2025 ACTA UNIVERSITATIS CAROLINAE PAG. 157–175
PHILOLOGICA 3

ACOUSTIC ANALYSIS OF VOWELS IN CZECH DISYLLABIC WORDS PRODUCED BY L1-GERMAN SPEAKERS

ANNA CHABROVÁ

ABSTRACT
This study presents the results of an acoustic analysis of vowels produced by L1-German speakers learning Czech as a foreign language and provides a brief overview of vowel behaviour in Czech and German. The analysed vowels are /a aː ɪ iː u uː/ in Czech disyllabic words, and the speakers are eight women with varying levels of proficiency. Vocalic formants F1 and F2 were analysed, and the differences in formant values between long and short vowels were calculated. Furthermore, vowel duration was measured, and differences between the durations of long and short vowels were assessed. The results are compared with reference values for female native speakers of Czech and German and indicate that speakers in the present study do not sufficiently distinguish between short and long vowels in Czech. Additionally, the absolute vowel durations observed in this study are longer than the reference values for Czech speakers.

Keywords: Czech as L2; German as L1; vowel length; vowel duration; vowel formants; disyllabic words; acoustic analysis

1. Introduction

Pronunciation is one of the key areas that speakers must master when learning a new language. Incorrect pronunciation can significantly hinder communication and may even lead to complete misunderstanding. The present study is part of a broader research project conducted at the Institute of Phonetics at the Faculty of Arts, Charles University, which focuses on the phonetic characteristics of non-native Czech speech. Experimental findings have confirmed, among other things, difficulties in the realization of vowel length in non-native speakers whose first language is Russian, Ukrainian, or Polish (e.g., Palková et al., 2020; Veroňková & Bořil, 2020a, 2020b; Veroňková et al., 2020). The present experiment focuses on vowel length in Czech disyllabic words produced by native German speakers.

Vowels in German and Czech differ in several aspects. Czech has five short monophthongs, five long monophthongs, and three diphthongs. Short and long vowels form phonologically distinctive pairs, and all Czech vowels are primarily lax and non-nasalized (Palková, 1994: 172). The short vowels are /ɪ e a o u/ and the long vowels are /iː eː aː oː uː/. Vowel pairs /e eː/, /a aː/ and /o oː/ do not differ significantly in quality, but there are

https://doi.org/10.14712/24646830.2025.26

noticeable differences in the vowels /ɪ iː/ and /u uː/. In the case of /ɪ iː/, the difference is so substantial (cf. e.g. Podlipský et al., 2009; Skarnitzl & Volín, 2012; Paillereau & Chládková, 2019) that separate transcription symbols are now used to distinguish the short and long variants not only for comparative purposes, but also for transcription within the Czech phonological system. Skarnitzl and Volín (2012) worked with recordings of several dozen university students reading a text in a recording studio, and their results serve as reference values for Czech native speakers. Therefore, I use their study as a reference in this paper.

Regarding vowel quantity, older publications report that long vowels are twice as long as short ones (Palková, 1994: 179), but more recent findings suggest a different picture. Skarnitzl (2012), who analysed recordings of professional speakers in broadcast news and worked with non-normalized durations in his study, determined the following duration ratios between the long and short variants of each vowel (see Table 1, on the left). Paillereau and Chládková (2019) who analysed normalized durations in spontaneous speech produced by non-professional speakers, reported different ratios, particularly for /ɪ iː/ (see Table 1, on the right). Their participants came from diverse social backgrounds and the recordings were based on spontaneous speech.

However, both studies agree that long vowels are less than twice as long as their short counterparts, and that the durational contrast is smaller for /ɪ iː/ and /u uː/ compared to the other vowel pairs. The reason is the qualitative difference between the short and long variants of these vowels, which facilitates perceptual discrimination and thus allows for a reduction in durational contrast (Skarnitzl, 2012: 151).

Table 1 Duration ratios between the long and short vowel in each vowel pair according to Skarnitzl (2012) and Paillereau and Chládková (2019).

	V: / V ratio	
vowel pair	Skarnitzl	Paillereau & Chládková
iː / ɪ	1.29	1.66
eː / e	1.72	1.78
aː / a	1.79	1.73
oː / o	1.73	1.87
uː / u	1.60	1.65

In German, the situation is different and vowel quality and quantity are closely interconnected. Unlike in Czech, vowel length in German is not indicated by diacritical marks. German also has more vowel qualities than Czech, and these are often directly linked to a specific realization of length, e.g., the vowel /ʊ/ is always phonologically short (Becker, 2012: 31). However, some vowel qualities exist in both short and long forms, and in such cases the speaker must know whether the vowel in a given word should be pronounced long or short, e.g., the vowel /u/ can occur in both long and short forms (Becker, 2012: 31).

Vowel quality and quantity in German are also linked to word stress. The distribution of certain vowels is restricted to either stressed or unstressed syllables, e.g., the tense long vowel /i:/ occurs only in stressed syllables, whereas the reduced vowel /ə/occurs only in unstressed ones (Kleiner & Knöbl, 2015: 32). According to many authors (e.g., Kleiner & Knöbl, 2015: 32), only short vowels can occur in unstressed syllables in native German words. However, vowel duration in German should be considered not only in absolute (i.e., short vs. long) but also in relative terms. Some authors (e.g., Becker, 2012; Jessen, 1993) assume that stressed syllables may be relatively longer than unstressed ones, and that this relatively longer duration contributes to their prominence. This may have significant implications for L1-German speakers learning Czech, as such a situation does not occur in Czech.

In Czech, word stress is fixed, typically falling on the first syllable of the word, and it has no influence on vowel length or quality (Ashby & Maidment, 2015: 135; Skarnitzl, 2008: 199–200). Both short and long vowels can occur in stressed as well as unstressed syllables (e.g., /pla:nɪ/ 'plans' and /plani:/ 'wild'). In the case of disyllabic words, German and Czech typically agree in placing stress on the first syllable (Kleiner & Knöbl, 2015: 59).

The present study is part of a broader experiment, the aim of which was to obtain data on the realization of vowel length and quality in Czech disyllabic words produced by native German speakers, using both a perception test and acoustic analysis. Deviations from the canonical form are expected, and different structural types are likely to exhibit different patterns of behaviour.

This paper presents the results of the acoustic analysis of vowel duration and formant measurements. The results of the perception test, along with summarised findings on vowel duration, were already published in Chabrová & Veroňková (2022). The present study expands and completes the analysis of vowel duration.

The study by Chabrová & Veroňková (2022) confirmed the assumption that vowel length in Czech disyllabic words poses a challenge for native German speakers – native Czech speakers perceived only 42% of the words in accordance with the original text, which was read by the German speakers. For the remaining 58%, Czech listeners perceived length other than the originally correct and intended one. A summary of the perception test results is presented in Table 2. Rows, labelled original, indicate vowel length in the original disyllabic word as written in the text read by German speakers. Columns, labelled perceived as, show the vowel length perceived by Czech listeners in the perception test. Bolded cells correspond to words perceived in agreement with the original text, i.e., pronounced correctly.

Table 2 Overall percentage distribution of how the original structures SS/SL/LS/LL were perceived. S = short vowel, L = long vowel. In Chabrová & Veroňková (2022).

(%)	perceived as SS	perceived as SL	perceived as LS	perceived as LL
original SS	**40.3**	25.1	21.8	12.8
original SL	28.1	**27.1**	26.7	18.1
original LS	12.3	9.0	**64.8**	13.9
original LL	17.5	14.1	32.5	**35.9**

Speakers achieved the highest success rate with the LS structure (65%), i.e., a long vowel in the first, stressed syllable and a short vowel in the second, unstressed syllable. On the contrary, the lowest success rate was found for the SL structure (27%), where the first syllable contained a short vowel and the second syllable a long vowel. The success rate for short vowels alone was 65% and for long vowels alone, 61%. The following factors influenced the success of vowel realization: the type of short/long structure in the target word (i.e., short-short = SS, short-long = SL, long-short = LS, long-long = LL), vowel quality, the position of the vowel in the first or second syllable (for the analysed items, this data could only be obtained for vowels /ɪ iː/, since in the material used, /a aː/ occurred only in the first syllable and /u uː/ only in the second syllable). The frequency of occurrence was mapped for the word forms used and their lemmas. Due to the smaller size of the spoken language corpus, which did not include many of the target items, the SYN2020 corpus of written language was used (Křen et al., 2020). The frequency of the lexemes/word forms may have influenced the evaluation in individual cases, but did not affect the overall trends.

2. Method

2.1 Speakers

The source material consisted of read-aloud recordings from eight native speakers of German. All participants were women[1] (aged 21–38) with varying levels of Czech language proficiency (five intermediate speakers, estimated level A2–B1 according to the CEFR; three advanced speakers, estimated level B2–C1 according to the CEFR). The intermediate speakers had been learning Czech for one to two years at the time of recording, had completed at least one year of Czech Studies, and were residing in the Czech Republic during the recording period. The advanced speakers had been learning Czech for four or more years at the time of recording and were long-term residents of the Czech Republic. One of the speakers was from Austria, and the remaining seven were from Germany.

2.2 Material

The recordings[2] were made in the studio of the Institute of Phonetics at the Faculty of Arts, Charles University, using an AKG C 4500 B-BC microphone. The audio was record-

[1] Only female speakers were included in the experiment, as recordings from too few male speakers were available, and their voices would have been too easily identifiable among the female speakers in the perception test.

[2] The recordings used in this study came from two sources: original recordings made by the author and recordings from the corpus of non-native Czech speech compiled at the Institute of Phonetics, Faculty of Arts, Charles University. The material for this corpus was recorded as part of the Czech science foundation grant project GA ČR 18-18300S *Zvukové vlastnosti češtiny v komunikaci nerodilých a rodilých mluvčích*. The recordings conducted by the author served as an extension of the corpus using identical texts.

ed at a sampling rate of 48 kHz with 16-bit quantization and was saved and processed in WAV format.

The recording text consisted of sentence pairs. Each pair included a first sentence, which served merely to establish a context, and a second sentence containing the target phenomenon. The two sentences within each pair were semantically related, whereas different sentence pairs were not related to each other. Each pair was presented on a new line, and the speakers read one A4 page at a time, meaning they could only see one sheet at once. They were allowed to read and prepare this sheet just before recording, then handed it in and received the next one for preparation. This procedure was chosen to minimize the risk of speakers noticing patterns between words or sentences and identifying the focus of the study. At the same time, the short preparation ensured fluent delivery with minimal hesitations or errors. Informal interviews conducted after the recordings confirmed that the speakers did not identify the focus of the study.

Suitable carrier sentences containing the target words were selected from the recordings. The target items were disyllabic words forming groups of four, three or two that differed only in vowel length (e.g., a group *sazi*, *sazí*, *sází*). All syllables in the target words were open, to avoid any potential influence of syllable structure (open vs. closed) on vowel duration. A total of 27 carrier sentences were used, each containing one target word in an unambiguous context. The target words were never located at the very beginning or end of a sentence, and sentence pairs containing target words from the same group (e.g., the group *sazi*, *sazí*, *sází*) were not placed close to one another to mask the target phenomenon.

The set of target words includes the following items:

a) 3 groups of four: myli, milí, mýlí, mílí; platu, platů, plátu, plátů; valy, valí, vály, válí;
b) 3 groups of three: sazi, sazí, sází; spali, spály, spálí; vazu, vazů, vázu;
c) 3 groups of two: kraji, krájí; planý, plány; sliby, slíbí.

An example of a group of four words platu / platů / plátu / plátů used in carrier sentences:

(1) Nebyla na tom špatně. K jeho **platu** dostala ještě podporu od pojišťovny. (SS)
*She wasn't doing badly. In addition to his **salary**, she also got support from the insurance company.*
(2) Ředitel se dohodl s odbory. V prosinci dostali šest **platů** jako bonus. (SL)
*The director reached an agreement with the unions. In December, they received six **salaries** as a bonus.*
(3) Desku tvaroval podle vzoru. Na jednom **plátu** oceli pracoval dva dny. (LS)
*He shaped the plate according to the pattern. He worked on one steel **plate** for two days.*
(4) Potřebovali vyplnit mezeru. Dohromady spojili pět **plátů** železa. (LL)
*They needed to fill the gap. Altogether, they joined five iron **plates**.*

Despite their different spelling, the letters y/ý and i/í are pronounced identically in Czech: y and i are both realized as the vowel [ɪ], while ý and í are both pronounced as [iː]. Similarly, the letters ú and ů both represent the long vowel [uː]; the difference between them is only graphical and depends on the vowel's position within the word.

Target words were extracted from the recordings, and after excluding items with disturbing, irremovable noise, a set of 203 stimuli was obtained, which was an adequate

number for the length of the perception test. The set contains a balanced representation of the four possible combinations of vowel length in disyllabic words (SS, SL, LS, LL) from all eight speakers. The vowel quality combinations in the target words are a–i, a–u, and i–i. Words containing e and o were intentionally excluded from the experiment because the long vowels /o:/ and /e:/ are located on the periphery of the phonological system (Vachek, 1968: 30–34).

2.3 Acoustic analysis

For the acoustic analysis, vowel durations in the words from the perception test were measured using Praat (Boersma & Weenink, 2020), and their formant values were obtained through automatic extraction followed by manual verification. For duration measurements, vowel boundaries were marked in two different ways. The first segmentation approach followed the recommended guidelines for segmenting phonemes (Machač & Skarnitzl, 2009), with vowel boundaries placed according to the formant structure. The second segmentation approach placed greater emphasis on perception and was guided visually by oscillogram: in relevant cases, the vowel boundary was shifted to include a voicing offset following the end of the vowel's articulation itself (Machač & Skarnitzl, 2009: 136–137).

Vowel duration was normalized relative to the articulation rate of the entire word using the method employed by Veroňková & Bořil (2020b), in order to allow for comparisons across different speakers. For each target word, the articulation rate in syllables per second was calculated, based on which the average articulation rate for each speaker was determined. Normalized duration was obtained by multiplying the actual duration of the word by average articulation rate of the respective speaker, and the resulting value was then divided by the overall average articulation rate across all speakers.

In the acoustic analysis of formants, only the first variant of segmentation was used, that is, the one according to Machač & Skarnitzl (2009). Formants F1 and F2 were first measured automatically in Praat in the middle third of vowel duration. The formants were extracted using the Burg method (time steps: automatic; maximum number of formants: 5; formant ceiling: 5500 Hz; window length: 0.025 s; pre-emphasis from: 50 Hz) using a script (Bořil, 2015). Since errors can occur in the automatic extraction of formant values, the obtained values were manually checked by comparing them with reference data. The reference used was the study by Skarnitzl & Volín (2012), which includes data from dozens of speakers and offers high-quality reference values for male speakers, including standard deviations. Based on these values, I calculated reference ranges for F1 and F2 of the respective vowel for female speakers. Any values outside these ranges were then manually verified and, if necessary, corrected. The reference ranges were calculated as follows: one standard deviation was added to and subtracted from the average male formant values, resulting in a reference range for male speakers. In accordance with generally observed patterns, formant frequencies for female speakers are approximately 15–20% higher than those for male speakers (Skarnitzl & Volín, 2012: 8). In order to obtain the reference ranges for female speakers, I increased the values of the male reference ranges by 17.5%.

3. Results

3.1 Formant analysis results

Table 3 presents the results of formant analysis in both hertz and ERB. For values in ERB, the percentage difference between formants of short and long vowels is also given. The conversion to ERB was performed using formula[3] $21.4 \times \log10(0.00437 \times f + 1)$ (Glasberg & Moore, 1990). Columns labelled F1/F2 difference show the difference in F1/F2 between short and long variants of the respective vowel.

Table 3 Mean F1 and F2 values in the analysed material.

	F1 (Hz)	F2 (Hz)	F1 (ERB)	F2 (ERB)	F1 difference (% of ERB)	F2 difference (% of ERB)
a	840.27	1457.23	14.33	18.56	2.16	0.84
a:	877.22	1429.26	14.64	18.41		
ɪ	370.81	2540.66	8.95	23.17	5.12	0.70
i:	343.30	2589.92	8.52	23.34		
u	373.64	1024.67	9.00	15.81	5.67	5.88
u:	343.15	911.70	8.51	14.93		

Differences between the formants of short and long variants of the respective vowels are small. For /a a:/, the long variant has a slightly higher F1, while there is almost no difference in F2. In case of /ɪ i:/, the short variant has a somewhat higher F1 and a lower F2, but in absolute terms, these differences remain small. The most differentiated are vowels /u u:/, where the short variant has both higher F1 and F2, and is therefore slightly more centralized than the long variant.

Tables 4a and 4b present formants of vowels /ɪ i:/, divided according to whether the vowel occurs in the first or the second syllable.

Table 4a Mean values of F1 and F2 for vowels /ɪ i:/ in the *first* syllable.

1st syll.	F1 (Hz)	F2 (Hz)	F1 (ERB)	F2 (ERB)	F1 difference (% of ERB)	F2 difference (% of ERB)
ɪ	362.05	2531.45	8.82	23.14	3.41	0.80
i:	343.87	2587.26	8.53	23.33		

[3] For the conversion from Hz to ERB, an online converter was used (ERB-rate scale converter, n.d., University College London. Retrieved 5. 6. 2025 from http://www.homepages.ucl.ac.uk/~sslyjjt/speech/erb.html).

Table 4b Mean values of F1 and F2 for vowels /ɪ iː/ in the *second* syllable.

2nd syll.	F1 (Hz)	F2 (Hz)	F1 (ERB)	F2 (ERB)	F1 difference (% of ERB)	F2 difference (% of ERB)
ɪ	373.42	2543.39	8.99	23.18	5.63	0.67
iː	343.13	2590.72	8.51	23.34		

While values for long /iː/ are the same in the first and second syllable, the values for short /ɪ/ differ slightly. In the second syllable, the short vowel shows slightly higher F1 and F2 values, resulting in a slightly greater percentage difference between the short and long vowel in F1 compared to the first syllable. Overall, however, the differences between the first and second syllable remain small.

The formant values of all vowels are graphically displayed in Figure 1.

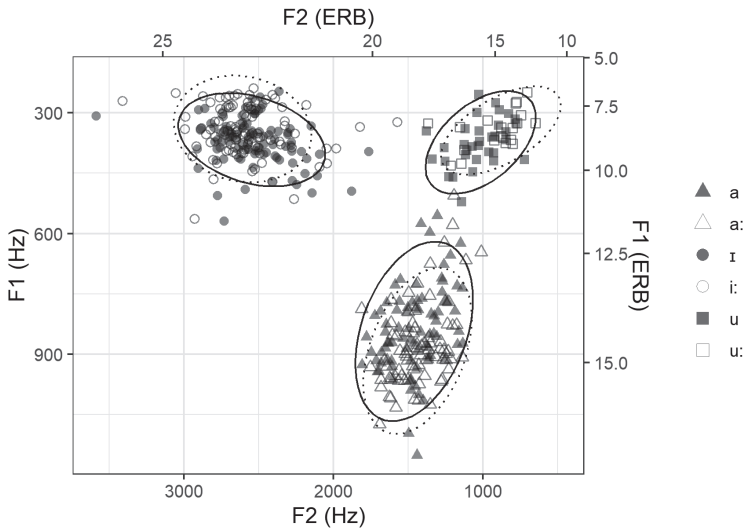


Figure 1 Formant values of vowels in Hz and ERB; ellipses cover 95% of the values. Solid ellipses represent short vowels and dotted ellipses represent long vowels.

The results were further filtered in Tables 5a–5d by combinations of vowel quality and quantity in order to analyse whether the vowel quantity patterns SS (short-short), SL (short-long), LS (long-short), and LL (long-long) differ from each other. Values are presented in both hertz and ERB.

F2 formants for /a aː/ are again nearly identical, but differences can be observed in F1. While in the SS pattern the values are closer to those of the short variant, in SL, LS, and LL patterns they correspond more to the long variant. This would be expected in the LS and LL patterns, where a long vowel is indeed supposed to occur, but in the SL pattern, this represents a deviation, as a short vowel is expected. As for /u uː/, the values are relatively consistent: the short variant shows relatively higher F1 and F2 values, while the long variant shows relatively lower ones. Vowels /ɪ iː/ behave quite consistently as well.

Table 5 Formant values for the a) SS, b) SL, c) LS, d) LL quantity pattern grouped by vowel quality combinations. S = short vowel, L = long vowel.

a) SS	1st vowel				2nd vowel			
	Hz		ERB		Hz		ERB	
	F1	F2	F1	F2	F1	F2	F1	F2
a - ɪ	824.10	1492.03	14.19	18.75	377.79	2526.79	9.06	23.13
a - u	809.07	1402.47	14.05	18.25	376.40	1048.60	9.04	15.98
ɪ - ɪ	361.14	2565.50	8.80	23.26	344.79	2552.86	8.54	23.21

b) SL	1st vowel				2nd vowel			
	Hz		ERB		Hz		ERB	
	F1	F2	F1	F2	F1	F2	F1	F2
a - i:	872.83	1489.83	14.61	18.74	349.87	2653.52	8.62	23.54
a - u:	853.71	1390.29	14.44	18.18	340.50	948.64	8.47	15.22
ɪ - i:	363.63	2471.88	8.84	22.94	337.25	2650.25	8.42	23.53

c) LS	1st vowel				2nd vowel			
	Hz		ERB		Hz		ERB	
	F1	F2	F1	F2	F1	F2	F1	F2
a: - ɪ	874.09	1430.27	14.62	18.41	386.00	2523.95	9.19	23.12
a: - u	896.38	1415.88	14.80	18.33	374.94	996.25	9.02	15.59
i: - ɪ	343.00	2583.38	8.51	23.32	376.13	2622.50	9.04	23.44

d) LL	1st vowel				2nd vowel			
	Hz		ERB		Hz		ERB	
	F1	F2	F1	F2	F1	F2	F1	F2
a: - i:	871.06	1439.03	14.59	18.46	341.19	2549.29	8.48	23.20
a: - u:	871.38	1415.38	14.59	18.33	347.63	887.25	8.59	14.73
i: - i:	344.33	2589.33	8.53	23.34	340.33	2554.73	8.47	23.22

The short variant tends to have higher F1 and lower F2 values, while the long variant tends to have lower F1 and higher F2 values. However, these differences are relatively small, and particularly for F2, the pattern is less regular, with more deviations from the described tendency.

3.2 Duration measurement results

All values presented below are normalized (see Method). Table 6a shows durations according to the first segmentation variant (i.e., based on formant structure), while Table 6b presents durations according to the second segmentation variant (i.e., based on oscil-

logram). In addition to the mean duration, standard deviation and the ratio between long and short variant of the respective vowel are provided. Vowels in the table are classified as short or long based on the length indicated in the original text, not on perception.

Table 6a Normalized duration of individual vowels, segmentation based on formants (variant 1).

segmentation 1	mean dur. (ms)	SD (ms)	V:/V ratio
a	121.35	34.82	1.36
a:	165.41	42.70	
ɪ	113.88	37.20	1.08
i:	122.56	41.73	
u	111.13	43.77	1.30
u:	144.62	49.62	

Table 6b Normalized duration of individual vowels, segmentation based on oscillogram (variant 2).

segmentation 2	mean dur. (ms)	SD (ms)	V:/V ratio
a	131.03	35.76	1.33
a:	174.39	39.32	
ɪ	130.77	45.83	1.08
i:	141.29	47.27	
u	135.84	49.56	1.27
u:	172.59	53.52	

Although the specific duration values differ for segmentation variants 1 and 2, the ratios between the duration of long and short vowels are very similar for both. A higher ratio can be observed for /a a:/ and /u u:/, and a lower ratio for /ɪ i:/. For this reason, I will continue to present results only according to the first segmentation variant, which follows the established segmentation rules (see Method) and is therefore considered default.

Vowels /ɪ i:/ appeared in both the first and second syllable in the analysed material (see Table 7). The data show that the behaviour of /ɪ i:/ in the first and second syllable differed. In the first syllable, the difference between short and long vowel variants is noticeable, and the long to short vowel duration ratio is comparable to the values in Table 3. However, in the second syllable, short and long vowel variants have the same duration (which corresponds to the duration of long vowel in the first syllable). I was interested in whether these differences would also be reflected in listeners' perception (for more information about the listeners and the perception test see Chabrová & Veroňková, 2022). However, the agreement between listeners' perception and the original text for /ɪ i:/ vowels, considered separately in the first and second syllable, differs only slightly (66% in the first syllable and 59% in the second syllable). The behaviour of /ɪ i:/ according to the position is also noted by Podlipský, Skarnitzl, & Volín (2009), although in their study this concerned the final position of an utterance compared to non-final positions. In the present experiment, the words were never positioned at the edges of sentences.

Table 7 Normalized mean duration of /ɪ i:/ vowels grouped by first and second syllable.

	1st syllable			2nd syllable		
	mean dur. (ms)	SD (ms)	V:/V	mean dur. (ms)	SD (ms)	V:/V
ɪ	87.19	20.81	1.35	121.82	37.30	1.02
i:	117.88	26.35		123.98	45.27	

The results of the duration analysis were further divided according to quantity patterns (SS, SL, LS, LL) and, within each pattern, by vowel quality combinations, as shown in Tables 8a–8d. In order to present all vowel combinations within a single table, I use the labels V1 (first vowel of the word) and V2 (second vowel of the word) instead of specifying the exact vowel quality.

Table 8 Normalized durations of vowels in the a) SS, b) SL, c) LS, d) LL quantity pattern. S = short vowel, L = long vowel.

a) SS	duration V1 (ms)	duration V2 (ms)
a - ɪ	119.21	113.99
a - u	102.71	119.20
ɪ - ɪ	133.49	131.31

b) SL	duration V1 (ms)	duration V2 (ms)
a - i:	133.49	122.50
a - u:	125.81	124.10
ɪ - i:	90.52	95.66

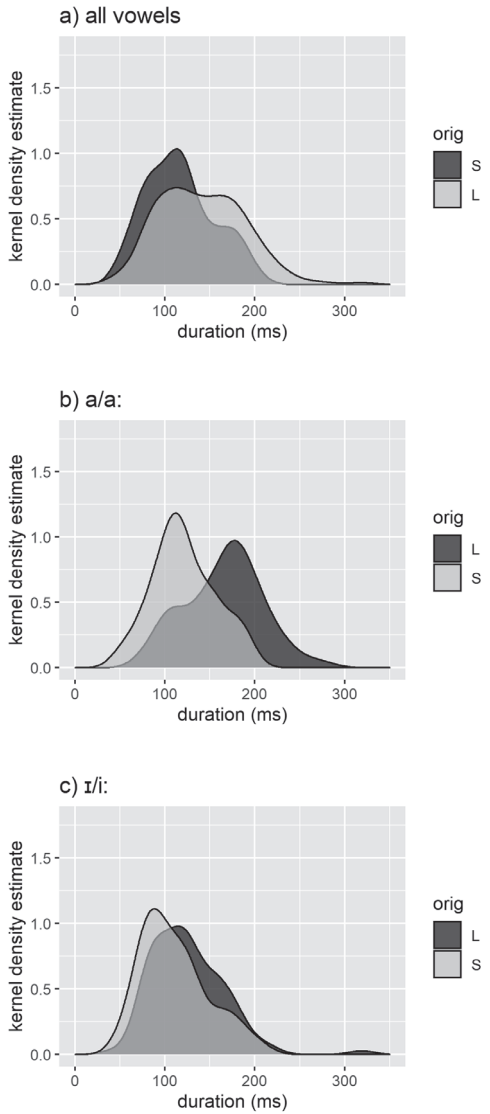
c) LS	duration V1 (ms)	duration V2 (ms)
a: - ɪ	179.88	121.68
a: - u	155.45	108.81
i: - ɪ	140.48	132.08

d) LL	duration V1 (ms)	duration V2 (ms)
a: - i:	169.92	121.43
a: - u:	128.06	161.69
i: - i:	110.31	143.00

The data show that individual vowels behave differently across various quantity patterns (SS, SL, LS, LL) and even within different quality combinations of the same pattern. The vowels /a a:/, which always occur in the first syllable in the analysed words, are longer in all quantity patterns when followed by /ɪ i:/ in the second syllable than when followed by /u u:/ (however, the vowels were not adjacent; there was always a consonant between them). The intended short /a/ is overall longer in the SL pattern than in the SS pattern, and the intended long /a:/ is overall longer in the LS pattern than in the LL pattern.

Vowels /u u:/, which always occur in the second syllable, behave in the opposite manner. Short /u/ is longer in the SS pattern compared to LS, and the long /u:/ is longer in the LL pattern compared to SL. The vowels /ɪ i:/ behave inconsistently, and no clear patterns can be observed.

Figures 2a–2d show a graphical representation of the relationship between vowel duration and whether the vowel was originally supposed to be short or long according to the source text. The graph demonstrates a significant overlap between the areas of short and long vowels, indicating that vowels originally classified as short and long were pronounced with similar durations.



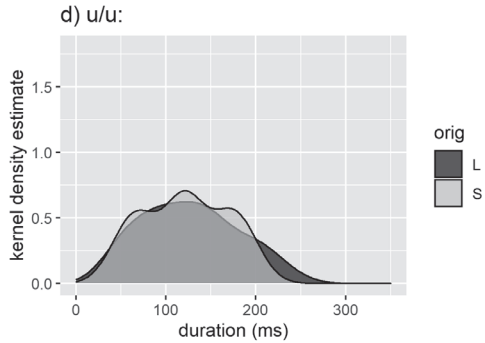
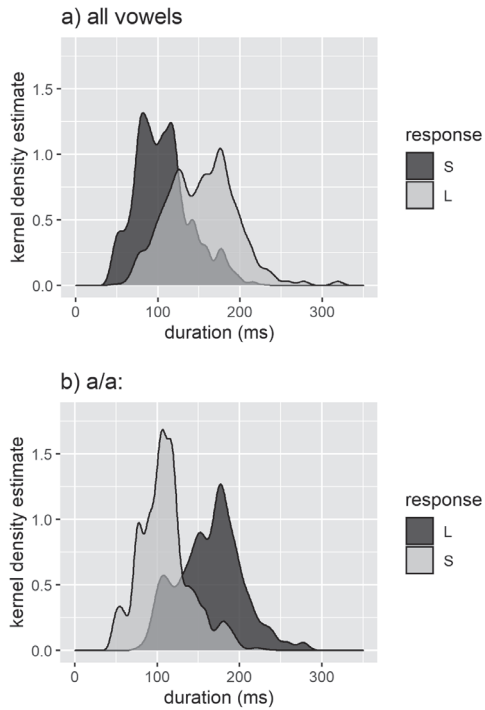


Figure 2 Relationship between normalized vowel duration (ms) of a) all vowels, b) /a aː/, c) /ɪ iː/, d) /u uː/ and the original classification of vowels as short (S) or long (L) according to the source text.

Figures 3a–3d illustrate the relationship between vowel duration and whether the vowel was perceived by the native Czech listeners as short or long. Compared to Figures 2a–2d, a difference is noticeable: the areas for short and long vowels are more distinct, and both variants show clearer peaks. This indicates that listeners were largely guided by duration when perceiving vowel length. However, the S and L areas still overlap to a large extent, which means that vowels of the same duration were often evaluated both as short and long – this could be either due to low inter-listener agreement, or because listeners also relied on other cues, such as vowel quality.



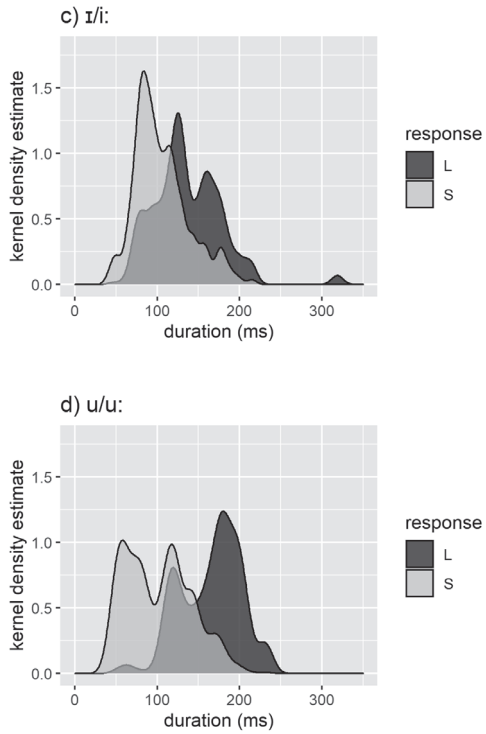


Figure 3 Normalized vowel duration (ms) of a) all vowels, b) /a a:/, c) /ɪ i:/, d) /u u:/ according to whether it was perceived by Czech listeners as short (S) or long (L).

4. Discussion

The previous section presented the results of acoustic analysis of formants and duration. I will now compare my results with reference values. As for the formant measurements, in this experiment I worked with reference data for male speakers published by Skarnitzl & Volín (2012), which I adjusted by increasing the values by 17.5%. At the time the experiment was conducted, this was the best available option. However, I now also have access to reference values for female speakers, which come from the above-mentioned study and have not yet been published, see Table 9. I would like to thank R. Skarnitzl for kindly providing these unpublished reference values (Hz) for female speakers. The conversion of values from Hz to ERB and calculation of the percentage difference between formants of the short and long vowel variants were carried out by the author of this paper; for details on the conversion method, see the Method section.

The comparison of the experimental results (see Table 3) with the reference values above shows that in the analysed material, native speakers of German do not exhibit such large differences in quality between the short and long vowel variants. As for /a a:/, the native speakers of German approximate the reference values well and pronounce the

Table 9 Unpublished F1 and F2 reference values for selected vowels in female Czech speakers, measured in the study by Skarnitzl & Volín (2012), rounded to the nearest ten. The ERB values and F1/F2 difference were calculated by the author of this paper.

	F1 (Hz)	F2 (Hz)	F1 (ERB)	F2 (ERB)	F1 difference (% of ERB)	F2 difference (% of ERB)
a	770	1500	13.70	18.79	1.93	2.40
a:	800	1420	13.97	18.35		
ɪ	490	2250	10.64	22.14	28.19	5.26
i:	330	2600	8.30	23.37		
u	420	1140	9.69	16.62	14.54	22.21
u:	340	760	8.46	13.60		

short and long variants essentially the same. Problems arise with vowels /ɪ iː/ and /u uː/. The results of the present study show that the speakers are going in the right direction when it comes to the realisation of the quality of these vowels – a difference in F1 values can be observed for /ɪ iː/, and in both F1 and F2 values for /u uː/. However, in order for the difference in quality to approach the reference values for native speakers, it would have to be much more pronounced.

An interesting comparison can also be made with the reference values for native German speakers (Sendlmeier & Seebode, n.d.), see Table 10. All six vowel qualities analysed in this study occur in German as well, with the difference that short /u/ is transcribed as /ʊ/ in German, as it reflects a different qualitative character of the sound.

Table 10 F1 and F2 reference values for selected vowels in female native German speakers (Sendlmeier & Seebode, n.d.). The ERB values and F1/F2 difference were calculated by the author of this paper.

	F1 (Hz)	F2 (Hz)	F1 (ERB)	F2 (ERB)	F1 difference (% of ERB)	F2 difference (% of ERB)
a	836	1586	14.29	19.25	3.45	1.91
a:	896	1517	14.8	18.89		
ɪ	433	2095	9.87	21.54	26.21	6.95
i:	302	2533	7.82	23.15		
ʊ	442	1081	10	16.21	17.10	6.09
u:	345	956	8.54	15.28		

When comparing the German and Czech reference values, we see that the vowels /a aː/ generally have higher F1 and F2 formants in German. The vowel /ɪ/ is slightly higher in German, while /iː/ is slightly lower. The vowel /ʊ/ shows only minor differences, and a more pronounced distinction can be observed in the F2 of /uː/. These disparities naturally also affect the F1 and F2 differences between short and long vowels. An important question is therefore: Could the different nature of the target vowels in German have influenced their non-canonical pronunciation in Czech? Based on my data, this can nei-

ther be confirmed nor ruled out. It might be possible for the vowels /a a:/, whose F1 values indeed correspond more closely to the German reference values, but their F2 values, on the other hand, tend to be closer to the Czech reference values. As for the other four vowels, the differences do not seem to be caused by different vowel qualities in the two languages, but rather by the fact that the Czech short vowels /ɪ/ and /u/ are pronounced very similarly to their long counterparts. The only case where the experimental results approach the German reference values is in the F2 of /u:/. Speakers are thus more likely unaware of the differences in vowel quality of these sounds.

As for the absolute vowel duration and the duration ratio between long and short vowels, the results of this study (Table 6a and 6b) can be compared with the reference values presented in the Introduction (Table 1). Whether we take the study by Skarnitzl (2012) or that by Paillereau and Chládková (2019) as a reference, it is clear that the duration differences between short and long vowels are insufficient. Once again, however, it can be observed that the native German speakers are moving in the right direction – the largest duration difference is found for /a a:/, a slightly smaller difference for /u u:/, and only a very small difference for /ɪ i:/, which corresponds primarily to the reference values reported by Skarnitzl (2012). As with vowel quality, it can therefore be concluded that the principle according to which native German speakers in this study distinguish between short and long vowels in Czech words is heading in the right direction, but the vowels need to be differentiated more strongly in order to achieve greater proportional differences.

An interesting comparison can also be made in terms of absolute vowel duration, see Table 11.

Table 11 Duration of native Czech vowels (ms) in Skarnitzl (2012) and Paillereau & Chládková (2019).

	Skarnitzl	Paillereau & Chládková
a	63.1	75
a:	113.0	126
ɪ	53.5	61
i:	68.9	98
u	57.3	74
u:	91.4	119

German speakers in this study produced all vowels with a longer duration than native Czech speakers in the reference studies. This may be because non-native speakers tend to have a slower speech rate, as they require more time to plan their utterances; in the case of reading aloud, they may also read more slowly because they need to concentrate more. In the present study, the average articulation rate of the speakers ranged from 3.29 to 5 syllables per second, with the overall mean articulation rate being 4.20 syllables per second (calculated based on target words, not full sentences). Skarnitzl (2014: 99), who measured speech rate in semi-spontaneous dialogues, found a significant effect of gender in his data, with an average articulation rate of 6.48 syllables per second for female speakers. Different values are reported by Palková (1994: 317–318), who gives a mean

speech rate of 4.98 syllables per second for both genders combined (this value results from averaging five studies, each of which worked with a different type of data). Considering that Skarnitzl worked with articulation rate and Palková with speech rate, the average articulation rates of the speakers in the present study are indeed slightly below average, but the differences are not substantial enough to explain the fact that the vowel durations are sometimes more than twice as long as the reference values.

It is also important to note that the vowels /ɪ iː/ behaved differently in the first and second syllable. Regarding vowel quality, the differences are small, on the order of a few percentage points (the F1 difference is 2.22% lower in the first syllable compared to the second, while the F2 difference is 0.13% higher in the first syllable than in the second; see Tables 5a and 5b). However, for vowel length, the differences are substantial: in the first syllable, the long variant is 35% longer than the short one, whereas in the second syllable, it is only 2% longer (see Table 8). Since /ɪ iː/ occurred more frequently in the second syllable, the overall results show a smaller duration difference between the long and short vowel. The results in Table 8 also suggest that lengthening occurred in the second syllable, as the duration of short /ɪ/ here approximately corresponds to the duration of long /iː/ in both syllables. Since some authors argue that the stressed syllable is relatively longer than the unstressed one in German (see Introduction), I also focused on the absolute duration of vowels in the first syllable compared to the second. However, this hypothesis was not confirmed in my data, see particular Tables 8a–8d for summary results and Table 8 for the vowels /ɪ iː/ in the Results section.

5. Conclusion

This study presented an acoustic analysis of formants and durations of Czech vowels /a aː ɪ iː u uː/ in the speech of native German speakers. Eight native German speakers (all female) were recorded in a studio while reading a text containing sentences with target words. The target words formed groups of two to four words differing only in vowel quantity (e.g., the group of four words platu, platů, plátu, plátů). Formants F1 and F2 were measured, and the differences in formant values between long and short variants of the respective vowels were calculated. The duration of target vowels was also measured, and the extent to which the long vowel variant was longer than the short one was determined. The results were then compared with reference values for Czech and German speakers. The speakers in the present experiment produced correct patterns in terms of how they should distinguish between the long and short vowel variants (differences in quality or duration). However, the differences between long and short vowels were not sufficiently distinct (for example, the vowel qualities /ɪ/ and /iː/ merge together, although they should be two separate vowels, or the duration difference between long /aː/ and short /a/ is not distinct enough). For the vowels /ɪ iː/, which were the only ones found both in the first and second syllables, differences in vowel duration depending on their position in the word were also observed. In the second syllable, lengthening was present, and the duration of both vowels almost completely merged. At the same time, the absolute duration of all vowels was considerably longer than the reference values, to which the speakers' slower articulation rate may have contributed. Differences in formant values when comparing

the results and the reference data for Czech native speakers may or may not be caused or influenced by slightly different formant values for these vowels in the two languages. This study confirmed the assumption that native German speakers tend to struggle with the realisation of vowel length in Czech, and that this phenomenon requires increased attention when acquiring the language.

Acknowledgements

I would like to thank Jitka Veroňková for her valuable consultations during the implementation of the experiment and writing of this article. I also thank Radek Skarnitzl for providing the reference formant values for female native speakers of Czech and Tomáš Bořil for providing initial scripts for the acoustic analysis, which were subsequently adapted for the purposes of this study. This experiment used recordings from the corpus of non-native Czech speech, which was compiled by the Institute of Phonetics at the Faculty of Arts, Charles University, and collected as part of the Czech Science Foundation project GA ČR 18-18300S *Zvukové vlastnosti češtiny v komunikaci nerodilých a rodilých mluvčích*.

REFERENCES

Ashby, M., & Maidment, J. A. (2015). *Úvod do obecné fonetiky*. Karolinum.

Becker, T. (2012). *Einführung in die Phonetik und Phonologie des Deutschen*. WBG Wissenschaftliche Buchgesellschaft.

Boersma, P., & Weenink, D. (2022). *Praat: doing phonetics by computer* (Version 6.2.10). Retrieved 26. 3. 2022 from http://www.praat.org.

Bořil, T. (2015). *Výpočet formantů*. Retrieved 26. 5. 2025 from https://fu.ff.cuni.cz/AKU/script_01.html#1.

Chabrová, A., & Veroňková, J. (2022). Percepce kvantity vokálů v českých dvojslabičných slovech v řeči rodilých mluvčích němčiny. *Nová čeština doma a ve světě*, *2*, 25–37.

Glasberg, B. R., & Moore, B. C. J. (1990). Derivation of auditory filter shapes from notched-noise data. *Hearing research*, *47*(1), 103–138.

Jessen, M. (1993). Stress conditions on vowel quality and quantity in German. *Working Papers of the Cornell Phonetics Laboratory*, *8*, 1–27.

Kleiner, S., Knöbl, R. (2015). *Duden, das Aussprachewörterbuch* (7th ed.). Dudenverlag.

Křen, M., Cvrček, V., Henyš, J., Hnátková, M., Jelínek, T., Kocek, J., Kováříková, D., Křivan, J., Milička, J., Petkevič, V., Procházka, P., Skoumalová, H., Šindlerová, J., & Škrabal, M. (2020). *SYN2020: reprezentativní korpus psané češtiny*. Ústav Českého národního korpusu FF UK. Retrieved from http://www.korpus.cz.

Machač, P., & Skarnitzl, R. (2009). *Principles of phonetic segmentation*. Epocha.

Paillereau, N., & Chládková, K. (2019). Spectral and temporal characteristics of Czech vowels in spontaneous speech. *Acta Universitatis Carolinae. Philologica*, *2019*(2), 77–95.

Palková, Z. (1994). *Fonetika a fonologie češtiny: s obecným úvodem do problematiky oboru*. Karolinum.

Palková, Z., Bořil, T., & Veroňková, J. (2020). Difficulties in adjacent vowel length of L1 Russian speakers in Czech. In Botinis, A. (ed.), *Proceedings of 11th International Conference of Experimental Linguistics* (pp. 149–152). ExLing.

Podlipský, V. J., Skarnitzl, R., & Volín, J. (2009). High front vowels in Czech: a contrast in quantity or quality? In *Proceedings Interspeech 2009* (pp. 132–135).

Pompino-Marschall, B. (2003). *Einführung in die Phonetik* (2nd ed.). Walter de Gruyter.

Rausch, R., & Rausch, I. (1998). *Deutsche Phonetik für Ausländer* (5th ed.). Langenscheidt.
Sendelmeier, W. F., & Seebode, J. (n.d.). *Formantkarten des deutschen Vokalsystems*. Retrieved 25. 3. 2022 from https://www.tu.berlin/kw/forschung/projekte/formantkarten.
Skarnitzl, R. (2012). Dvojí i v české výslovnosti. *Naše řeč*, *95*(3), 141–153.
Skarnitzl, R. (2014). *Fonetická identifikace mluvčího*. Filozofická fakulta Univerzity Karlovy.
Skarnitzl, R. (2018). Fonetická realizace slovního přízvuku u delších slov v češtině. *Slovo a slovesnost*, 79(3), 199–216.
Skarnitzl, R., & Volín, J. (2012). Referenční hodnoty vokalických formantů pro mladé dospělé mluvčí standardní češtiny. *Akustické listy*, *18*(1), 7–11.
Vachek, J. (1968). *Dynamika fonologického systému současné spisovné češtiny*. Academia, nakladatelství Československé akademie věd.
Veroňková, J., & Bořil, T. (2020a). Czech vowel quantity in Polish speakers as perceived by Moravian-Silesian listeners. In *Speech Research conference* (pp. 89–91). Hungarian Research Institute for Linguistics.
Veroňková, J., & Bořil, T. (2020b). Phonological length of L2 Czech speakers' vowels in ambiguous contexts as perceived by L1 listeners. In Karpov, A., & Potapova, R. (eds.), *Speech and Computer. 22nd International Conference, SPECOM 2020*. *Lecture Notes in Computer Science*, vol. 12335 (pp. 624–635). Springer.
Veroňková, J., Bořil, T., Palková, Z., & Poukarová, P. (2020). Délka českých samohlásek u polských mluvčích v taktech s různou strukturou kvantity. In Bogoczová, I. (ed.), *Area Slavica 3 (Jazyk na hranici – hranice v jazyku)* (pp. 51–61). Ostravská Univerzita, Filozofická fakulta.
Wiese, R. (2011). *Phonetik und Phonologie*. Wilhelm Fink.

RESUMÉ

Předkládaná studie je součástí širšího experimentu, který se zaměřuje na percepční a akustickou analýzu vokálů /a a: ɪ i: u u:/ v českých dvojslabičných slovech u rodilých mluvčích němčiny, učících se češtinu jako cizí jazyk. Studie prezentuje výsledky akustické analýzy a poskytuje stručný přehled o chování vokálů v češtině a němčině. Materiál tvoří dvojice, trojice a čtveřice slov, které se navzájem liší pouze kvantitou vokálu (např. čtveřice platu, platů, plátu, plátů). Nahrávky pochází od 8 žen různých jazykových úrovní (A2–B1 podle SERR). V cílových slovech byly analyzovány vokalické formanty F1 a F2 a vypočteny rozdíly ve formantových hodnotách mezi krátkými a dlouhými vokály. Dále bylo změřeno trvání vokálů a byly vypočteno, o kolik jsou dlouhé vokály delší než krátké. Výsledky studie jsou následně porovnány s referenčními hodnotami pro ženské rodilé mluvčí češtiny a ukazují, že mluvčí v předkládané studii nerozlišují krátké a dlouhé vokály dostatečně silně, a to jak kvalitativně, tak kvantitativně, obecně ale volí správné strategie. V oblasti trvání jsou největší rozdíly mezi trváním krátkého a dlouhého vokálu u dvojice /a: a/, následuje /u: u/ a nejmenší rozdíl je u /ɪ i:/. Také z hlediska formantů dochází správně k odlišování rozdílné kvality u dlouhého a krátkého /ɪ i:/, v menší míře také /u u:/. Například u vokálů /ɪ i:/ by ale podle referenční studie měly být hodnoty F1 u dlouhého vokálu o 28 % nižší než u krátkého, zatímco u mluvčích v této studii byl rozdíl pouze 5 %. Hodnoty vokalických formantů u mluvčích v této studií mohou, ale nemusí být ovlivněny mírně odlišnými hodnotami formantů v němčině ve srovnání s češtinou. Výsledky také ukazují, že absolutní trvání vokálů je delší, než uvádí referenční hodnoty pro české mluvčí, což by mohlo být způsobeno nižším artikulačním tempem. Studie potvrdila předpoklad, že vokalická kvantita je jev, kterému by rodilí mluvčí němčiny při osvojování češtiny měli věnovat zvýšenou pozornost.

Anna Chabrová
Institute of Phonetics
Charles University, Faculty of Arts
Prague, Czech Republic
anna.chabrova@ff.cuni.cz