

# JAK ČASTO SE VYSKYTUJÍ (VYSKYTUJOU) TZV. HOVOROVÉ TVARY 1. OS. J. Č. A 3. OS. MN. Č. V ČESKÉM NÁRODNÍM KORPUSU?<sup>1</sup>

Neil Bermel

**Abstract: How often do the so-called “colloquial” 1 sg. and 3 pl. forms appear in the Czech National Corpus?**

This article takes as its starting point two places in Czech morphology where a contrast exists between the “standard” form, considered to be marked for its formality, and the “colloquial” form, considered to be marked for its informality. This arrangement leaves no place for a neutral form acceptable in all contexts, and it is thus of interest what the actual distribution of these forms is in written Czech. Following on a quantitative study of the appearance of “colloquial” vs. “standard” forms in common written texts, as exemplified in the Czech National Corpus (Bermel, in press), this study contrasts a purely quantitative approach with one that focuses on context and collocations in individual examples. Its findings in part show how contextual and collocational features complement, amplify, or undercut the quantitative results, and in part raise questions about the level of detail that is desirable, acceptable, and useful in these descriptions.

## 1. Úvod

Při kopírování názvu tohoto příspěvku z programu konference do textového souboru jsme si všimli, že v ní je pravopisná chyba. Textový editor Microsoft Word nám ukázal klikatou červenou čarou, že tvar *vyskytují* není správný, i když pak nedokázal nám doporučit lepší možnost. Přitom stálo hned vedle slovo *vyskytují*, podle pravopisného podprogramu tvar plně adekvátní.

Co je dovoleno a co dovoleno není se rychle zjišťuje pomocí empirického výzkumu. Podle Wordu

- já *píši* i *píšu*, žáci *píší* i *píšou*,
- já *vymaži* i *vymažu* program z počítače, správci počítačové sítě denně *vymaží* i *vymažou* programy,
- já *kouši* i *koušu* cizí lidi, jakož i zlí psi je *kouší* i *koušou*,
- já *piji* pivo, Češi *pijí* pivo, ale v žádném případě to pivo *nepiju*, ani Češi ho *nepijou*,
- já *nakupuji* v hypermarketu, přesně tak jak mnozí Češi tam *nakupují*, ale není přípustné, když *nakupuju*, ani když *nakupujou* Češi.

Touto volbou tvarů nejrozšířenější textový editor na světě nás vrátí ke kodifikačním příručkám 40. a 50. let minulého století. V nich F. Trávníček (1941, 1951) do-

<sup>1</sup> Tento výzkum vznikl s podporou British Academy (č. grantu RB104187) a účast na konferenci Korpus jako zdroj dat o češtině sponzoroval University of Sheffield Learned Societies Fund.

poručoval až na pár sloves dávat přednost tvarům tradičně spisovným (na *-i* a *-í*). Tvary analogické (na *-u* a *-ou*) připouštěl ve vzorech *psát* a *mazat* jen pro neformální kontexty, ve vzorech *pít* a *kupovat*, tj. po souhlásce *-j-*, je ale zakazoval po celé šíři spisovných kontextů.

V novějších příručkách -- začínaje s B. Havránkem a A. Jedličkou (viz např. 1963, 1998) a konče *Příruční mluvnici* brněnského Ústavu českého jazyka (1995) -- se ale k této volbě přistupuje méně kategoricky. Zdůrazňuje se přednost analogických tvarů v některých slovesných třídách a pokračující rozšiřování vhodnosti analogických tvarů v třídách jiných. Všimáme si v nich posunu z *nařizování* jedné či druhé varianty k *popisování* jejich užití. Cílem tohoto článku je zjistit, do jaké míry stav užívání těchto tvarů odpovídá předpisům.

Mluvnice od Trávníčkových až po nejnovější zachycují celou řadu faktorů, které ovlivňují přijatelnost tvarů:

**i. Slovesná třída.** Podle tradiční klasifikace jsou analogické tvary nepřijatelnější v 5. třídě. Nejméně přijatelné jsou v třídě první. Tato stupnice se zachovává ve všech příručkách, i když v detailech se jejich doporučení liší.

**Tabulka 1. Tvary a slovesné třídy**

Třída	SČ: 1. os. j. č., 3. os. mn. č. (tvary „tradiční“)	OČ: 1. os. j. č., 3. os. mn. č. (tvary „analogické“)
I.	kryji, kryjí	kryju, kryjou
V.	maži, maží	mažu, mažou
VI.	kupuji, kupují	kupuju, kupujou

**ii. Osoba.** 1. osoba j. č. na *-u* představuje tvar rozšířený po celém území České republiky. 3. osoba mn. č. na *-ou* je naopak tvarem charakteristickým jenom pro Čechy a západní Moravu: jinde v místních dialektech a v městské mluvě ho nahrazují tvary na *-ó* a na *-ú*. Toto geografické omezení z něj dělá tvar úžeji spjatý s *češtinou obecnou*, a proto je občas považován za méně vhodný v spisovných kontextech. Ne každá příručka ale rozlišuje mezi užitím 1. os. a os. 3.

**iii. Registr slova.** Zmiňuje se v různých příručkách o tom, že sémantika jednotlivých slov působí na jejich otevřenost vůči různým koncovkám. Trávníček např. zdůvodňuje frekvenci analogických koncovkách u sloves typu *kousat* tím, že vyjadřují každodenní fyzické děje, a proto jsou s nimi přípustné koncovky analogické. Novější příručky upozorňují, že u sloves vyšších stylů, obzvláště v 6. třídě, jsou analogické koncovky spíše pocíťovány jako nevhodné či příliš mluvené.

**iv. Registr textu.** Mluvnice nás upozorňují, že v textech vyšších registrů můžou být analogické tvary pocíťovány jako nevhodné, a rovněž pro texty všedního typu varují před některými příliš knižními tvary.

Na problematiku pronikání tzv. „hovorové“ tvary do spisovného jazyka se poukazuje v řadě prací, včetně empirického výzkumu v L. Hammer 1985 a 1993, K. Kravčíšínová a B. Bednářová 1968, J. Šonková 1995 a Ch. Townsend 1990. Na základě těchto nálezů se celá záležitost dostala do prací zaměřených také na přehodnocení spisovné normy, jako např. P. Sgall 1998, P. Sgall a J. Hronek 1992, P. Sgall et al. 1992, aj.

Tento výzkum spadl do dvou částí, které budeme konfrontovat. V první fázi (viz N. Bermel v tisku) jsme provedli analýzu příruček a zjistil rozložení tvarů v korpusu. Druhá, současná fáze představuje analýzu jednotlivých příkladů z korpusu.

## **2. Rozložení tvarů v ČNK**

Korpus SYN2000, který tvoří jádro Českého národního korpusu, je stamilionový reprezentativní korpus. Zásady a následky jeho složení jsou popsány jednak v příručce J. Kocka et al., eds. 2000, jednak v různých článcích, např. M. Šulc 2001, J. Králík 2001. Náměty pro jeho užívání při zkoumání morfologických jevů jsou v posledních letech zastoupeny v širším okruhu článků věnovaných práci s korpusem (viz např. F. Štícha 2001, T. Berger 2004, N. Bermel 2004, J. Šimandl 2004).

Potřebovali jsme zjistit rozsah variace jak pro slovesné tvary, tak pro jednotlivá slovesa. Výsledky zkoumání ukazují, že rozložení koncovek se jasně liší podle slovesné třídy.

Tabulka 2. Slovesné třídy a koncovky

Třída	Tvar	Doklady (%)	Distinktivní tvary (dokladů/tvar)	Distribuce v SYN2000
<b>I</b>	<b>vše</b>	<b>11 571 (100%)</b>	<b>376 (30.8 d/t)</b>	
	-ji	906 (7,8 %)	69 (13,1 d/t)	Nekonzistentní, více v první čtvrti.
	-ju	1 634 (14,1 %)	92 (17,8 d/t)	Nekonzistentní, více v první čtvrti.
	-jí	8 678 (75,0 %)	155 (60,0 d/t)	Nekonzistentní, více v druhé čtvrti.
	-jou	353 (3,1 %)	60 (5,9 d/t)	Nekonzistentní, nejvíce v první čtvrti, další ostrůvky později.
<b>V</b>	<b>vše</b>	<b>9 152 (100 %)</b>	<b>324 (28,2 d/t)</b>	
	-i	460 (5,0 %)	42 (11,0 d/t)	Jen trochu nekonzistentní, o něco více v první třetině.
	-u	2 629 (28,7 %)	104 (25,3 d/t)	Nekonzistentní, nejvíce v první čtvrti, později málo.
	-í	3 584 (39,2 %)	69 (51,9 d/t)	Jen trochu nekonzistentní, málo v první čtvrtině, jinak rovnoměrně rozdělené
	-ou	2 479 (27,1 %)	109 (22,7 d/t)	Konzistentní (ostrůvky všude)
<b>VI</b>	<b>vše</b>	<b>227 341 (100 %)</b>	<b>6 976 (32,6 d/t)</b>	
	-uji	26 307 (11,6 %)	1 515 (17,4 d/t)	Nekonzistentní: na začátku je mnoho, jinde výrazně méně.
	-uju	8 232 (3,6 %)	1 018 (8,1 d/t)	Nekonzistentní: mnoho na začátku, později jen menší ostrůvky.
	-ují	191 913 (84,4 %)	4 005 (47,9 d/t)	Nekonzistentní: po nízkém výskytu na začátku pár ostrůvků a pak rovnoměrná distribuce.
	-ujou	889 (0,4 %)	438 (2,0 d/t)	Velmi nekonzistentní: Valná většina výskytů hned na začátku, později jen ojedinělé ostrůvky.

Tento stav potvrzuje některé naše předpoklady. Jsou značné rozdíly mezi výskytu analogických tvarů v různých slovesných třídách. Analogické tvary 1. os. j. č. se vyskytují častěji než analogické tvary 3. os. mn. č. Data jsou ovšem poněkud zkruslena funkčním uplatněním těchto dvou tvarů v psaných textech: 1. os. se vyskytuje nejčastěji v citátech mluvené řeči, jinak je jeho užití řídké, a proto očekáváme, že bude více dokla-

dů analogických tvarů na *-u*. 3. os. mn. č. má oproti tomu širší uplatnění ve všech textových žánrech a kontextech a je tedy menší procento kontextů, do kterých se hodí jejich analogický tvar.

Tedy ale k dalším dvěma předpokladům: o vlivu registru sloves a registru textů. V Bermel (v tisku) jsme poukázali na náznaky o textovém rozložení těchto tvarů. V korpusu SYN2000 jsou texty archivovány tak, aby vyhledávky probíhaly vždy stejnou trasou, tj. vyhledávka po každé začne, proběhne a končí stejnými texty v pořadí, které respektuje textové žánry a zdroje. Zadáním dotazu o rozložení výskytů v korpusu obdržíme malý graf, který zobrazuje relativní frekvenci dokladů.

Toto pevné pořadí začíná literárními texty, pokračuje v jiných žánrech a končí publicistikou. Je proto pozoruhodné, že jsou značné rozdíly v rozložení tvarů, a že se tyto rozdíly zdají odpovídat hranicím v korpusu mezi texty literárními a neliterárními.<sup>2</sup>

V korpusu SYN2000 jsme vyhledali relevantní tvary 23 sloves, která celkem vzato poskytují sondu do různých lexikálních kategorií. Slovesa 1. a 5. třídy tvoří uzavřenou skupinu vysokofrekventovaných lexémů, jsou vesměs součástí základní slovní zásoby jazyka. Na první pohled nás může překvapit fakt, že jsou relativně málo zastupována v korpusu, ale to snad proto, že korpus obsahuje převážně texty formálního charakteru, ve kterých se vyskytují každodenní děje jen občasně. Z těchto skupin jsme vybrali slovesa prefigovaná i neprefigovaná: *česat, dokázat, kousat, namazat, napsat, vypít, výt, zabít, zakrýt*.<sup>3</sup>

V 6. třídě nás nejvíce zajímala stylistické rozvrstvení této skupiny, která je zcela nejfrekventovanější a nejproduktivnější. Slovesa *nakupovat, kupovat, milovat* a *zamilovat se* pocházejí z každodenního života. *Archivovat, absolvovat, citovat, reprezentovat, charakterizovat* a *organizovat* patří do vědecké a administrativní vrstvy jazyka; a *naservírovat, koncertovat* a *vytunelovat* představují slova běžně mluveného jazyka, která se často vyskytovala i v novinách druhé poloviny 90. let. Díky jeho uplatnění v různých ustálených společenských replikách a situacích tvoří sloveso *děkovat* zcela speciální případ.

### 3. Variace na úrovni lexikálních jednotek

Rozhodli jsme se zatím omezit na sondy do jednotlivých sloves a tímto způsobem odhalit některé problémy nebo nepřesnosti, které vznikají během hromadné analýzy dat uvedené v tabulce 2.

Pro 1. třídu jsme se zaměřovali na hlubší analýzu sloves na *-krýt* a *-bít*. Obecně pro slovesa 1. třídy panovaly analogické tvary v 1. os. j. č., kdežto pro 3. os. mn. č. převládaly tvary tradiční:

---

<sup>2</sup> Nový korpusový manažer Bonito, který se zpřístupnil až na závěr tohoto výzkumu, umožní jednoduché určení žánrových rozdílů sestavením subkorpusů podle textového typu, ve kterých pak spustíme relevantní vyhledávání.

<sup>3</sup> Slovesa prefigovaná byla vždy ověřena v porovnání se všemi tvary tohoto kořene, aby nedošlo ke zkreslení výsledků na základě jednoho slova.

Tabulka 3. Slovesa 1. třídy

kořen	1. sg. -u	1. sg. -i	3. pl. -ou	3. pl. -í
-krýt (-křejt)	22 (55 %) <sup>4</sup>	18 <sup>5</sup> (45 %)	5 (1 %)	538 (99 %)
-pít	319 (78 %)	92 (22 %)	58 (9 %)	588 (91 %)
-bít	258 (93 %)	19 (7 %)	94 (16 %)	485 (84 %)
-výť	7 (87 %)	1 (13 %)	9 (89 %)	1 (11 %)

Co se týče slovesa *krýt*, výzkum jednotlivých dokladů potvrzuje okrajovou povahu tvarů typu *kryjou*. Z pěti dokladů jeden je zavádějící (je to místní jméno *Skryja* v 7. pádě). Další, pod (1), je z písně a má se rýmovat s předchozím veršem. (2) je z dramatu Josefa Topola *Masopust*, které představuje mluvenou řeč. (3) a (4) mají hovorový ráz, stylizují psaný text do mluvené podoby. Obzvlášť výraz *krýt někomu záda* se hodí k užití hovorového tvaru.

1. <doc S|SON|1998|zalman>svou , ptáci pod křídly je <ukryjou> a pak v objetí s nimi odletí
2. <doc S|SCR|1968|masopust> Panebože , těch je ! Úplně <zakryjou> nebe . Kde se tu vzaly ?
3. <doc S|PUB|1998|mf980921> . I tak se náklady zřejmě <nepokryjou> úplně . My si ale myslíme
4. <doc S|PUB|1998|pr980702> jen kvůli tomu , že někomu <kryjou> záda , " vyjádřil se jeden

Srovnáme tyto příklady s neutrálními užitími tradičního tvaru *kryjí*. Doklady uvedené níže svědčí o širokém uplatnění tohoto tvaru (vybrané tvary jsou spíš typické než výjimečné). Všimáme si dokladů z prózy, představených tu Arnoštem Lustigem:

5. <doc S|NOV|1995|brizy> , on i ostatní , kteří ho <kryjí> . Mluvil o Tetovaném , že

Dále jsou vybrány doklady z vědecké literatury a populárních časopisů:

6. <doc S|SCI|199g|sloaslv> tradičními rámcovými označeními se <kryjí> náměty velice subtilní ;
7. <doc S|POP|199f|lekar05> případě dělat , a příznaky se <kryjí> s příznaky dřívějšího onemocnění

Dva doklady z novin představují mluvenou řeč a popis filmu, takže tento tvar se zdá být jak pro reprezentaci mluveného textu, tak pro psaný text dostačující.

8. <doc S|PUB|1997|mf970212> pojišťoven naši režii zdaleka <nepokryjí> , " řekl .</doc><doc>Učitelé stávkovali
9. <doc S|PUB|1999|film0199> prohledají . Chlapci se však <ukryjí> ve dvojitém dně a vyváznou

<sup>4</sup> Včetně 3 dokladů *-křejt*.

<sup>5</sup> Z nichž je 8 dokladů z jedné znělky.

Všimáme si tu spojení ryze spisovných tvarů, u kterých by tvar *kryjou* zněl neadekvátně: máme tu na mysli 7. pád mn. č. u (6) a (7) a tvar *kterí* u (5). Tu hraje roli jak registr textů, tak souhra jiných tvarů spojených v češtině s pojmem „spisovnosti“.

Ačkoli jsme tvrdili, že výraz *krýt někomu záda* se hodí k užití méně formálního tvaru *kryjou*, stojí za povšimnutí, že výraz *krýt někomu záda* objevujeme třikrát i s tvarem *kryjí*:

10. <doc S|PUB|1998|pr980627> penězi a mediální podporou , <kryjí> záda švédským odpůrcům elektráren
11. <doc S|PUB|1998|pr981023> ale ty si často vzájemně <kryjí> záda , takže nesplní původně
12. <doc S|PUB|1999|ln99088s> střílí góly , další dvě jí <kryjí> záda . V základní části hrál

Získáním seznamu nejfrekventovanějších kolokací se slovesy na *-kryjí* zjistíme, že slovo *záda* je páté nejčastější podle relativní frekvence a ze sémanticky významných slov jen *náklady* a *asistenční* mají vyšší absolutní frekvenci.<sup>6</sup> Tento fakt argumentuje pro širší uplatnění tradičního tvaru, který se zdá hodit i do kontextů hovorového rázu. Srovnáme to s uplatněním tvarů 1. os. j. č., kde je distribuce jiná. V dalších dokladech najdeme sprostější variantu tohoto výrazu:

13. <doc S|NOV|1997|marcin> Mluvení tě nic nestojí , Dicku ! Vždyť já ti tady <kreju> tvůj zadek ! " V tom měl samozřejmě pravdu . " Dobře
14. <doc S|NOV|1996|klima> nohy , a někdo jiný račič mu sednout na záda , která <přikryju> já svým pláštěm . " - Suk byl vyražen , tlustý jeden

Nenajdeme s tímto výrazem žádné doklady s tvarem *kryjí* a oba doklady s *kryju*, *kreju* jsou z vyprávění literárních postav. Na rozdíl od 3. os. mn. č. si tu všimáme koexistence analogického tvaru s jinými tvary spisovného rázu, jako např. tvar *kerá* u (13). Kolokace s tvary *kryju*, *kreju*, *kryjí*, *kryjí*, *kryjou* tedy celkově potvrzují závěry hromadné analýzy dat, že spisovné tvary 3. os. mn. č. na *-í* mají mnohem širší uplatnění, než jejich spisovné protějšky 1. os. mn. č. na *-i*.

Jinou distribuci tvarů najdeme u sloves s kořenem *-bít*, u kterého jsou zastoupeny analogické tvary o něco častěji. V 1. os. j. č. naprosto dominují tvarům tradičním, i v 3. os. mn. č. tvoří šestinu až čtvrtinu všech dokladů. Za to asi může silná expresivita sloves s tímto kořenem a tedy i vhodnost těchto tvarů pro citáty z mluvené řeči:

15. <doc S|NOV|1996|vykupne> jestli se odtamtud pohnou , <zabiju> je . " Flynn se naklonil
16. <doc S|PUB|1992|ln92200> Všechny si vás pamatuju . <Zabiju> vás a nebo vám utrhnu hlavu

Výraz *já tě zabiju* je zastoupen 13krát v korpusu a *zabiju tě* 15krát, většinou jako samostatná věta:

17. <doc S|NOV|1991|mirak> najednou mě to taky vzrušilo . <Zabiju> tě , řekl jsem , ale ona

---

<sup>6</sup> Při nastavení standardních parametrů. S nastavením většího rozsahu se do výsledků pletou další slova, která nejsou přímo spojená se slovesem, ale spíše svědčí o stylistickém a obsahovém charakteru textu.

18. <doc S|NOV|1994|top3> nemá . Kde má hadry . Já tě <zabiju> . Ne , polknul . Ne . To

Oproti tomu všechny čtyři doklady s kolokací *tě zabiji* tvoří součásti delších souvětí.<sup>7</sup>

19. <doc S|COL|1997|staprap> si můžeš být jist , že tě <zabiji> . ' A dotkl jsem se jeho  
20. <doc S|COL|1997|staprap> ji , bud' si jist , že tě <zabiji> . " I šel hvězdný chlapec  
21. <doc S|FAC|1996|orten> nenamlouvej jiného , nebo Tě <zabiji> . &lsqb; Kutná Hora &rsqb;

Navíc můžou představovat citáty z dopisů:

22. <doc S|PUB|1997|story97> herečce JODIE FOSTEROVÉ : " <Zabiji> pro vás prezidenta ! " Tímto

Rozložení těchto tvarů v korpusu jsme zjistili očíslováním uložených netříděných výsledků, které se ukládají podle žánrů zdrojů.

**Tabulka 4. Frekvence tvarů slovesa zabít**

tvar	doklady v beletrii	frekvence vzhledem k velikosti lit. části (15 %)	doklady v publicistice	frekvence vzhledem k velikosti pub. části (60 %)	poměr frekvencí beletrie: publicistika
zabiju	118	786,67	53	88,33	8,91:1
zabiji	4	26,67	3	5	5,33:1
zabijou	30	200,00	21	35,00	5,71:1
zabijí	39	260,00	94	156,67	1,60:1

Frekvence těchto je tedy vždy vyšší v literárních textech než v textech publicistických, podle údajů je tento nepoměr mezi texty literárními a neliterárními však nižší u tradičně spisovných tvarů než u tzv. tvarů „hovorových“.

Kolokace s těmito tvary mnoho nenapovídají, jelikož se spojují jen se zájmeny (tě, ho, je) a interpunkci.

Slovesa 6. třídy vykazují jasné preference pro tradiční tvary až na pár výjimek. V Bermel (v tisku) jsme odhadli, že se tu tvoří stupnice podle registru jednotlivých sloves, tj. že hlavní roli tu hraje náš třetí činitel (viz výše). Všední slovesa mají sklon k vyššímu výskytu analogických tvarů, kdežto slovesa ryze akademická či oficiální silně upřednostňují varianty tradiční. V tomto argumentu zařazení sloves *milovat*, *zamilovat se* a *naservírovat* do první skupiny odpovídá jejich emocionálně naladěné povaze. Jakmile se ale podíváme do jednotlivých dokladů, začíná se do té analýzy plést vliv kontextu a žánru. Přehled těchto tvarů je v tabulce 5.

<sup>7</sup> Není vůbec zastoupena kolokace *zabiji tě*.

**Tabulka 5. Slovesa 6. třídy.**

	1. sg. -u	1. sg. -i	3. pl. -ou	3. pl. -í
<i>naservírovat</i>	3 (100%)	0	0	6 (100%)
<i>-servírovat</i>	5 (50%)	5 (50%)	1 (2%)	62 (98%)
<i>zamilovat se</i>	15 (65%)	8 (35%)	3 (4%)	65 (96%)
<i>milovat</i>	814 (48%)	880 (52%)	29 (5%)	602 (95%)
<i>-milovat</i>	830 (48%)	892 (52%)	32 (5%)	677 (95%)

Jako příklad vezměme nejdříve sloveso *naservírovat*. Všimneme si, že se tvar *naservíruju* vyskytuje jenom v románech a povídkách. Nejsou doklady tvaru *naservíruji*, takže nemůžeme nic soudit o jeho možném uplatnění.

23. <doc S|COL|1992|prokleti> pohlédl pozorně na Ivanse .</doc><doc>"  
<Naservíruju> mu to , jak budu moct horké

V 3 os. mn. č. tvar *naservírují* je doložen i v obvyklých novinových článcích:

24. <doc S|PUB|1996|ln96292> přijímat jen to , co mi politici <naservírují> na náměstích . Navíc když

25. <doc S|PUB|1997|ln97090> distributory , kteří jim vše <naservírují> na stříbrném podnose , "

26. <doc S|PUB|1999|mf990213> že jim dědeček s babičkou <naservírují> k posteli vychlazené šampaňské

Doklady tvaru *naservírujou* chybí úplně, takže můžeme předpokládat, že tradiční tvar *naservírují* má neutrální uplatnění. Vidíme ho v kontextech formálních, ale nechybí úplně náznaky hovorovosti, jako např. fráze *naservírují na stříbrném podnose*.

Do stejné skupiny patří slovesa *milovat* a *zamilovat se*. Neprefigované sloveso je hojně doloženo v SYN2000, proto se dá udělat závěry na základě korpusových kolokací. Zadáním dotazu v rozsahu jednoho místa doprava a doleva zjistíme, že jsou malé ale nicméně zajímavé rozdíly v jejich kolokacích, uvedených podle relativní frekvence:

**Tabulka 6. Kolokace s tvary slovesa milovat.**

tvar	kolokace prvních 20 lemmat (interpunkce vynechána)
miluju	šileně, ty, Irena, strašně, smírně, dodnes, pořád, voda, sbírka, moc, někdo, skutečně, vlastně, zase, on, tvůj, stejně, tady, také, více
miluji	vroucně, &num;, dvojče, ty, novina, pořád, kino, tenhle, bůh, protože, přece, žena, on, velmi, více, tvůj, život, moc, dítě, muž
milujou	člověk, on, se <sup>8</sup>
milují	vášnivě, doopravdy, romantický, kočka, fotbal, Američan, bůh, sníh, zvíře, komunista, tolik, opravdu, příroda, svoboda, hudba, dítě, kdo, žena, divák, mladý

---

<sup>8</sup> Kvůli nízkému výskytu tohoto tvaru jsou jenom tři kolokace.

Rýsuje se tu částečné stylistické rozvrstvení tvarů spisovných a hovorových. Vysoká relativní frekvence lemmat *šileně*, *strašně*, *smírně*,<sup>9</sup> *skutečně* u hovorového tvaru 1. os. j. č. oproti *vroucně*, *novina*,<sup>10</sup> *kino*, *bůh*, *život* u spisovného tvaru svědčí o kontextech, do kterých tyto tvary spadají. Situace je méně jasná u 3. os. mn. č. pro velice nízký výskyt hovorového tvaru, které znemožňuje smysluplné hledání kolokací.

Sloveso *zamilovat se* má nižší frekvenci, a proto je přiměřenější se dívat ne na kolokace, ale přímo na doklady. Tvar *zamiluju* se uplatňuje v různých žánrech:

27. <doc S|NOV|1998|palom> přesvědčeně . " Proč je pro vás důležité ,  
jestli se já zase <zamiluju> . " " Chtěla bych vědět , jestli je to možné .  
" "
28. <doc S|PUB|1997|story97> stejně nepředpokládala , že se ještě někdy  
do někoho <zamiluju> a zároveň jsem chtěla něco zásadního ve svém  
životě

Nicméně je doloženo např. v publicistice výhradně v ženských časopisech *Vlasta* a *Story* a v novinových přílohách, kde najdeme obvykle delší interview s kulturními postavami. Oproti tomu je tvar *zamiluji se* doložen osmkrát:

29. <doc S|PUB|1997|story97> tajně nebo , jak mi v poradně říkali dis-  
krétně , <zamiluji> do nějaké spíše paní nežli slečny a ta láska je nějak

Je těžké ale rozlišovat mezi volbami *zamiluju se* a *zamiluji se*. I když tradiční tvar se uplatňuje v širším okruhu textů, nicméně jsou kontexty dost podobné těm, které vidíme u tvaru analogického.

V plurálu je tvar *zamilují se* zjevně základní oproti ojedinělým výskytům tvaru *zamilujou se*. Všimáme si značně rozdílných kontextů analogického tvaru *zamilujou se*:

30. <doc S|NOV|1990|prapor> prohlásil Bamza , " jak se dva <zamilujou>  
a von jí přivede do jinýho
31. <doc S|NOV|1993|svet> milionáře , kteří se do ní <zamilujou> a v ráj  
jí život promění
32. <doc S|PUB|1993|refl9315> Samozřejmě se do sebe hrozně <zamilu-  
jou> . Aby Jack Lily získal ,

(30) je stylizace obecněčeského textu, s náznaky mluvenosti jako *von*, *jinýho*. (31) je text spisovnější, jak jasně ukazuje tvar *kteří*. (32) je text hovorový, psaný ale neformální, což vyplývá ze slov jako *samozřejmě*, *hrozně*. SYN2000 nám těmito doklady poskytuje škálu kontextů, do kterých se analogický tvar hodí.

Tvar tradiční *zamilují se* naopak má užší škálu kontextů:

33. <doc S|FAC|1999|touha\_9> Protože když se zdejší dívky <zamilují> ,  
ztratí soudnost . Jistě
34. <doc S|PUB|1997|story97> pacientky do svých lékařů <zamilují> . A  
stávalo se to i plastickému
35. <doc S|NOV|1995|kantilen> rozdíl od války ) šileně <zamilují> v  
okamžiku , kdy získají

<sup>9</sup> Evidentně se tím rozumí i „záporný“ tvar *nesmírně*.

<sup>10</sup> Chybná lemmatizace pro *noviny*.

V (33) a (34) jsou slova spisovnějšího rázu jako *dívky* a *lékaři*, trochu hovorověji zní (35) se slovem *šíleně*.

## 9. Závěry

Projekt „Možnosti a meze české gramatiky ve světle ČNK“ se zaměřuje jednak na orientační sondy do problematiky těžení gramatických informací z ČNK, jednak na budování metodologie pro popis našich nálezů v gramatice velkého formátu. Výsledky našeho bádání přesvědčuje, že korpusová data mají co dodat do velké gramatiky, problematický však zůstává fakt, že korpus pro tento účel nabízí široký výběr dat a metod pro výklady, které z nutnosti budou stručné a obecné, nikoli podrobné. Otázkou tedy je, které z množství možných přístupů se hodí k výstižnému popisu nějakého jevu vhodnému do použitelné mluvnice.

Ve své analýze jsme konfrontovali již popsany přehled korpusových dat s několika kolokacemi jednotlivých tvarů a podrobnou analýzou dokladů v kontextu. Přehled dat nám poskytuje návod k použitelnosti různých tvarů a popisuje pravděpodobnost výskytu jednoho či druhého tvaru v korpusu. Kontextová analýza a zkoumání kolokací naopak zdůrazňují možnosti volby a souhru jiných textových faktorů, např. frazeologismů, přítomnosti přímé řeči a vlivu jiných jevů spisovného a mluveného jazyka.

Tyto přístupy se ovšem vzájemně nevylučují. Je ale na místě se zeptat, kterému z nich dáme přednost ve stručné analýze, či jak je můžeme integrovat do jednotného popisu jevu. Je pořád účelné se ptát, jak daleko do detailů a jednotlivých dokladů musíme dojít, abychom získali přehled užití, který je víceméně úplný, ale zároveň i průhledný.

Pro dané jevy se zdá, že data z korpusu SYN2000 zdůrazňují spíš možnosti volby nežli meze užívání. Celkový přehled v číslech na data týkající se každého slovesa a jejich podrobné srovnání napovídá o tom, že dominantní roli hraje registr slova. Jinak řečeno, jsou vzory, podle kterého určujeme nejfrekventovanější koncovky a slovesa se do těchto vzorů zařazují podle sémantických kritérií. Detailnější analýza vysvětluje, jak kontext a žánr přispívají k volbě jednoho či druhého tvaru. Pro morfologické jevy tedy navrhuje, že pečlivě zvolené kvantitativní vyhledávky tvoří jádro dat, které vedou eventuelně k stručnému popisu daného jevu, ale jsou nutně zmírněné, doplněné a obohacené kvalitativní analýzou kolokací, frekvencí a kontextů.

## Literatura

- BERGER, T. Osud nepravidelných substantiv v češtině. In: Karlík, P. - Hladká, Z. (eds.): *Čeština - univerzália a specifika*, 5, 2004, Praha: Lidové noviny, s. 246-254.
- BERMEL, N.: *Register variation and language standards in Czech*. Munich: Lincom Europa 2000.
- V *korpusu* nebo v *korpusu*? Co nám řekne (a neřekne) ČNK o morfologické variaci v tvarech lokálu. In: Karlík, P. - Hladká, Z. (eds.): *Čeština - univerzália a specifika*, 5, 2004, Praha: Lidové noviny, s. 163-171.

- Standard and standardized verb forms in the Czech National Corpus. In: Blatná, R.–Petkevič, V. (eds.), *Jazyky a jazykověda*. Praha: v tisku.
- ČERMÁK, F. 1987. Relations of spoken and written Czech. *Wiener slawistischer Almanach*, 20, 1987, s. 133-150.
- Spoken Czech. In: Eckert, E. (ed.): *Varieties of Czech*. Atlanta: Rodopi 1993, s. 27-41.
- ČESKÝ NÁRODNÍ KORPUS - SYN2000. Praha: Ústav Českého národního korpusu FF UK. <<http://ucnk.ff.cuni.cz>>.
- HAMMER, L. . *Prague Colloquial Czech: A case study in code switching*. Indiana University 1985: nepublikovaná disertace.
- The function of code switching in Prague Colloquial Czech. In: Eckert, E. (ed.): *Varieties of Czech*. Atlanta: Rodopi 1993, s. 63-78.
- HAVRÁNEK, B. - JEDLIČKA, A.: *Česká mluvnice*. Praha: Státní pedagogické nakladatelství 1963.
- Stručná mluvnice česká*. 26. vydání. Praha: Fortuna 1998.
- KOCEK, J. - KOPŘIVOVÁ, M. - KUČERA, K. (EDS.): *Český národní korpus: úvod a příručka uživatele*. Praha: ÚČNK FF UK 2000.
- KRÁLÍK, J.: Vyvážení zdrojů Synchronního korpusu češtiny SYN2000. *Slovo a slovesnost*, 62, 2001, s. 38-53
- KRAVČIŠINOVÁ, K. - BEDNÁŘOVÁ, B. 1968. Z výzkumu běžné mluvené češtiny. *Slavica Pragensia*, 10, 1968, s. 305-320.
- KUČERA, K.: The Czech National Corpus: Principles, design, and results. *Literary and linguistic computing*, 17 (2), 2002, s. 245-257.
- MLUVNICE ČEŠTINY (= AMČ). Díl 2: Tvarosloví. Praha: Academia 1985. .
- PŘÍRUČNÍ MLUVNICE ČEŠTINY (= PMČ). Praha: Nakladatelství Lidové noviny 1995.
- SGALL, P.: Neochuzujeme spisovnou češtinu. *Český jazyk a literatura*, 49, 1998, s. 29-35
- SGALL, P. – HRONEK, J.: Čeština bez příkras. Praha: Nakladatelství H&H 1992.
- SGALL, P. – HRONEK, J. – STICH, A. – HORECKÝ, J.: Variation in Language: Code Switching in Czech as a Challenge for Sociolinguistics. Amsterdam a Filadelfie: John Benjamins 1992.
- ŠIMANDL, J. Korpus jako (jediný) zdroj? In: Karlík, P. - Hladká, Z. (eds.): *Čeština - univerzálie a specifika*, 5, 2004, Praha: Lidové noviny, s. 50-57.

ŠONKOVÁ, J.: *Lingvistické zpracování jazyka na počítači: morfologie mluvené češtiny*.

Filozofická fakulta, Univerzita Karlova: Ph.D, 1995.

ŠTÍCHA, F.: Kritéria dramatickosti (korpus jako argument a inspirace). *Slovo a slovesnost*, 62, 2001, s. 161-165.

ŠULC, M.: Tématická reprezentativnost korpusů. *Slovo a slovesnost*, 62, 2001, s. 53-61.

TOWNSEND, CH.: *A Grammar of Spoken Prague Czech*. Columbus, OH: Slavica 1990.

TRÁVNÍČEK, F.: *Stručná mluvnice česká*. Praha: František Borový 1941.

—*Mluvnice spisovné češtiny. Část I: Hláskosloví – Tvoření slov – Tvarosloví*. Praha: Slovanské nakladatelství 1951.