

Aleš Bičan

KVANTITATIVNÍ ANALÝZA SLABIKY V ČESKÉM LEXIKONU

ABSTRACT

*The paper provides a quantitative analysis of the syllable in contemporary Czech in a corpus of 146,703 syllables contained in Czech words recorded in *Slovník spisovné češtiny*. After discussing the general structure of the Czech syllable and constraints on its structure (in terms of the allowed number of non-vocalic phonemes within a syllable), it examines the distribution of syllabic types according to the type of syllabic nucleus and according to different parts of speech. It shows that the most common are open syllables and syllables beginning with a single consonant. In total, the most favored syllabic type is CV. These tendencies hold for all syllabic nuclei except for the diphthongal vowels, but here we assume the influence of the chosen data (no inflected forms of words are included) rather than that it would be a special characteristic of this vowel class. All parts of speech show the same preferences, and there are only minor differences between them. However, it has been found out that pronouns and uninflectable words have a less complex syllabic structure than the other parts of speech. A partial comparison with other samples of the Czech vocabulary is offered, too.*

KEYWORDS

Phonotactics; syllable; phonological corpus; lexical corpus; syllabic consonants; parts of speech.

1. Úvod

V tomto článku nabízíme kvantitativní analýzu slabičné struktury slov v moderní češtině; zdrojem dat je pro nás fonologický lexikální korpus. Zajímat nás bude struktura slabiky jako takové a procentuální zastoupení jednotlivých slabičných typů, a to jak v celém vzorku zvolené slovní zásoby, tak u jednotlivých slovních druhů.

Slabičná struktura češtiny byla již dříve podrobena kvantitativním analýzám (KUČERA – MONROE 1968, LUDVÍKOVÁ 1972a, 1972b, 1976, 1978, souhrnně pak Těšitelová et al. 1985), avšak zmíněné analýzy byly omezeny technickými možnostmi

své doby.¹ Kučerova databáze vycházela z psaného textu a obsahovala 42 217 slabik. Ludvíková také založila svou analýzu na textech několika stylů, především na textech mluvených. Její databáze čítala pouze 5 000 slabik, avšak v článku LUDVÍKOVÁ (1972a) zohlednila i texty psané a získala vzorek 12 475 slabik. Následně mohly být pro texty stanoveny frekvence rozličných slabičných typů, výskyt slabik v různých částech slov, výskyt slov podle počtu slabik a další charakteristiky. Doposud se však nikdo nevěnoval popisu slabičné struktury v českém lexikonu a u konkrétních slovních druhů.

Zdrojem našich dat je Fonologický lexikální korpus češtiny.² Tento korpus je anotovanou fonologickou databází slovní zásoby češtiny skládající se z lexikálních položek zaznamenaných v nejdůležitějších publikovaných slovnících češtiny (především PSJČ, SSJČ, SSČ, ale též SN₁, 2 či VSČ; viz literatura). Jelikož celý lexikální korpus obsahuje mnoho archaických, dialektických a okazionálních slov (zaznamenaných především v PSJČ), rozhodli jsme se popis omezit pouze na slova vyskytující se ve *Slovníku spisovné češtiny* (SSČ). SSČ obsahuje 45 978 různých fonologických slov (viz dále), která se skládají ze 146 703 slabik. Mluvíme-li tedy o češtině, máme na mysli především češtinu spisovnou. V některých případech ovšem naše pozorování ověřujeme na datech z celého lexikálního korpusu. V něm je obsaženo 257 855 fonologických slov skládajících se z 950 949 slabik. K otázce, zda se z hlediska slabičné struktury liší slovní zásoba SSČ od zbytku lexikálního korpusu, se vrátíme v závěru článku.

V našem lexikálním korpusu je každé položce přiřazena fonologická interpretace její předpokládané ortoepické výslovnosti (podle pravidel formulovaných v HÁLA 1968). Fonologická interpretace se řídí principy funkční fonologie, jejíž základy rozpracovala pražská škola (viz především TRUBECKOJ 1939), a již následně rozvinul André Martinet (např. MARTINET 2011) a Jan Mulder (MULDER 1989; srov. i BIČAN 2013).

Fonologická transkripce zohledňuje jak segmentální, tak suprasegmentální strukturu slov. Fonologická forma každé lexikální jednotky je přepsána jako posloupnost fonémů, která je rozdělena do slabik a fonologických slov. Slabikou zde rozumíme soběstačnou (autonomní) fonotaktickou jednotku, tedy jednotku, na které lze vyčerpávajícím způsobem popsat distribuci fonémů v daném jazyce (MULDER *op. cit.*). V tomto pojetí je slabika čistě fonologickou entitou, nikoliv entitou fonetickou.

Fonologické slovo je jednotka definovaná přízvukem (není proto větší než přízvukový takt) a/nebo výskytem hraničních signálů jako např. rázu (BIČAN 2014a). Ráz neboli glotální okluzíva se v češtině vyskytuje jen před vokály, a to 1) na za-

1 Některé informace o frekvenci slabičných typů přináší i CHURAŇOVÁ (2014), její popis je však fonetický, nikoliv fonologický, což znamená, že mezi konsonanty počítá i ráz.

2 Viz *Český fonologický lexikální korpus* [online]. [cit. 1.6.2015]. Dostupné z: <http://www.ujc.cas.cz/phword>.

čátku gramatických slov (např. *k [ʔ]autu*), 2) mezi prefixem a slovním základem (např. *na[ʔ]učit*), nebo 3) na hranici kompozit (např. *indo[ʔ]evropský*). Jako v mnoha jiných jazycích (DIXON – AIKHENVALD 2003) nelze tudíž pro češtinu stanovit totožnost mezi fonologickými slovy definovanými fonologicky a gramatickými slovy definovanými gramaticky (morfológicky či syntakticky, tedy jinak než fonologicky).³ Naše analýza vychází z fonologických slov. Budeme-li dále bez jakéhokoliv přívlastku o slovech mluvit, míníme tím slova fonologická.

Při dělení slov na slabiky („slabikování“) jsme se řídili pravidly navrženými KUČEROU a MONROEM (1968, 52); podobná pravidla použila i Ludvíková (viz např. TĚŠITELOVÁ 1985, 22–23). Oproti nim jsme však nebrali v úvahu hranici morfémů, jelikož dělení slov na slabiky je fonologická operace, kdežto hranice morfémů jsou stanoveny morfológickou analýzou. Vycházeli jsme tedy z předpokladu, že hranici slabik lze stanovit nezávisle na hranicích morfémů uvnitř slov. Ostatně při vysokém počtu slov (více než 250 000) nebylo ani v našich silách provést jejich morfológickou analýzu. Slova jsme slabikovali automaticky podle předem stanovených pravidel slabikování. Znamená to, že každá sekvence konsonantů mezi dvěma vokály je vždy rozdělena stejným způsobem. Přístup Kučery a Ludvíkové umožňuje několikrát dělení, a to podle toho, zda a kde se objevuje hranice morfémů.

Při stanovení pravidel slabikování jsme v prvé řadě volili jen takové dělení, jehož výsledkem je slabika doložená jinde (tj. na začátku či na konci slov, kde je slabičná hranice jasná), popř. slabika, jejíž struktura není v rozporu se strukturou doložených slabik (obecně k tomuto běžně uznávanému pravidlu viz KURYŁOWICZ 1948 a BELL 1976). Pokud bylo možné více takových dělení, tak jsme se podobně jako Kučera a Monroe řídili následující hierarchií dělení, jež odráží frekvenci jednotlivých slabičných struktur (V = vokál, C = konsonant, tečka = slabičná hranice):

(1) *Hierarchie pravidel slabičného dělení*

VCV	→ V.CV > VC.V
VCCV	→ VC.CV > V.CCV > VCC.V
VCCCV	→ VC.CCV > V.CCCV > VCC.CV > VCCC.V
VCCCCV	→ VCC.CCV > VC.CCCV > V.CCCCV > VCCC.CV
VCCCCCV	→ VCC.CCCV > VC.CCCCV > VCCC.CCV
VCCCCCVV	→ VCC.CCCV > VCCC.CCCV

Znamená to, že např. tři konsonanty mezi dvěma vokály jsou rozděleny tak, že jeden konsonant náleží levé slabice, dva konsonanty pravé slabice. Pokud je výsledkem nepřipustná slabika, volí se dělení, které je v hierarchii na druhém místě.

³ Gramatická slova tedy nejsou jen neplnovýznamová slova jako předložky, spojky či částice, ale obecně takové entity, které jsou vymezeny morfológickými nebo syntaktickými kritérii. Pro zjednodušení lze říct, že v češtině jsou gramatická slova totéž co slova ortografická.

Zde je třeba upozornit na dvě důležitá specifika naší analýzy. Za prvé, afrikáty ve slovech jako *cesta*, *čaj*, *dzinkat*, *džez* interpretujeme jako kombinace dvou konsonantů, tj. jako kombinace alveolárních okluzív + /s/, /š/, /z/, /ž/. Hlavní důvod této analýzy vyplývá z faktu, že čeština na začátku a konci slabik nedovoluje kombinace alveolárních okluzív se zmíněnými sykavkami, ačkoliv okluzívy jiného místa artikulace (kromě palatálních, které nestojí před žádným konsonantem) se s nimi kombinují, a proto je tak analyzujeme (srov. *psát*, *kšandy*, *zips*, *jakžtakž*). Máme za to, že afrikáty tyto kombinace suplují (pro podrobnější výklad viz Bičan 2013).⁴ Uvnitř slov uvedené kombinace nikdy nerozdělujeme do různých slabik. V opačném případě bychom museli říct, že se slova *podšít* a *počít* skládají ze stejných slabik, byť jsou fonologicky odlišná.

Za druhé, pracujeme s neutralizací znělosti. Jak známo, na konci slov (resp. přízvukových taktů) není v češtině rozdíl mezi neznělými a znělými konsonanty fonologicky relevantní. Máme-li se tedy řídit pravidlem slabičného dělení, které zapovídá slabikování vedoucí k nepřipustným tvarům slabik, musíme např. formu /podlaha/ dělit na /po-dla-ha/, nikoliv na /pod-la-ha/, poněvadž slabika /pod/ není kvůli neutralizaci znělosti na konci slov přípustná.

2. Frekvence slabičných typů bez ohledu na typ slabičného jádra

Jádrem slabiky může být v češtině krátký, dlouhý nebo diftongální vokál či slabičná sonanta /r/ nebo /l/. Je-li jádrem slabiky vokál, může sám o sobě tvořit slabiku (srov. *a*). Před slabičnými sonantami však musí stát minimálně jeden konsonant (srov. *(být) hr*) a zpravidla alespoň jeden konsonant následuje (srov. *trn*, *vlk*). Na začátku slabiky se vyskytuje jeden až pět konsonantů (srov. *on*, *jen*, *prach*, *strach*, *vzpřímený*, *vzkvět*, tj. /FSkvjeT/),⁵ zatímco na konci slabiky stojí maximálně tři konsonanty (srov. *to*, *pes*, *kost*, *zábst*). Kombinací těchto možností získáme 20 různých slabičných typů – viz tabulka 1 (např. *strach* odpovídá typu CCCVC).⁶ Typům, které jsou označeny hvězdičkami, se budeme věnovat v následujícím oddíle.

4 Fakt, že afrikáty *c* a *č* alternují např. s *k* (*ruka* ~ *ruce* ~ *ručka*), nemluví pro dvoufonémovou interpretaci, poněvadž není důvod, proč by dva fonémy nemohly alternovat s fonémem jedním (srov. alternace /b/ ~ /bj/ v *bába* ~ *bábě* oproti *voda* ~ *vodě*).

5 Velká písmena označují archifonémy jako produkty neutralizace znělosti. Jsou to okluzívy a frikativy, pro něž není v daném kontextu znělost fonologicky relevantní.

6 Stranou necháváme slovo *boršč*, jehož slabičná kóda je z hlediska struktury jiných slabičných kód anomální. Ať už totiž analyzujeme afrikáty jako dva fonémy, nebo foném jeden, setkáme se zde s kombinací šč, která se jinde v kódě nevyskytuje.

Tab. 1: Logicky možné slabičné typy v češtině

		Konec slabiky			
		-∅	-C	-CC	-CCC
Začátek sl.	∅-	V	VC	VCC	VCCC
	C-	CV	CVC	CVCC	CVCCC
	CC-	CCV	CCVC	CCVCC	CCVCCC
	CCC-	CCCV	CCCVC	CCCVCC	CCCVCCC
	CCCC-	CCCCV	CCCCVC	CCCCVCC	*CCCCVCCC
	CCCCC-	CCCCCV	CCCCCVC	*CCCCCVCC	*CCCCCVCCC

Procentuální zastoupení doložených slabičných typů v SSČ pro celou slovní zásobu bez ohledu na slovní druh nabízíme v tabulce 2. Slabiky s počátečními pěti konsonanty nebereme v úvahu, jelikož nejsou v SSČ doloženy. Je zřejmé, že otevřené slabiky (69,99%) jsou častější než slabiky zavřené (30,01%). Dále jsou preferovány slabiky, které začínají na jednoduchý konsonant (69,44%) spíše než na konsonant žádný nebo na konsonantů více (30,56%). Nejfrekventovanějším slabičným typem je CV. Hierarchie frekvence pěti nejčastějších typů je CV (48,05%) > CVC (18,17%) > CCV (17,11%) > CCVC (5,51%) > CVCC (3,13%).

Tab. 2: Celkové procentuální zastoupení slabičných typů (celkem 146 703 slabik)

Všechna jádra		Konec slabiky				Celkem
		-∅	-C	-CC	-CCC	
Začátek sl.	∅-	3,02%	1,41%	0,09%	0,0007%	4,52%
	C-	48,05%	18,17%	3,13%	0,08%	69,44%
	CC-	17,11%	5,51%	0,96%	0,03%	23,61%
	CCC-	1,76%	0,43%	0,16%	0,005%	2,36%
	CCCC-	0,05%	0,02%	0,001%	0%	0,07%
Celkem		69,99%	25,54%	4,34%	0,12%	100%

Uvedené výsledky se prakticky shodují s předchozími kvantitativními popisy založenými na textech.⁷ Pro mluvené texty je nejrozšířenější typ CV, ale jeho procentuální zastoupení je vyšší (59,76%). Celkově je v mluvených textech vyšší i výskyt otevřených slabik; Ludvíková uvádí poměr 75,14% otevřených slabik oproti 24,86% zavřených slabik (TĚŠITĚLOVÁ et al. 1985, 23–24). V mluvených textech jsou preferovány stejné slabičné typy, ale hierarchie se částečně liší: CV (59,76%) > CVC (17,18%) > CCV (10,06%) > V (4,54%) > CCVC (3,76%). V lexiku má tedy větší frekvenci typ

7 Naše výsledky budeme srovnávat především s výsledky Ludvíkové (viz citovaná literatura).

CCVC (5,51%), který je v textech až na 8. místě s frekvencí 0,62%. I přesto je zřejmé, že preference určitých slabičných typů je stabilním rysem češtiny, a to jak ve slovní zásobě jako celku, tak v textech, které obsahují různé morfologické tvary slov.

3. Frekvence slabičných typů podle typu slabičného jádra

Jak již bylo řečeno, jádrem slabiky mohou být v češtině čtyři typy fonémů: krátké, dlouhé a diftongální vokály a dále sonanty /r/ a /l/.⁸ Jelikož se předchozí popisy rozdílů mezi slabičnými jádry nevěnovaly, v tabulkách 3–6 nabízíme frekvence slabičných typů podle typu jádra.

Tab. 3: Procentuální zastoupení slabičných typů pro krátké vokály (celkem 113 878 slabik)

Krátké vokály		Konec slabiky				Celkem
		-∅	-C	-CC	-CCC	
Začátek sl.	∅-	3,52%	1,69%	0,1%	< 0,001%	5,32%
	C-	47,27%	19,9%	3,65%	0,07%	70,89%
	CC-	15,96%	4,87%	1,04%	0,01%	21,88%
	CCC-	1,28%	0,4%	0,2%	0,002%	1,88%
	CCCC-	0,03%	0,01%	< 0,001%	0%	0,05%
Celkem		68,07%	26,87%	4,99%	0,08%	100%

Tab. 4: Procentuální zastoupení slabičných typů pro dlouhé vokály (celkem 28 022 slabik)

Dlouhé vokály		Konec slabiky				Celkem
		-∅	-C	-CC	-CCC	
Začátek sl.	∅-	1,21%	0,5%	0,05%	0%	1,76%
	C-	53,06%	12,13%	1,4%	0,16%	65,75%
	CC-	22,2%	5,1%	0,75%	0,07%	28,12%
	CCC-	3,69%	0,56%	0,05%	0,01%	4,31%
	CCCC-	0,05%	0,004%	0,004%	0%	0,06%
Celkem		79,21%	18,3%	2,24%	0,25%	100%

8 Pro zjednodušení necháme stranou okrajové případy slov, v nichž je jádrem slabiky /m/ (*sedm*, *osm*). Na rozdíl od /r/ a /l/ se totiž ve spisovné češtině slabičné /m/ může realizovat s doprovodným vokálem. Proto slova *sedm* a *osm* přepisujeme jako /sedum/ a /osum/. Poznamenejme, že sonanty /r/ a /l/ nejsou na začátku slabiky před konsonantem slabičné čili /rti/ (*rtý*) je jedna slabika začínající na dva konsonanty stejně jako /Sto/ (*sto*). Ke slabičnosti /r/ a /l/ viz BIČAN (2014b).

Tab. 5: Procentuální zastoupení slabičných typů pro diftongální vokály (celkem 2 929 slabik)

Diftongální vokály		Konec slabiky				Celkem
		-∅	-C	-CC	-CCC	
Začátek sl.	∅-	2,8%	0,2%	0%	0%	3%
	C-	26,56%	13,49%	1,26%	0%	41,31%
	CC-	15,33%	36,94%	0,38%	0,38%	53,02%
	CCC-	1,98%	0,61%	0%	0,07%	2,66%
	CCCC-	0%	0%	0%	0%	0%
Celkem		46,67%	51,25%	1,64%	0,44%	100%

Tab. 6: Procentuální zastoupení slabičných typů pro slabičné sonanty (celkem 1 874 slabik)

Sonanty		Konec slabiky				Celkem
		-∅	-C	-CC	-CCC	
Začátek sl.	∅-	0%	0%	0%	0%	0%
	C-	69,21%	10,94%	0,85%	0%	81%
	CC-	13,82%	1,92%	0,05%	0%	15,8%
	CCC-	1,55%	0,16%	0,05%	0%	1,76%
	CCCC-	1,01%	0,43%	0%	0%	1,44%
Celkem		85,59%	13,45%	0,96%	0%	100%

Porovnáním výskytu slabičných typů pro různé typy slabičných vrcholů dojdeme k několika závěrům. Především k tomu, že u drtivé většiny slabik je jádrem krátký vokál (77,62%); druhým nejčastějším slabičným jádrem jsou dlouhé vokály (19,1%), dále následují diftongální vokály (2%) a nejméně časté jsou slabičné sonanty (1,28%). Pro mluvené texty (rozhovor) uvádí LUDVÍKOVÁ (1976, 93) procentuální zastoupení krátkých vokálů 82,16% a pro dlouhé vokály (kam patrně řadí i vokály diftongální) 17,84%; slabičné sonanty jsou zastoupeny 0,42%. Počítáme-li poměr mezi krátkými a dlouhými vokály stejně jako Ludvíková, pak získáme poměr 78,63% krátkých vokálů oproti 21,37% dlouhých vokálů. V lexikonu je tedy o něco menší zastoupení krátkých vokálů, ale zároveň je tam větší zastoupení slabičných sonant.

Kromě diftongálních vokálů jsou rozdíly v preferenci slabičných typů mezi různými slabičnými typy jen malé. Nejčastější je typ CV, dále CVC nebo CCV, dále CCVC a CVCC nebo CCCV. Opět kromě diftongálních vokálů převažují otevřené slabiky nad zavřenými (viz dále). Stejně tak převažují slabiky začínající na jednoduchý konsonant nad slabikami bez počátečního konsonantu nebo začínajícími na více než jeden konsonant. Tyto typy slabik jsou v největší míře doložené u slabičných sonant.

Diftongální vokály se oproti ostatním slabičným jádrům vyznačují několika specifiky. Za prvé, nejfrekventovanějším slabičným typem pro ně není CV, ale CCVC (36,94%; CV odpovídá 26,56%). Vysvětlení je však nasnadě: drtivá většina slov s tímto slabičným typem jsou slovesa zakončená na *-out* (např. *zašeptnout*), a tudíž uvedený výsledek bude velmi pravděpodobně ovlivněn vzorkem materiálu (lexikální korpus obsahuje slovesa jen v infinitivu). Bude zajímavé porovnat výsledky se situací v konkrétních textech, ale jak již bylo zmíněno, takové údaje zatím chybí.

Povahou materiálu si vysvětlujeme i fakt, že oproti jiným slabičným jádrům převažují u diftongálních vokálů zavřené slabiky nad slabikami otevřenými (srov. tabulka 7). Je to právě typ CCVC, který přispívá k vysoké frekvenci zavřených slabik.

Tab. 7: Poměr mezi otevřenými a zavřenými slabikami podle slabičného jádra

Slabiky	Krátké vokály	Dlouhé vokály	Diftongální vokály	Sonanty	Celkem
Otevřené	68,08%	79,21%	46,67%	85,59%	69,99%
Zavřené	31,93%	20,79%	53,33%	14,41%	30,01%

Zvolenému materiálu, byť z jiného důvodu, musíme též přiřknout fakt, že není doložena ani jedna slabika, v níž by před diftongálním vokálem stály čtyři konsonanty. Takové slabiky v češtině možné jsou, jen se nevyskytují v SSČ. V celém lexikálním korpusu najdeme 11 příkladů slov, kde před diftongálním vokálem stojí čtyři konsonanty (např. *pstroužek*, *štknouti* – obě v SSJČ).

Zajímavá je též struktura slabik se slabičnými sonantami. Především pro ně neexistují slabiky bez počátečního konsonantu, jak už jsme zmínili v předchozím oddíle. Důležitější jsou ovšem dvě další okolnosti. Především není doložena ani jedna slabika, v níž by za slabičnou sonantou následovaly tři konsonanty. To však platí jen pro zvolenou slovní zásobu, poněvadž existuje slovo *šmrnc* (SSJČ) – zde je však nutné si uvědomit, že afrikáty interpretujeme jako kombinace dvou konsonantů (viz oddíl 2).

Daleko závažnější je zjištění, že pro slabičné sonanty nejsou doloženy slabiky typů CCCVCCC, CCCCVCC a CCCCVCCC, a to ani v celém lexikálním korpusu. Uvedené typy jsou všechny typy slabik, v nichž by se kromě slabičných sonant vyskytovalo dalších šest konsonantů. Tento fakt lze zobecnit následujícím distribučním omezením (viz Bičan 2013, 121):

(2) *Omezení počtu nevokalických fonémů v české slabice*

V češtině žádná slabika neobsahuje více než šest nevokalických fonémů.

Nevokalickými fonémy jsou jak konsonanty, tak sonanty /r/ a /l/, ať už slabičné, či nikoliv. Uvedené pravidlo má širší dosah než na slabičné sonanty. Zde se musí-

me vrátit k tabulce 1, v níž byly slabičné typy CCCCVC, CCCCCV, CCCCCV-CCC označeny hvězdičkou. Ani jeden z nich není totiž v celém lexikálním korpusu (o 950 000 slabikách) doložen. Jejich existenci nepotvrzují ani jiné zdroje, a rovněž ani Ludvíková, ani Kučera je nezmiňují. Společným rysem uvedených typů je přítomnost více než šesti nevokalických fonémů. Znamená to, že je-li jádrem slabiky vokál, slabika v češtině může obsahovat maximálně šest nevokalických fonémů; je-li jádrem slabiky slabičná sonanta, snižuje se tento počet na pět, aby byl maximální počet neslabičných fonémů zachován. Pravidlo (2) tedy zohledňuje obě možnosti.

Dva z nemožných slabičných typů mají na začátku pět konsonantů. Pětikonsonantické začátky slabik jsou sice velmi vzácné (v SSČ se nevyskytují), ale nikoliv nedoložené. V celém lexikálním korpusu najdeme tři příklady: *vzkvět* (tj. /FSkvjeT/, SSJČ), *vzkvětáváti* (PSJČ), *zčtvernásobiti* (SSJČ; zde samozřejmě, jen pokud afrikáty pokládáme za kombinace dvou konsonantů). Myslitelný je i tvar *vzdměte* (tj. /FSd-mněte/) jako imperativ od slovesa *vzdmouti*.

4. Frekvence slabičných typů podle slovního druhu

Dalším problémem, který doposud zůstával stranou kvantitativních analýz slabiky, je otázka, zda je slabičná struktura stejná ve všech typech gramatických slov, tj. zda některé slovní druhy preferují určitý typ slabik. Intuitivně můžeme předpokládat, že se fonologická struktura ohebných gramatických slov bude lišit od struktury slov neohebných, avšak v případě slabičné struktury nebyla doposud tato intuitivní představa kvantitativně potvrzena. Jelikož je slovní zásoba zařazená do lexikálního korpusu označována tak, aby ji bylo možné třídit podle slovních druhů, zaměřme se na tento aspekt. Procentuální zastoupení slabičných typů pro jednotlivé slovní druhy je uvedeno v tabulkách 8–17. Poznamenejme, že pokud se gramatické slovo určitého slovního druhu skládá z více fonologických slov, pak všechna tato fonologická slova řadíme pod daný slovní druh.

Tab. 8: Procentuální zastoupení slabičných typů pro podstatná jména (59 756 slabik celkem)

Podstatná jména		Konec slabiky				Celkem
		-∅	-C	-CC	-CCC	
Začátek sl.	∅-	2,99%	1,49%	0,1%	0,002%	4,58%
	C-	47,83%	17,9%	5,31%	0,06%	71,11%
	CC-	14,65%	5,81%	1,72%	0,01%	22,18%
	CCC-	1,21%	0,52%	0,33%	0,002%	2,07%
	CCCC-	0,03%	0,02%	0,002%	0%	0,06%
Celkem		66,71%	25,75%	7,46%	0,07%	100%

Tab. 9: Procentuální zastoupení slabičných typů pro přídavná jména (39 913 slabik celkem)

Přídavná jména		Konec slabiky				Celkem
		-∅	-C	-CC	-CCC	
Začátek sl.	∅-	2,54%	1,19%	0,09%	0,07%	3,89%
	C-	57,68%	10,14%	2,39%	0%	70,22%
	CC-	19,64%	2,87%	0,48%	0,01%	22,99%
	CCC-	2,5%	0,3%	0,04%	0%	2,85%
	CCCC-	0,05%	0,01%	0%	0%	0,06%
Celkem		82,41%	14,51%	3%	0,08%	100%

Tab. 10: Procentuální zastoupení slabičných typů pro zájmena (318 slabik celkem)

Zájmena		Konec slabiky				Celkem
		-∅	-C	-CC	-CCC	
Začátek sl.	∅-	1,57%	0,31%	0%	0%	1,89%
	C-	50,63%	18,55%	2,83%	0%	72,01%
	CC-	21,07%	4,72%	0,31%	0%	26,1%
	CCC-	0%	0%	0%	0%	0%
	CCCC-	0%	0%	0%	0%	0%
Celkem		73,27%	23,58%	3,14%	0%	100%

Tab. 11: Procentuální zastoupení slabičných typů pro číslovky (397 slabik celkem)

Číslovky		Konec slabiky				Celkem
		-∅	-C	-CC	-CCC	
Začátek sl.	∅-	3,78%	0%	0%	0%	3,78%
	C-	59,7%	10,08%	4,79%	1,76%	76,32%
	CC-	13,1%	2,02%	1,26%	0,5%	16,88%
	CCC-	2,52%	0%	0%	0%	2,52%
	CCCC-	0,5%	0%	0%	0%	0,5%
Celkem		79,6%	12,09%	6,05%	2,27%	100%

Tab. 12: Procentuální zastoupení slabičných typů pro slovesa (33 770 slabik celkem)

Slovesa		Konec slabiky				Celkem
		-∅	-C	-CC	-CCC	
Začátek sl.	∅-	3,76%	1,58%	0,05%	0%	5,39%
	C-	36,45%	30,41%	0,42%	0,12%	67,39%
	CC-	15,77%	8,73%	0,4%	0,09%	24,98%
	CCC-	1,65%	0,44%	0,05%	0,02%	2,15%
	CCCC-	0,07%	0,02%	0%	0%	0,09%
Celkem		57,69%	41,17%	0,91%	0,23%	100%

Tab. 13: Procentuální zastoupení slabičných typů pro příslovce (12 203 slabik celkem)

Příslovce		Konec slabiky				Celkem
		-∅	-C	-CC	-CCC	
Začátek sl.	∅-	2,71%	1,34%	0,11%	0,07%	4,22%
	C-	49,27%	12,05%	2,34%	0%	63,67%
	CC-	24,76%	4,04%	0,37%	0,03%	29,2%
	CCC-	2,38%	0,39%	0,08%	0%	2,85%
	CCCC-	0,04%	0,01%	0,01%	0%	0,06%
Celkem		79,17	17,82%	2,91%	0,1%	100%

Tab. 14: Procentuální zastoupení slabičných typů pro předložky (146 slabik celkem)

Předložky		Konec slabiky				Celkem
		-∅	-C	-CC	-CCC	
Začátek sl.	∅-	4,79%	1,37%	0%	0%	6,16%
	C-	54,11%	12,33%	3,42%	0%	69,86%
	CC-	15,75%	5,48%	1,37%	0%	22,6%
	CCC-	0,68%	0,68%	0%	0%	1,37%
	CCCC-	0%	0%	0%	0%	0%
Celkem		75,34%	19,86%	4,79%	0%	100%

Tab. 15: Procentuální zastoupení slabičných typů pro spojky (143 slabik celkem)

Spojky		Konec slabiky				Celkem
		-∅	-C	-CC	-CCC	
Začátek sl.	∅-	3,5%	2,1%	2,1%	0%	7,69%
	C-	48,95%	22,38%	2,8%	0%	74,13%
	CC-	13,99%	3,5%	0%	0%	17,48%
	CCC-	0,7%	0%	0%	0%	0,7%
	CCCC-	0%	0%	0%	0%	0%
Celkem		67,13%	27,97%	4,9%	0%	100%

Tab. 16: Procentuální zastoupení slabičných typů pro částice (94 slabik celkem)

Částice		Konec slabiky				Celkem
		-∅	-C	-CC	-CCC	
Začátek sl.	∅-	5,32%	3,19%	0%	0%	8,51%
	C-	47,87%	28,72%	5,32%	0%	81,91%
	CC-	7,45%	1,06%	0%	0%	8,51%
	CCC-	0%	1,06%	0%	0%	1,06%
	CCCC-	0%	0%	0%	0%	0%
Celkem		60,64%	34,04%	5,32%	0%	100%

Tab. 17: Procentuální zastoupení slabičných typů pro citoslovce (236 slabik celkem)

Citoslovce		Konec slabiky				Celkem
		-∅	-C	-CC	-CCC	
Začátek sl.	∅-	3,39%	2,12%	0%	0%	5,51%
	C-	53,81%	13,98%	1,69%	0%	69,49%
	CC-	17,37%	6,36%	0,42%	0%	24,15%
	CCC-	0%	0,42%	0,42%	0%	0,85%
	CCCC-	0%	0%	0%	0%	0%
Celkem		74,58%	22,88%	2,54%	0%	100%

Z tabulek 8–17 lze vyvodit několik závěrů. Především všechny slovní druhy preferují otevřené slabiky. Největší frekvencí otevřených slabik se vyznačují přídavná jména (82,41%) a číslovky (79,6%), zatímco nejméně frekventované jsou otevřené slabiky u částic (60,64%) a sloves (57,69%). Toto zjištění není překvapující, jelikož přídavná jména, až na několik výjimek (např. *blond*), končí otevřenou slabikou a slovesa v infinitivu končí zpravidla na slabiku zavřenou.⁹ Stejně tak jsou ve všech slovních

9 Mimo archaické tvary na *-ti*, které se však v SSČ nevyskytují.

druzích nejfrekventovanější slabiky s počátečním jednoduchým konsonantem, byť se konkrétní procentuální zastoupení liší. Nejčastější jsou u částic (81,91%) a číslovky (76,32%), nejméně časté u sloves (67,39%) a příslovčí (63,67%).

Celkově se jednotlivé slovní druhy shodují v preferenci stejných slabičných typů, a to CV, CCV, CVC, CVCC, CCVC a V. Pro přehled shrnujeme hierarchie pro všechny slovní druhy (celková hierarchie je CV > CVC > CCV > CCVC > CVCC, viz výše):

- (3) *Hierarchie slabičných typů pro jednotlivé slovní druhy*
- | | |
|------------------|--|
| Podstatná jména: | CV > CVC > CCV > CCVC > CVCC |
| Přídavná jména: | CV > CCV > CVC > CCVC > V |
| Zájmena: | CV > CCV > CVC > CCVC > CVCC |
| Číslovky: | CV > CCV > CVC > CVCC > V |
| Slovesa: | CV > CVC > CCV > CCVC > V |
| Příslovce: | CV > CCV > CVC > CCVC > V |
| Předložky: | CV > CCV > CVC > CCVC > V |
| Spojky: | CV > CVC > CCV > CCVC, V ¹⁰ |
| Částice: | CV > CVC > CCV > CVCC, V |
| Citoslovce: | CV > CCV > CVC > CCVC > V |

Slovní druhy se tedy liší pouze v pořadí uvedených šesti nejpreferovanějších typů. Pochopitelně se rovněž liší v procentuálním zastoupení, ale s ohledem na různý počet slabik u jednotlivých druhů (např. 59 756 slabik u podstatných jmen oproti 94 slabikám u částic) nelze rozdílu přikládat velkou váhu.

Ve všech případech je nejčastějším slabičným typem CV, tedy otevřená slabika s jedním počátečním konsonantem. Obvykle se tvrdí, že takové slabiky existují ve všech jazycích (MADDIESON 2013), s kterýmžto tvrzením však musíme nakládat opatrně, poněvadž data ze všech jazyků nejsou a nemohou být známa a konkrétní podoba slabiky závisí na způsobu slabikování, jenž nemá univerzální charakter. Ačkoliv není zcela jasné, do jaké míry je typ CV univerzálně preferovaný, v češtině je rozhodně favorizovaný. U všech slovních druhů je zhruba polovina všech slabik tohoto typu, což je při existenci 20 různých slabičných typů obrovské číslo.

Tabulky 8–17 dále jasně ukazují, že ne všechny slabičné typy jsou doloženy pro každý slovní druh. Neohebná slova (předložky, spojky, částice a citoslovce) vykazují jednodušší slabičnou strukturou než slova ohebná. Neohebná slova nedovolují slabiky se čtyřmi počátečními konsonanty a slabiky se třemi koncovými konsonanty. Takové slabiky nejsou u neohebných slov doloženy ani v celém lexikálním korpusu. Rovněž v celém lexikálním korpusu nenajdeme pro tyto slovní druhy slabiky typu CCCVCC. Uvedená fakta můžeme postihnout následujícími distribučními omezeními, která se částečně překrývají:

10 Tyto typy mají stejnou frekvenci.

(4) Omezení počtu konsonantů ve slabikách u neohebných slovních druhů

- (a) Slabika nemůže obsahovat více než 4 konsonanty.
- (b) Slabika nemůže končit na více než 2 konsonanty.

Ze všech slovních druhů vůbec nejjednodušší slabičnou strukturu mají zájmena, poněvadž se u nich objevují jen slabiky začínající nebo končící maximálně na dva konsonanty. To potvrzují i data z celého lexikálního korpusu. Jinými slovy, platí tato distribuční omezení, která se opět částečně překrývají (omezení (5b) je shodné se (4b)):

(5) Omezení počtu konsonantů ve slabikách u zájmen

- (a) Slabika nemůže začínat na více než 2 konsonanty.
- (b) Slabika nemůže končit na více než 2 konsonanty.

Mezi zájmeny a neohebnými slovy je tedy určitá fonologická podobnost, jelikož všechna tato slova mají oproti ostatním jednodušší slabičnou strukturu. Shodují se také v tom, že jejich celková délka měřená podle průměrného počtu slabik ve slově je menší než u jiných slovních druhů. Toto dokládají údaje v tabulce 18, kde jsou uvedeny průměrné délky slov podle počtu slabik (řádek *Slab./slovo*). Zájmena a neohebné slovní druhy se průměrně skládají ze dvou slabik, zatímco ostatní slovní druhy ze slabik tří. Jsou tedy průměrně o jednu slabiku kratší. Nejkratší jsou citoslovce (1,74 slabiky) a nejdelší přídavná jména (3,68 slabiky).

Tab. 18: Průměrný počet slabik a průměrná délka slabik pro jednotlivé slovní druhy

Průměrný počet	Podstatná jména	Přídavná jména	Zájmena	Číslovky	Slovesa
Slab./slovo	3,02	3,68	2,19	3,18	3,04
Foném/slab.	2,63	2,46	2,54	2,51	2,68

	Příslovce	Předložky	Spojky	Částice	Citoslovce	Celkem
Slab./slovo	3,21	2,15	2,13	2,09	1,74	3,19
Foném/slab.	2,55	2,49	2,49	2,47	2,48	2,59

Celková průměrná délka slov je 3,19 slabiky. Údaj platí jen pro lexikon, protože v mluvených textech je slabičná délka výrazně menší (LUDVÍKOVÁ 1976, 93): 2,28 (vědecká zpráva), popř. 1,95 (rozhovor).¹¹ Tento výrazný rozdíl mezi lexikonem a texty je jednoduše vysvětlitelný. Na jedné straně se v lexikonu vyskytuje mnoho víceslabičných slov, na straně druhé se mluvené texty vyznačují velkým výskytem

11 Ludvíková ovšem vychází z gramatických (ortografických) slov, nikoliv ze slov fonologických.

jednoslabičných slov (srov. *je, jsem, předložky* atd.). Jednoslabičná slova se navíc v mluvených textech opakují a jejich celkový výskyt je velký (LUDVÍKOVÁ *ibid.* pro rozhovor uvádí zastoupení 43,2%, ve vědecké zprávě však jen 29,32%).

Tabulka 18 též nabízí hodnoty průměrného počtu fonémů ve slabice pro každý slovní druh (řádek *Foném/slab.*). Jak patrně, mezi jednotlivými slovními druhy nejsou markantní rozdíly a především u neohebných druhů je délka prakticky stejná. Celková průměrná hodnota pro všechny slovní druhy je 2,59 fonému ve slově. LUDVÍKOVÁ (*ibid.*) pro texty uvádí hodnotu 2,53 (vědecká zpráva) a 2,38 (rozhovor).

Podívejme se konečně na procentuální zastoupení slov podle počtu slabik. Tabulka 19 nabízí pro každý slovní druh zastoupení slov obsahující 1 až 5 slabik. Delší slova nebereme v potaz. Ze součtu celkového procentuálního zastoupení v posledním řádku tabulky je patrné, že více než pětislabičná slova tvoří toliko 2,47% lexika. Šedé zbarvení zvýrazňuje nejčastěji doložená slova. Jak u zájmen, tak u neohebných slov je největší zastoupení slov dvojslabičných (kromě citoslovců, u nichž převládají jednoslabičná slova), zatímco u ostatních slovních druhů s výjimkou číslovek jsou to slova tříslabičná. Jak již bylo řečeno, zájmena a neohebná slova průměrně obsahují dvě slabiky, a proto není největší výskyt dvojslabičných slov překvapivý. Nicméně z čistě matematického hlediska nutně neplatí, že slova s průměrnou délkou budou nejvíce frekventovaná. To vidíme například u číslovek, u nichž je průměrná délka 3,18 slabik, ale nejvíce jsou zastoupena dvojslabičná slova. Konečně je z údajů v tabulce 19 zřejmé, že jak zájmena, tak spojky a částice nemají slova o pěti slabikách; nemají ani slova o více slabikách.

Tab. 19: Procentuální zastoupení slov o daném počtu slabik pro jednotlivé slovní druhy

	Procentuální zastoupení slov o daném počtu slabik				
	1	2	3	4	5
Podstatná j.	5,80%	27,88%	36,67%	20,7%	6,77%
Přídavná j.	0,44%	9,78%	36,46%	34,77%	13,73%
Zájmena	24,83%	39,31%	27,59%	8,28%	0%
Číslovky	5,60%	28%	27,2%	24,8%	11,2%
Slovesa	2,01%	24,54%	46,31%	22,74%	3,51%
Příslovce	2,16%	21,07%	40,75%	27,55%	6,87%
Předložky	32,35%	38,24%	16,18%	8,82%	4,41%
Spojky	20,9%	46,27%	31,34%	1,49%	0%
Částice	33,33%	35,56%	20%	11,11%	0%
Citoslovce	49,27%	34,56%	11,03%	3,68%	1,47%
Celkem	3,44%	22,33%	39,18%	25,01%	7,57%

Procentuální zastoupení slov o určitém počtu slabik pro jednotlivé slovní druhy uvádějí i BARTOŇ et al. (2009, 114). Své výpočty zakládají na datech z Českého národního korpusu (SYN2005), který automaticky transkribovali. Jejich výsledky se však s našimi neshodují, což je dáno tím, že autoři do své analýzy zahrnují i různé morfologické tvary gramatických slov. My naopak vycházíme z lexika a základní jednotkou je pro nás fonologické slovo. Například celkově nejfrekventovanějšími jsou v ČNK gramatická slova čtyřslabičná (32,08%). Podrobnosti viz ve zmíněné publikaci.

5. Závěr a shrnutí

Cílem našeho článku bylo kvantitativně popsat slabičnou strukturu slov v češtině, a to především z hlediska doložených slabičných typů. Kvantitativně orientované analýzy zaměřené na slabičnou strukturu lexikonu pro jiné jazyky existují (ROUSSET 2004), pro češtinu však doposud chyběly. Zdrojem popisu byla lexikální slovní zásoba zaznamenaná ve *Slovníku spisovné češtiny*, jež je součástí Fonologického lexikálního korpusu. Náš vzorek obsahuje celkem 45 978 fonologických slov a 146 703 slabik. Celý lexikální korpus se potom v současnosti skládá z 257 855 slov a z 950 949 slabik.

Čeština se vyznačuje oblibou určitých slabičných typů. Tato obliba se projevuje u všech slovních druhů a potvrzují ji i předchozí kvantitativní popisy založené na textech (psaných, či mluvených). Je to tedy konstantní vlastnost češtiny. Na jedné straně tento jazyk preferuje otevřené slabiky a na straně druhé slabiky začínající na jednoduchý konsonant. Vůbec nejfrekventovanějším slabičným typem je CV. Uvedené preference platí pro všechny typy slabičných jader kromě diftongálních vokálů. U nich nicméně předpokládáme, že odchylky nejsou dány specifícností této vokalické třídy, nýbrž že se zde projevuje zvolený materiál. U sloves jsou totiž zahrnuty pouze infinitivy (srov. jejich časté zakončení na *-out*), ale už ne jejich různé tvary.

Uvedená zjištění nemusí být na první pohled nijak zvlášť překvapivá. Otevřené slabiky, a zvlášť slabiky typu CV, jsou oblíbené i v jiných jazycích (jak potvrzuje např. ROUSSETOVÉ (*ibid.*) srovnání lexikonů 16 jazyků) a mnoho jazyků má jenom je (MADDISON 2013). Ve slovanských jazycích je tendence k otevřeným slabikám silná již od praslovanštiny.¹² Na druhou stranu je však v češtině na začátku a konci slabiky doloženo velké množství konsonantických kombinací. Na začátku se mohou objevit kombinace až pěti konsonantů; na konci pak mohou stát kombinace až tří konsonantů. V našem materiálu najdeme 337 různých počátečních konsonantických kombinací a 73 různých koncových kombinací.¹³ I přes tyto velké kombinační možnosti na konsonantickou kombinaci začíná pouze 26,04% slabik a pouze 4,44%

12 O praslovanštině se předpokládá, že dovolovala pouze otevřené slabiky (tzv. „zákon otevřených slabik“), byť bylo toto pravidlo na konci praslovanského období následkem zániku jevů porušováno.

13 To však zcela nevyčerpává všechny možnosti. V celém lexikálním korpusu takových kombinací

slabik na ni končí (na jednoduchý konsonant začíná 69,44% slabik a na žádný 4,52% slabik; na jeden konsonant končí 25,54% slabik a na žádný 69,99% slabik). To znamená, že čeština sice podle počtu dovoluje mnoho komplexních slabičných iniciál a kód, ale jejich frekvence je v lexiku oproti jednoduchým slabikám typů V, CV, VC, CVC malá.

Díky velkému množství počátečních a koncových konsonantických kombinací má slabika v češtině poměrně komplexní strukturu. Naše data ukazují, že tato struktura má svá omezení. Především počet počátečních konsonantů nepřesahuje pět a stejně tak počet koncových konsonantů nepřesahuje tři. Kromě toho existuje pro počet konsonantů ve slabice ještě další omezení, a to takové, že celkový počet nevokalických fonémů (tj. konsonantů včetně slabičných sonant) je omezen na maximální počet šest. Jinými slovy, je-li slabičným jádrem vokál, součet počátečních a koncových konsonantů nebude nikdy větší než šest. Je-li slabičným jádrem sonanta /r/ nebo /l/, pak součet konsonantů nepřesahuje počet pěti (šestým nevokalickým fonémem je sonanta). Slabiky porušující uvedené pravidlo nezmiňují ani předchozí popisy, ani se nevyskytují v celém lexikálním korpusu s téměř milionem slabik. Důvody takového omezení mohou souviset s artikulační náročností, ale mohou též vyplývat z morfologické struktury češtiny (malá přítomnost neslabičných afixů).

Ačkoliv se výše uvedené preference slabičných typů shodně projevují u všech slovních druhů, naše data naznačují, že se některé slovní druhy řídí dalšími pravidly omezujícími počet konsonantů ve slabice. U neohebných slovních druhů se neobjevují slabiky obsahující celkově více než čtyři konsonanty; zároveň u nich nejsou doloženy slabiky končící na tři konsonanty. Obě omezení opět potvrzují data z celého lexikálního korpusu. Vůbec nejjednodušší slabičnou strukturou se vyznačují zájmena. U nich slabiky nezačínají více než dvěma konsonanty a stejně jako u neohebných slovních druhů nekončí na tři konsonanty. Obecně tedy můžeme konstatovat, že slova s primárně gramatickou funkcí (v širším slova smyslu) mají v češtině jednodušší slabičnou strukturu než slova plnovýznamová. To nemusí být náhoda, ale může to souviset s tím, že přinejmenším v indoevropských jazycích mají zájmena a další slova s gramatickou funkcí jednodušší strukturu (např. pro praindoevropštinu se předpokládá, že tato slova byla tvořena kořeny s otevřenou slabikou, zatímco jména a slovesa se odvozují od kořenů tvořených zavřenou slabikou, srov. GAMKRELIDZE – IVANOV 1995, 188).

Konečně poznamenejme, že existuje vztah mezi slabičnou strukturou a polohou slabiky ve slově. Již Ludvíková (TĚŠITELOVÁ et al. 1985, 153) tento vztah prokázala na textech a data z lexikálního korpusu ho jen dokazují. Stručně řečeno, výskyt a frekvence slabičných typů se liší v iniciálních, mediálních a finálních slabikách slov.

existuje daleko více a v naší předchozí práci (BIČAN 2013) jsme popsali 442 počátečních kombinací a 104 koncových kombinací.

Otevřené slabiky jsou nejméně časté ve finální pozici (55,14%), nejčastější v pozici iniciální (75,63%), i když v mediální pozici jsou téměř stejně frekventované (72,14%).¹⁴ Naopak v iniciálních slabikách se nejméně objevují slabiky s jedním počátečním konsonantem (57,33%); takové slabiky jsou nejčastější v mediální pozici (73,05%) a dále ve finální pozici (71,6%). Ve všech pozicích je opět shodně nejfrekventovanější typ CV, který má největší zastoupení v pozici mediální (52,53%), dále v pozici iniciální (43,64%) a nejmenší v pozici finální (37,58%). Zde je však třeba si uvědomit, že uvedená zjištění jsou částečně závislá na způsobu slabikování slov. Např. slovo *podlaha* obsahuje mediální slabiku s komplexní iniciálou, poněvadž ho dělíme na /po.dla.ha/ (viz výše), ale jinak zvolená kritéria slabikování by ho mohla dělit na /pod.la.ha/.

Náš článek rozhodně nevyčerpává všechny možnosti, jak slabičnou strukturu jazyka popsat. Pro češtinu byly některé z nich již popsány Henry Kučerou a především Marií Ludvíkovou. I pro ostatní jazyky existuje bohatá literatura. V našem článku jsme se zaměřili na aspekty, které nebyly doposud pro češtinu adekvátně popsány. Tyto výsledky bude nutné ověřit na datech z textů, byť některá zjištění, jako ta týkající se neohebných slov, budou platit i tam.

Jako jedno z možných témat dalšího zkoumání se nabízí otázka vztahu frekvence slov a slabičné struktury a vztahu slabičné struktury a typu lexika. Je slabičná struktura odlišná u méně frekventovaných slov, u archaických slov, u expresivních slov nebo u slov cizího původu? Na závěr proto porovnejme frekvence slabičných typů pro podstatná jména ze tří vzorků, a to u slov doložených jen v PSJČ, ale nikoliv v SSJČ, SSČ nebo jinde, což jsou převážně archaismy, okazionalismy a dialektismy (= tabulka 20), u slov zaznamenaných v ROMPORTL (1978), tedy u slov cizího původu (= tabulka 21) a konečně u slov sesbíraných v SN_{1, 2} a CSN, tedy u poměrně nové slovní zásoby (= tabulka 22). Porovnáme-li tyto tabulky se situací u podstatných jmen v SSČ (= tabulka 8), vidíme, že ve všech případech jsou vždy třemi nejfrekventovanějšími typy CV, CVC a CCV, byť se jejich procentuální zastoupení liší (např. ve VSČ je největší zastoupení typu CV). Stejně tak ve všech vzorcích převažují otevřené slabiky a slabiky začínající na jeden konsonant. I v tomto případě se konkrétní procentuální zastoupení liší. Zda jsou tyto rozdíly statisticky významné, bude však nutné ověřit statistickými testy. Přesto je zřejmé, že rozdíly nebudou nijak zásadní.

14 V našich datech je 44 397 iniciálních a finálních slabik a 76 869 mediálních slabik. O mediálních slabikách pochopitelně můžeme mluvit jen u tří- a vícslabičných slov. Monosylabická slova v tomto případě nebereme v úvahu.

Tab. 20: Procentuální zastoupení slabičných typů pro podstatná jména v PSJČ (109 125 slabik celkem)

		Konec slabiky				Celkem
		-∅	-C	-CC	-CCC	
Začátek sl.	∅-	2,58%	1,3%	0,07%	0%	3,94%
	C-	49,31%	17,26%	5,69%	0,06%	72,33%
	CC-	15,04%	5,53%	1,09%	0,01%	21,67%
	CCC-	1,34%	0,45%	0,19%	0,001%	1,98%
	CCCC-	0,05%	0,03%	0%	0%	0,08%
Celkem		68,32%	24,57%	7,04%	0,07%	100%

Tab. 21: Procentuální zastoupení slabičných typů pro podstatná jména ve VŠČ (58 980 slabik celkem)

		Konec slabiky				Celkem
		-∅	-C	-CC	-CCC	
Začátek sl.	∅-	3,72%	2,47%	0,12%	0%	6,3%
	C-	56,54%	19,8%	1,61%	0,05%	78%
	CC-	10,6%	4,48%	0,34%	0,01%	15,44%
	CCC-	0,16%	0,09%	0,02%	0%	0,26%
	CCCC-	0%	0%	0%	0%	0%
Celkem		71,02%	26,84%	2,08%	0,07%	100%

Tab. 22: Procentuální zastoupení slabičných typů pro podstatná jména v SN1, 2 a CSN (27 628 slabik celkem)

		Konec slabiky				Celkem
		-∅	-C	-CC	-CCC	
Začátek sl.	∅-	3,52%	1,94%	0,17%	0,01%	5,64%
	C-	50,94%	19,68%	4,79%	0,06%	75,47%
	CC-	12,01%	4,85%	0,96%	0,01%	17,83%
	CCC-	0,6%	0,39%	0,06%	0%	1,06%
	CCCC-	0,004%	0%	0%	0%	0,004%
Celkem		67,07%	26,86%	5,99%	0,08%	100%

LITERATURA

- BARTOŇ, Tomáš et al. 2009. *Statistiky češtiny*. Praha: Nakladatelství Lidové noviny – Ústav Českého národního korpusu.
- BELL, Alan. 1976. The Distributional Syllable. In: JUILAND, Alphonse, ed. *Linguistic Studies Offered to Joseph Greenberg, Vol. 2: Phonology*. Saragota, California: Anma Libri, s. 249–262.
- BIČAN, Aleš. 2013. *Phonotactics of Czech*. Frankfurt am Main: Peter Lang.
- BIČAN, Aleš. 2014a. K pojmu fonologické slovo v češtině. In: BOČEK, Vít – VYKYPĚL, Bohumil, eds. *Sophia Slavica. Sborník prací věnovaných PhDr. Žofii Šarapatkové k osmdesátým narozeninám*. Brno: Tribun, s. 13–23.
- BIČAN, Aleš. 2014b. On the Czech Nuclear /r/ and /l/. In: WITKOŚ, Jacek – JAWORSKI, Sylwester, eds. *New Insights into Slavic Linguistics*. Frankfurt am Main: Peter Lang, s. 21–34.
- CSN = SOCHOVÁ, Zdeňka – POŠTOLKOVÁ, Běla. 1994. *Co v slovnících nenajdete: novinky v současné slovní zásobě*. Praha: Portál.
- Český fonologický lexikální korpus [online]. [cit. 1.6.2015]. Dostupné z: <http://www.ujc.cas.cz/phword>.
- DIXON, R. M. W. – AIKHENVALD, Alexandra Y. 2003. Word: A Typological Framework. In: DIXON, R. M. W. – AIKHENVALD, Alexandra Y., eds. *Word: A Cross-linguistic Typology*. Cambridge: Cambridge University Press, s. 1–41.
- GAMKRELIDZE, Thomas V. – IVANOV, Vjačeslav V. 1995. *Indo-European and the Indo-Europeans*. Berlin – New York: Mouton de Gruyter.
- HÁLA, Bohuslav. 1968. *Výslovnost spisovné češtiny*. 2. vydání. Praha: Academia.
- CHURAŇOVÁ, Eliška. 2014. The Consonantal-Vocalic Structure of the Czech Word Stress Group. *Phonetica Pragensia*. 13, s. 79–90.
- KUČERA, Henry – MONROE, George K. 1968. *A Comparative Quantitative Phonology of Russian, Czech, and German*. New York: Elsevier.
- KURYŁOWICZ, Jerzy. 1948. Contribution à la théorie de la syllabe. *Bulletin de la Société Polonaise de Linguistique*. 8, s. 80–114.
- LUDVÍKOVÁ, Marie. 1972a. Quantitative Syllable Analysis of Words in Czech. *Prague Studies in Mathematical Linguistics*. 3, s. 27–34.
- LUDVÍKOVÁ, Marie. 1972b. Some Quantitative Aspects of the Czech Syllables. *Prague Studies in Mathematical Linguistics*. 4, s. 141–54.
- LUDVÍKOVÁ, Marie. 1976. On Some Statistical Differences in Two Spoken Texts on the Syllabic Level. *Prague Studies in Mathematical Linguistics*. 5, s. 91–104.
- LUDVÍKOVÁ, Marie. 1978. On the Occurrence of Syllables in Different Word Positions. *Prague Studies in Mathematical Linguistics*. 6, s. 39–45.
- MADDISON, Ian. 2013. Syllable Structure. In: DRYER, Matthew – HASPELMATH, Martin, eds., *The World Atlas of Language Structures Online* [online]. Leipzig: Max Planck Institute for Evolutionary Anthropology [cit. 1.6.2015]. Dostupné z: <http://wals.info/chapter/12>.
- MARTINET, André. 2011. *Éléments de linguistique générale*. 5. vydání. Paris: Armand Colin.
- MULDER, Jan. 1989. *Foundations of Axiomatic Linguistics*. Berlin – New York: de Gruyter.
- PSJČ = *Příruční slovník jazyka českého*. Díl I–VIII. Praha: Státní pedagogické nakladatelství, 1935–1957.
- ROUSSET, Isabelle. 2004. *Structures syllabiques et lexicales des langues du monde* [online]. Gre-



noble. Disertační práce. Université Stendhal [cit. 1.6.2015]. Dostupné z: <https://tel.archives-ouvertes.fr/tel-00250154>.

SN₁, 2 = *Nová slova v češtině. Slovník neologizmů 1, 2*. Praha: Academia, 1998, 2004.

SSČ = *Slovník spisovné češtiny pro školu a veřejnost*. Praha: Academia, 2003.

SSJČ = *Slovník spisovného jazyka českého I-IV*. Praha: Academia, 1960–1971.

TĚŠITELOVÁ, Marie et al. 1985. *Kvantitativní charakteristiky současné češtiny*. Praha: Academia.

TRUBECKOJ, Nikolaj S. 1939. *Grundzüge der phonologie*. = *Travaux du Cercle Linguistique de Prague*. 7.

VSČ = ROMPORTL, Milan et al. 1978. *Výslovnost spisovné češtiny*. Praha: Academia.

Poděkování: Příspěvek vznikl za podpory grantu 13-15361P *Problémy ve fonologii slova v češtině* (GAČR).

Aleš Bičan

Institute of the Czech Language of the AS CR, v. v. i.

Veveří 97, 602 00 Brno

Czech Republic

bican@phil.muni.cz

