America*, 126, 1379–1393.
- Fant, G. (1960). *The Acoustic Theory of Speech Production*. The Hague: Mouton.
- Fant, G. (1969). Formant frequencies of Swedish vowels. In: *Speech Transmission Laboratory, Quarterly Progress and Status Report 3* (pp. 94–99). Stockholm: Royal Institute of Technology.
- Frinta, A. (1909). *Novočeská výslovnost: pokus o soustavnou fonetiku jazyka českého*. Praha: Česká akademie císaře Františka Josefa pro vědy, slovesnost a umění.
- Frinta, A. (1925). *A Czech phonetic reader*. London: University of London Press.
- Flemming, E. (2019). Implications of [i] vowels for the theory of vowel inventories. Presented at *LSA Annual Meeting*, NYC.
- Fox, R. A., Flege, J. E. & Munro, M. J. (1995). The perception of English and Spanish vowels by native English and Spanish listeners: A multidimensional scaling analysis. *Journal of the Acoustical Society of America*, 97, 2540–2550.
- Gendrot, C., Adda-Decker, M. & Vaissière, J. (2008). Les voyelles /i/ et /y/ du français: focalisation et variations formantiques. In: *Proceedings of JEP*, 205–208.
- Grammont, M. (1933). *Traité pratique de prononciation française*. Paris: Delagrave.
- Hála, B. (1941). *Akustická podstata samohlásek*. Praha: Česká akademie věd a umění.
- Hála, B. (1955). *Výslovnost spisovné češtiny: její zásady a pravidla*. Praha: Nakladatelství Československé Akademie Věd.
- Hála, B. (1962). *Uvedení do fonetiky češtiny na obecně fonetickém základě*. Praha: ČSAV.
- Hirata, Y. & Tsukada, K. (2009). Effects of speaking rate and vowel length on formant frequency displacement in Japanese. *Phonetica*, 66, 129–149.
- Hirayama, M. (2003). Contrast in Japanese vowels. *Toronto Working Papers in Linguistics*, 20, 115–132.
- Jones, D. (1932). *An Outline of English Phonetics*. Leipzig: B. G. Teubner.
- Jones, L. G. (1953). The vowels of English and Russian: An acoustic comparison. *Word*, 9(4), 354–361.
- Kamiyama, T. & Vaissière, J. (2009). Perception and production of French close and close-mid rounded vowels by Japanese-speaking learners. *Acquisition et interaction en langue étrangère Aile... Lia*, 2, 9–41.

- Kamiyama, T. (2011). Pronunciation of French vowels by Japanese speakers learning French as a foreign language: Back and front rounded vowels /u y ø/. *Phonological Studies: Phonological Society of Japan*, 97–108.
- Kewley-Port, D. (1995). Thresholds for formant frequency discrimination of vowels in consonantal context. *Journal of the Acoustical Society of America*, 97(5), 3139–3146.
- Kuznetsova, A., Brockhoff, P. B. & Christensen, R. H. B. (2017). lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software*, 82(13), 1–26.
- Ladefoged, P. & Broadbent, D. E. (1957). Information conveyed by vowels. *Journal of the Acoustical Society of America*, 29, 98–104.
- Lambacher, S. G., Martens, W. L., Kakehi, K., Marasinghe, C. A. & Molholt, G. (2005). The effects of identification training on the identification and production of American English vowels by native speakers of Japanese. *Applied Psycholinguistics*, 26, 227–247.
- Lengeris, A. & Hazan, V. (2010). The effect of native vowel processing ability and frequency discrimination acuity on the phonetic training of English vowels for native speakers of Greek. *Journal of the Acoustical Society of America*, 128(6), 3757–3768.
- Lenth, R., Singmann, H., Love, J., Buerkner, P. & Herve, M. (2018). *Estimated marginal means, aka least-squares means*. R package version 1.2.4. Retrieved from <https://cran.r-project.org/package/emmeans>.
- Liljencrants, J. & Lindblom, B. (1972). Numerical simulation of vowel quality systems: The role of perceptual contrast. *Language*, 48, 839–862.
- Ludvíková, M. & Kraus, J. (1966). Kvantitativní vlastnosti soustavy českých fonémů. *Slovo a slovesnost*, 27, 334–344.
- Lyakso, E., Frolova, O. & Grigorev, A. (2009). The acoustic characteristics of Russian vowels in children of 6 and 7 years of age. In: *Proceedings of Interspeech 2009*, 1739–1742.
- Nicolaidis, K. (2003). Acoustic variability of vowels in Greek spontaneous speech. In: *Proceedings of the 15th International Congress of Phonetic Sciences*, 3221–3224.
- Niimi, S., Kumada, M. & Niitsu, M. (1994). Functions of tongue-related muscles during production of the five Japanese vowels. *Ann. Bull. R. I. L. P. Univ. Tokyo*, 28, 33–40.
- Padgett, J. (2004). Russian vowel reduction and dispersion theory. *Phonological Studies*, 7, 81–96.
- Pailhereau, N. (2016). ‘Identical’ vowels in L1 and L2? Criteria and implications for L2 phonetic teaching and learning. In: Liszka, S. A., Leclercq, P., Tellier, M. & Daniel, G. (Eds.), *EUROSLA Yearbook 2016* (pp. 144–178). Amsterdam: John Benjamins.
- Pailhereau, N. & Skarnitzl, R. (2019). An acoustic-perceptual study on Czech monophthongs. In: T. Radeva-Bork & P. Kosta (Eds.), *Current developments in Slavic Linguistics. Twenty years after (based on selected papers from FDSL11)* (pp. 453–466). Berlin: Peter Lang.
- Palková, Z. (1997). *Fonetika a fonologie češtiny*. Praha: Karolinum.
- Palmer, H. E. (1920). *A First Course of English Phonetics*. Cambridge: W. Heffer and Sons Ltd.
- Podlipský, V. J., Skarnitzl, R. & Volín, J. (2009). High front vowels in Czech: A contrast in quantity or quality? In: *Proceedings of Interspeech 2009*, 132–135.
- Podlipský, V. J., Chládková, K. & Šimáčková, Š. (in press). Spectrum as a perceptual cue to vowel length in Czech, a quantity language. *Journal of the Acoustical Society of America: Express Letters*.
- R Development Core Team (2008). *R: A language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing. Retrieved from <https://www.r-project.org>.
- Skarnitzl, R. & Volín, J. (2012). Referenční hodnoty vokálních formantů pro mladé dospělé mluvčí standardní češtiny. *Akustické listy*, 18, 7–11.
- Skarnitzl, R. (2012). Dvoji *i* v české výslovnosti. *Naše řeč*, 95, 141–153.
- Skarnitzl, R., Šturm, P. & Volín, J. (2016). *Zvuková báze řečové komunikace: fonetický a fonologický popis řeči*. Praha: Karolinum.
- Stevens, K. N. & House, A. S. (1963). Perturbations of vowel articulations by consonantal context: an acoustical study. *Journal of Speech and Hearing Research*, 6, 111–128.
- Šimáčková, Š., Podlipský, V. J. & Chládková, K. (2012). Czech spoken in Bohemia and Moravia. *Journal of the International Phonetic Association*, 42, 225–232.
- Vaissière, J. (2009). Articulatory modeling and the definition of acoustic-perceptual targets for reference vowels. *The Chinese Phonetics Journal*, 2, 22–33.

- Vaissière, J. (2011). On the acoustic and perceptual characterization of reference vowels in a cross-language perspective. In: *Proceedings of ICPHS 2011*, 52–59.
- Wells, J. (2001). IPA transcription systems for English. Retrieved from <https://www.phon.ucl.ac.uk/home/wells/ipa-english-uni.htm> (last accessed on May 16, 2019)

RESUMÉ

Článek se zaměřuje na akustickou analýzu českého vokálního systému a popisuje spektrální a délkové charakteristiky 10 českých monoftongů. Oproti dosavadní literatuře jsou studovány hlásky z projevu spontánního a sesbíraného jak od mužských tak od ženských mluvčích, kteří byli rekrutováni ze široké veřejnosti (nejedná se tedy pouze o studenty nebo reportéry médií, jejichž čtený projev byl předmětem zkoumání v předchozích studiích). Spektrální analýza zahrnuje první čtyři formanty (F1 až F4). Analýza hláskové délky se zaměřuje na poměrné trvání dlouhých vůči krátkým vokálům a srovnává trvání samohlásek normalizované pro délku slov. Výsledky potvrzují významný spektrální rozdíl u vysokých předních samohlásek, kde je fonologicky krátké /ɪ/ realizováno s vyšším F1 než dlouhé /i:/. Přední střední vokály /e/ a /ɛ:/ jsou realizovány s relativně vysokým F1, kvalitativně jsou tak dokonce blíže k nízkým samohláskám /a/ a /a:/ než k zadním středním samohláskám /o/ a /o:/. Novým zjištěním je, že krátké zadní samohlásky /o/ a /u/ mají vyšší F2 než jejich dlouhé protějšky: tento mírný posun dopředu lze pravděpodobně připsat spontánnímu stylu řeči a také převážně koronálnímu kontextu, ve kterém se samohlásky objevovaly. Na rozdíl od moderní literatury, která uvádí velmi malý poměr trvání dlouhých vůči krátkým vysokým vokálům, naše výsledky ukazují, že trvání konzistentně rozlišuje fonologickou délku napříč všemi pěti vokálními páry: dlouhé samohlásky jsou v průměru 1,76krát delší než samohlásky krátké. Diskuze je uzavřena zamyšlením se nad vztahem akustických vlastností hlásek a jejich fonologickou transkripcí.

Nikola Paillereau

Kateřina Chládková

E-mail: nikola.paillereau@ff.cuni.cz

Institute of Phonetics

Faculty of Arts, Charles University

Prague, Czech Republic

Institute of Psychology, Czech Academy of Sciences

Prague, Czech Republic