

Referenční hodnoty vokálních formantů pro mladé dospělé mluvčí standardní češtiny

Radek Skarnitzl a Jan Volín

Fonetický ústav – FF UK, nám. J. Palacha 2, 116 38 Praha 1
e-mail: [radek.skarnitzl; jan.volín]@ff.cuni.cz

Vowel formants, the resonances of the vocal tract during the production of vowels, are the most frequent parameters used for the description of vocalic systems of languages. Czech vowels and their formants were systematically last investigated several decades ago, with different researchers investigating vowels in different kinds of speech material. The main aim of this paper is therefore to provide reference values of vowel formants in contemporary Czech. Formant measurements are based on the analyses of eight instances of each short vowel and three instances of each long vowel in 75 speakers of Czech. Apart from providing the reference values, the paper suggests an interesting development in the Czech vocalic system: while the difference in quality between the short /ɪ/ and the long /i:/ has been previously documented, our results indicate a similar split in the close back vowels, /u/ and /u:/.

1. Úvod

Když Hermann von Helmholtz v roce 1863 použil speciálně sestavený rezonátor k analýze a imitaci samohláskových zvuků, nejspíš ani sám netušil, že o sto padesát let později budou rezonanční frekvence vokálního traktu, tzv. formanty, stále nejpoužívanějšími parametry pro popis vokálních systémů v jazycích světa.

Popis samohlásek pomocí formantů je efektivní a zároveň elegantní z několika důvodů. Zaprvé je schopen rozlišit samohlásky daného jazyka z hlediska fonologické kvality pomocí prvních dvou rezonancí, F1 a F2, a tedy umožňuje daný vokální systém zobrazit ve dvou rozměrech.¹ Výhodné je i to, že oba rozměry jsou zároveň vyjádřeny v jediné veličině, frekvenci. Oproti tomu pro popis systémů souhláskových potřebujeme deskriptivních parametrů více a tyto se navíc mohou pro jednotlivé konsonantické skupiny lišit. Důležité rovněž je, že výsledné zobrazení systému samohlásek intuitivním způsobem koresponduje s artikulačním nastavením jazyka: z vokálních čtyřúhelníků a uvedených popisků na obr. 1 například vyplývá, že [i:] je vysoký přední vokál, [a] je vokál nízký střední a [o] středový zadní.

Pozice jednotlivých samohlásek se liší nejen v různých jazycích, ale dokonce i v dialektech jednoho jazyka. Formantovými hodnotami českých samohlásek se zabýval především Bohuslav Hála, který měřil vokální formanty ve slovech, větách i v souvislé řeči u čtyř mluvčích [1]. Boro-
vičková a Maláč [3] zkoumali výslovnost vokálů v tzv. logatomech (pseudoslovech, která nemají lexikální význam, avšak jejich struktura odpovídá stavbě slov v daném jazyce). Český vokální systém je v těchto studiích prezentován jako relativně jednoduchý a symetrický, v němž se

fonologicky krátké a dlouhé vokály v dané dvojici z hlediska kvality (tzn. hodnot formantů) neliší. Výjimku představují vysoké přední vokály, /ɪ/ a /i:/, jež se liší právě i v kvalitě; to odráží i odlišný způsob jejich transkripce.

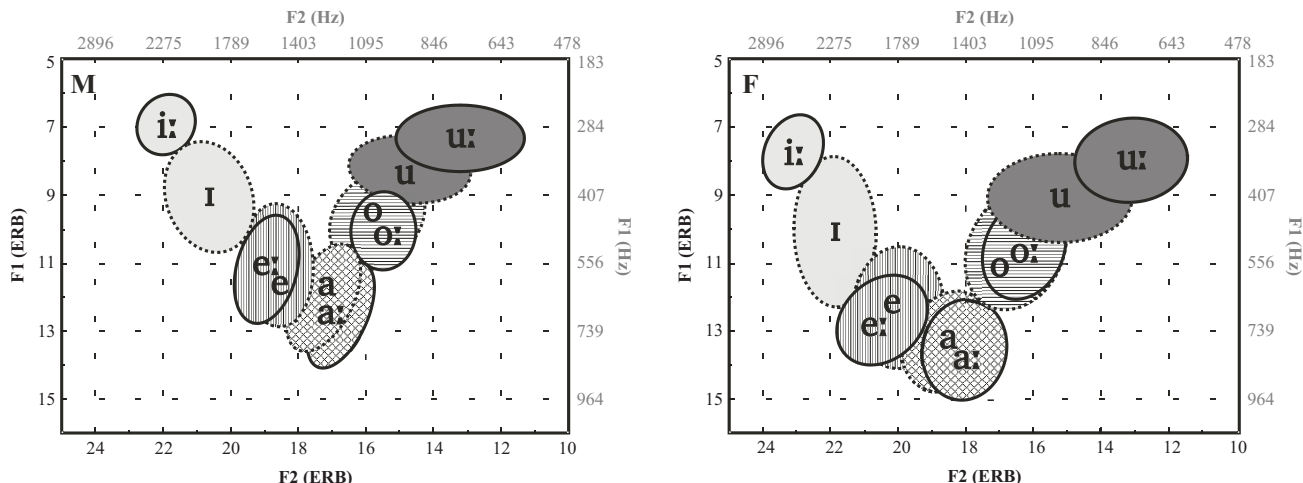
V poslední době české vokály zkoumali Volín a Studenovský [4], jejich cílem však nebylo poskytnout referenční formantové hodnoty: zjišťovali, jaká metoda nejefektivněji normalizuje hodnoty formantů tak, aby bylo možné porovnávat například mluvčí ženského a mužského pohlaví, ale aby zároveň nedošlo ke ztrátě lingvisticky relevantních detailů. Autoři zkoumali pouze krátké vokály, protože ty jsou v běžném textu průměrně tři- až čtyřikrát častější než samohlásky dlouhé [5]. Skarnitzl [6] se zabýval vztahem mezi trváním a formantovými hodnotami u dvojice /ɪ-/i:/.

Protože jsou vokální formanty důležitým údajem fonetického popisu, je samozřejmě vhodné, aby pro daný jazyk byly dostupné jejich referenční hodnoty. Pro češtinu doposud aktuální a zároveň reprezentativní data chyběla. Cílem našeho příspěvku je tuto mezeru zaplnit.

2. Metoda

Analýza formantů byla založena na nahrávkách 75 studentů lingvistických oborů na Filozofické fakultě (48 žen, 27 mužů); z části vycházíme z materiálu Volína a Studenského [4]. Jednalo se o osoby ve věku 20–30 let, bez řečových a sluchových vad. Studenti byli instruováni, aby po krátké přípravě neutrálně a přirozeně přečetli přibližně minutový souvislý text. Nahrávky byly pořízeny ve zvukově upraveném studiu Fonetického ústavu FF UK v Praze. Byl použit kondenzátorový mikrofon AKG C4500 B-BC, akustický signál byl nahráván přímo do zvukové karty SB Audigy 4 při vzorkovací frekvenci 32 kHz a 16bitovém rozlišení.

¹Stranou nyní necháváme vokály realizované s různými typy fonace, např. dyšné vokály v gudžarátštině či vokály s třepenou fonací v mexickém jazyce Jalapa Mazatec; v těchto jazycích mají zmíněné modifikace distinktivní platnost [2: 315nn.].



Obrázek 1: Formantová pole českých monoftongů, zvláště pro muže (M) a ženy (F). Elipsy naznačují jednu směrodatnou odchylku od průměru. Fonologicky dlouhé vokály jsou ohraničeny plnou čarou, krátké tečkovanou

Z textů bylo vybráno osm položek každého krátkého vokálu a tři položky každého dlouhého vokálu. Celkově jsme tedy analyzovali 4 125 vokálů (3 000 krátkých a 1 125 dlouhých). Vokály byly vybírány z různých segmentálních, prozodických i gramatických kontextů tak, aby byly vlivy těchto kontextů na hodnoty formantů vyvážené.

Frekvence prvních dvou formantů (F1 a F2) byly měřeny v programu Praat [7], nástroji určeném pro fonetické analýzy, pomocí Burgova algoritmu založeného na LPC. Formanty byly extrahovány při standardním nastavení pro každé z pohlaví: pro mužské mluvčí byly extrahovány tři formanty ve frekvenčním pásmu 0–3 000 Hz, pro ženské mluvčí v pásmu 0–3 300 Hz. Za frekvenci formantu byl považován aritmetický průměr sedmi ekvidistantních měření z prostřední třetiny vokálu. To ovšem neznamená měření rozdílným analyzačním oknem: u delších vokálů byly jednotlivé body měření poněkud dále od sebe. Protože neexistuje zcela spolehlivý extraktor formantů, naměřené hodnoty jsme následně zkontrolovali a případně opravili přímým odečtem ze spektrogramu.

Opravené hodnoty formantů jsme nakonec také převedli z hertzů do psychoakustické jednotky ERB (Equivalent Rectangular Bandwidth), která lépe vyjadřuje, jak výškové rozdíly vnímá lidské ucho [8]. Mezi objektivním fyzikálním měřením a jeho percepčním dopadem totiž existují zajímavé, nelineární vztahy.

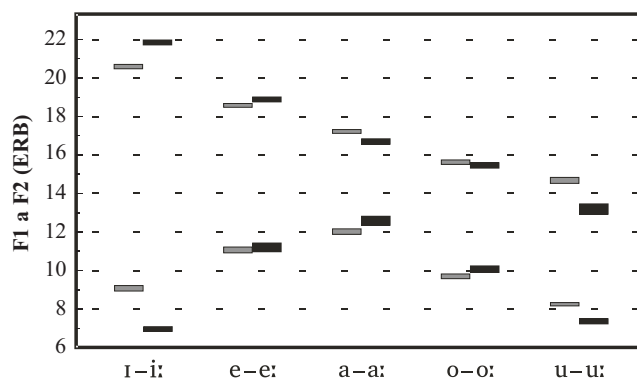
3. Výsledky a diskuse

Umístění všech českých monoftongů na erbové stupnici – kterou zde považujeme za primární – představuje obrázek 1, zvláště pro mužské a ženské mluvčí. Stupnice v hertzech jsou pro orientaci uvedeny v šedé barvě. Elipsy odpovídají jedné směrodatné odchylce od průměru, což znamená, že zahrnují přibližně 68 procent všech položek daného vokálu. U ženských i mužských mluvčích je patrné,

že zatímco u e-ových, a-ových a o-ových vokálů se formantová pole do značné míry překrývají, u u-ových je překryv jen částečný a u i-ových vokálů v podstatě žádný.

Jak naznačuje obrázek 1, formanty ženských mluvčích se od mužských zásadně neliší; v souladu s všeobecně platnými tendencemi jsou přibližně o 15 až 20 procent vyšší než formanty mužských mluvčích (elipsy jsou tedy na obrázku posunuty dolů a doleva). Protože se většinou jako referenční uvádějí hodnoty pro mužské mluvčí, představují následující zobrazení již pouze hodnoty pro mužské mluvčí.

Jak jsme již uvedli v úvodu, podle tradičních popisů se s výjimkou i-ových vokálů kvalita fonologicky krátkého vokálu od dlouhého v dané dvojici neliší. Zajímalo nás proto, zda tuto poučku potvrzuje i náš řečový vzorek mladších uživatelů češtiny. Obrázek 2 ukazuje F1 a F2 všech vokálů (krátké vokály jsou zobrazeny šedě, dlouhé černě) v podobě schematického spektrogramu.



Obrázek 2: Schematické spektrogramy českých monoftongů (pouze pro mužské hodnoty). Šedé pruhy odpovídají formantům fonologicky krátkých vokálů, černé pruhy formantům vokálů dlouhých. Tloušťka pruhů vyjadřuje 95% interval spolehlivosti

	F1 (Hz)	F2 (Hz)	F1 (ERB)	F2 (ERB)	rozdíl F1 %	rozdíl F2 %
I	414,9 (76,6)	1943,2 (202,0)	9,06 (1,08)	20,64 (0,87)	30,55	5,80
i:	281,7 (33,2)	2255,2 (151,6)	6,94 (0,58)	21,91 (0,56)		
e	566,3 (98,5)	1518,8 (127,7)	11,05 (1,18)	18,62 (0,70)	1,16	1,69
e:	575,7 (86,0)	1578,0 (122,9)	11,18 (1,03)	18,94 (0,63)		
a	648,5 (94,8)	1286,0 (113,5)	12,01 (1,03)	17,26 (0,72)	4,23	3,05
a:	698,6 (96,5)	1206,3 (97,2)	12,54 (1,01)	16,75 (0,65)		
o	457,7 (69,3)	1054,8 (129,8)	9,68 (0,93)	15,66 (0,92)	3,68	1,16
o:	483,7 (54,9)	1027,9 (82,0)	10,05 (0,75)	15,48 (0,62)		
u	359,2 (43,1)	936,6 (145,1)	8,25 (0,66)	14,71 (1,19)	12,55	11,36
u:	304,1 (37,7)	768,9 (132,4)	7,33 (0,64)	13,21 (1,24)		

Tabulka 1: Hodnoty formantů českých monoftongů (mužské hodnoty) a procentuální rozdíl mezi hodnotami pro krátký a dlouhý vokál jednotlivých párů. V závorkách pod hodnotami formantů jsou uvedeny směrodatné odchylky

Ilustrace samozřejmě naznačuje stejné tendence jako obr. 1; výhodou je však zobrazení obou formantů na jediné ose. Ačkoli u všech dvojic můžeme alespoň v případě jednoho formantu pozorovat odlišné hodnoty, je otázkou, zda je lidské ucho schopné takové rozdíly vůbec vnímat. Za nejmenší vnímatelný rozdíl se u formantů přirozených (tj. nikoli syntetických) vokálů považuje přibližně pět procent [9]. Tabulka 1 shrnuje průměrné hodnoty a v posledních dvou sloupcích uvádí procentuální rozdíl mezi hodnotou formantu pro krátký a dlouhý vokál daného páru, vypočítaný z psychoakustických hodnot (ERB).

Hodnoty procentuálního rozdílu naznačují, že kvalitu krátkého a dlouhého vokálu v e-ové, a-ové a o-ové dvojici skutečně můžeme považovat za stejnou. Dvojice i-ových vokálů se liší především ve vokalické výšce, tzn. v prvním formantu; rozdíl mezi /i/ a /i:/ v F2 se pohybuje mírně nad hranicí rozlišitelnosti. Zajímavá situace se ukazuje u vokálů u-ových, u nichž se doposud odlišná kvalita nepředpokládala, ale rozdíl mezi /u/ a /u:/ výrazně přesahuje nejmenší vnímatelný rozdíl jak v prvním, tak i druhém formantu.

Je otázkou, zda kvalitativní odlišení u-ových vokálů, jež naznačují naše data, můžeme považovat za probíhající

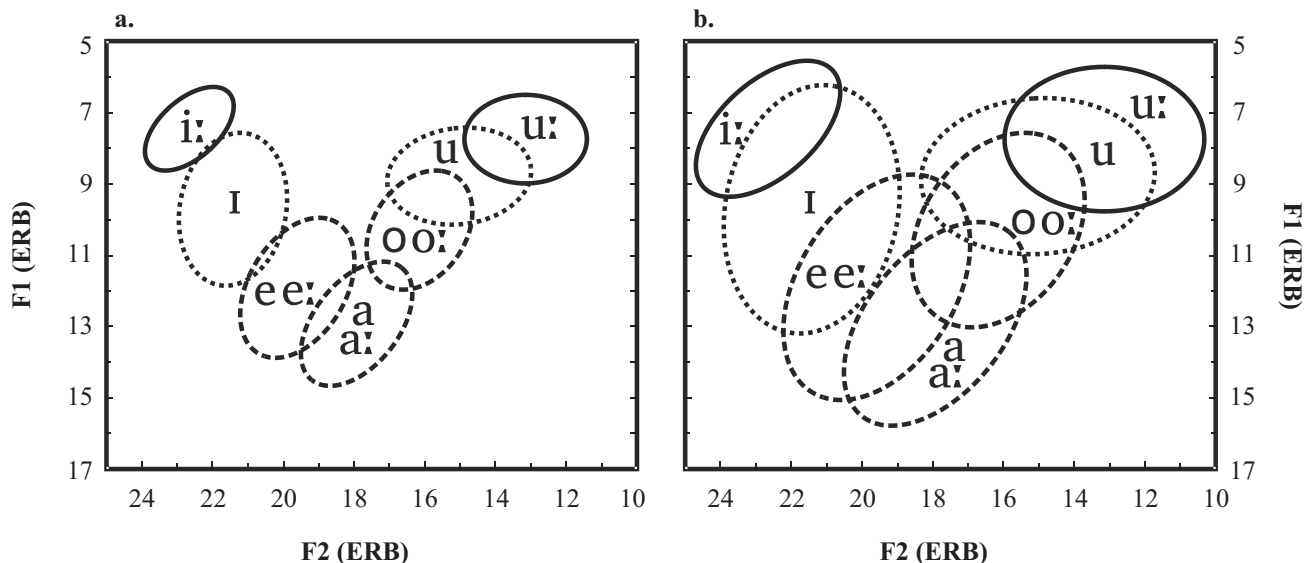
změnu v českém vokalickém systému. Pokud by tomu tak bylo, jeden z možných scénářů budoucího vývoje by mohl vést k plnému kvalitativnímu odlišení i-ových a u-ových, tedy vysokých vokálů. V této souvislosti je zajímavé podívat se na rozdíly mezi trváním krátkého a dlouhého vokálu v jednotlivých dvojicích: porovnání zobrazuje tabulka 2 (uvádíme trvání vokálů ze čteného zpravodajství Českého rozhlasu podle studie Skarnitzla [6]). Vidíme, že zatímco u e-ových, a-ových a o-ových vokálů je fonologicky dlouhý vokál vždy o 70 až 80 procent delší než vokál krátký, u vysokých vokálů je situace odlišná. Fonologicky dlouhé /i:/ je přibližně jen o 30 procent delší než krátké /i/. Je pravděpodobné, že značný spektrální rozdíl mezi /i/ a /i:/ umožňuje ekonomicky snížit rozdíl v oblasti trvání. Dlouhé /u:/ je průměrně o 60 procent delší než krátké /u/; u-ové vokály tak z hlediska trvání – podobně jako v oblasti vokalické kvality – leží mezi jednoznačně rozlišenými vokály i-ovými a mezi vokály nevysokými.

Můžeme se tedy domnívat, že temporální údaje souvisí s rozdíly spektrálními. Pokud by tento trend pokračoval a docházelo by k dalšímu vzájemnému odlišování u-ových vokálů, tedy k centralizaci krátkého /u/ a naopak ke zkracování fonologicky dlouhého /u:/, vedlo by to – možná paradoxně – k větší symetričnosti českého vokalického systému: vysoké vokály by byly odlišeny ve spektrální i temporální rovině, zatímco u vokálů nevysokých by rozlišující vlastností zůstalo pouze trvání. Pokud by se v budoucnu nastíněný scénář uskutečnil, bylo by zřejmě nutné u-ové vokály rozlišit i v oblasti transkripce: prepisovali bychom dlouhé /u:/ a krátké /u/, podobně jako například v angličtině či němčině.

vokál	trvání (ms)	poměr V: / V
I	53,5	1,29
i:	68,9	
e	53,5	1,72
e:	91,8	
a	63,1	1,79
a:	113,0	
o	51,7	1,73
o:	89,3	
u	57,3	1,60
u:	91,4	

Tabulka 2: Průměrné trvání a poměr trvání dlouhých a krátkých vokálů ve výslovnosti moderátorů Českého rozhlasu (podle: Skarnitzl, 2012)

Za zajímavý výsledek považujeme vztah mezi krátkým /a/ a dlouhým /a:/. Možná jste v posledních letech u mladých Čechů zaznamenali posun především dlouhého /a:/ dopředu, výslovnost, již bychom přepsali jako [a:]. Takové /a:/ se podobná anglickému vokálu /æ/, například ve slově *cat*, a můžeme se s ním rovněž setkat jako s regionální



Obrázek 3: Formantová pole českých vokálů se zahrnutím cca 68 % položek (a) a cca 95 % položek (b)

realizací na Plzeňsku. Je zajímavé, že toto neformální pozorování naše data nepotvrzují, ačkoli srovnání s Hálovými [1] hodnotami posun dopředu u krátkého /a/ i dlouhého /a:/ naznačuje.

Poslední aspekt, který zde zmíníme, je překryv jednotlivých samohlásek ve vokalickém prostoru. Podle dřívějších předpokladů se – samozřejmě i díky relativně nízkému počtu vokálů v češtině – formantová pole vokalických kvalit nepřekrývají. Obrázek 3 porovnává formantová pole při zahrnutí jedné směrodatné odchylky od průměru (cca 68 % všech položek) a dvou směrodatných odchylek (cca 95 % položek). Je patrné, že při druhém zobrazení, v němž jsou vypuštěny pouze extrémní hodnoty, se v našem vzorku formantová pole již překrývají do značné míry.

4. Závěr

Hlavním cílem této studie bylo aktualizovat hodnoty formantů českých samohlásek. Vokalické formanty byly analyzovány v nahrávkách souvislých čtených textů mladých dospělých mluvčích. Ačkoli jsme se nesnažili vzorek stu-

dentů vyvážit z hlediska dialektologického rozvrstvení češtiny, nahrávky zahrnují subjekty ze všech hlavních dialektologických oblastí České republiky.

Studie potvrdila některé tendence, jež naznačily předchozí sondy. Jedná se zejména o rozdílnou kvalitu českých i-ových vokálů [6] a o asymetrii v oblasti středových vokálů, u nichž je zřejmá zavřenější realizace zadních o-ových a otevřenější realizace předních e-ových vokálů [4]. Naše studie však především ukázala, že jednotlivé dvojice nevysokých samohlásek (e-ové, a-ové a o-ové) můžeme z hlediska vokalické kvality považovat za homogenní, zatímco krátké a dlouhé vysoké samohlásky (i-ové a u-ové) se ve formantech liší. To zohledňuje i tabulka 3, která uvádí cílové, referenční hodnoty českých formantů v zaokrouhlené podobě. Tyto hodnoty mohou být informativní i pro fonetické účely, kde je znalost modálních deskriptorů žádoucí.

Poděkování

Tento výzkum vznikl za podpory grantu GAČR 406/12/0298.

Reference

- [1] Hála, B.: *Akustická podstata samohlásek*, Česká akademie věd a umění, Praha, 1941.
- [2] Ladefoged, P., Maddieson, I.: *The Sounds of the World's Languages*, Blackwell, Oxford, 1996.
- [3] Borovičková, B., Maláč, V.: *The spectral analysis of Czech sound combinations*, Academia, Praha, 1967.
- [4] Volín, J., Studenovský, D.: Normalization of Czech vowels from continuous read texts, *Proc. of the 16th ICPHS*, p. 185–190, Saarbrücken, 2007.

	F1	F2
I	410	1940
i	280	2260
e e:	570	1550
a a:	670	1250
o o:	470	1040
u	360	940
u:	300	770

Tabulka 3: Zaokrouhlené hodnoty kmitočtů (Hz) českých monoftongů pro mluvčí mužského pohlaví

- [5] Těšitelová, M. a kol.: *O češtině v číslech*, Academia, Praha, 1987.
- [6] Skarnitzl, R.: Dvojí i v české výslovnosti, *Naše řeč*, 2012 (v tisku).
- [7] Boersma, P., Weenink, D.: *Praat: doing phonetics by computer* (Version 5.2). Staženo 1. listopadu 2010, <http://www.praat.org>
- [8] Moore, B. C. J., Glasberg, B. R.: Suggested formulae for calculating auditory-filter bandwidths and excitation patterns, *Journal of the Acoustical Society of America*, 74, p. 750–753, 1983.
- [9] Pols, L. C. W.: Flexible, robust, and efficient human speech processing versus present-day speech technology, *Proc. of the 14th ICPHS*, p. 9–16, San Francisco, 1999.

AKUSTICKÉ LISTY

České akustické společnosti
www.czakustika.cz

ročník 18, číslo 1

duben 2012

Obsah

Usnesení Valné hromady České akustické společnosti	3
prof. Ing. Felix KOLMER, DrSc. – devadesátiletý <i>Jaroslav Kyncl</i>	5
Referenční hodnoty vokálních formantů pro mladé dospělé mluvčí standardní češtiny Reference Values of Vowel Formants for Young Adult Speakers of Standard Czech <i>Radek Skarnitzl a Jan Volín</i>	7
Kirnbergerova temperatura v „Die Kunst des reinen Satzes in der Musik“ (1771) The Kirnberger Temperament published in „Die Kunst des reinen Satzes in der Musik“ (1771) <i>Martin Celhoffer</i>	12

ČESKÁ AKUSTICKÁ SPOLEČNOST