

Kolokační analýza českých rodných jmen z diachronní perspektivy¹

Keywords: anthroponyms, personal names, first names, Czech National Corpus, diachronic analysis, hypocorisms, collocation analysis

Klíčová slova: antroponyma, osobní jména, křestní jména, Český národní korpus, diachronní analýza, hypokoristika, kolokační analýza

Abstract

The study deals with the most common Czech first names in a developmental/diachronic perspective, (the 1993–2013 period). The material for the frequency analysis (always the three most frequent female and male first names for the years 1993, 1998, 2003, 2008, and 2013) is drawn from the data of the Ministry of the Interior of the Czech Republic and the Czech Statistical Office. We will subject the obtained set of personal names to a collocation analysis in the field of journalism of the Czech National Corpus SYN, version 11. The first goal is to analyse the socio-cultural contexts associated with personal names. The second goal is to find out in which thematic areas of the field of journalism the basic forms of the names and their hypocoristic forms are used, based on the comparison of the collocations of the basic forms and those of their hypocoristic forms.

Předkládaná studie se zabývá kolokační analýzou tří nejčastějších českých rodných jmen v diachronní perspektivě, konkrétně v letech 1993, 1998, 2003, 2008 a 2013. Data pochází z údajů Ministerstva vnitra České republiky a Českého statistického úřadu. Jako nástroj pro analýzu dat jsme využili Český národní korpus SYN, verze 11; zaměřili jsme se na pět nejčastějších kolokací zvolených jmen v publicis-

¹ Studie vznikla s podporou grantového projektu SGS s názvem *OnomOs – Ostravský korpus vlastních jmen* (reg. č. SGS02/FF/2023, řešeno na Filozofické fakultě Ostravské univerzity).

tických textech. Výzkum ukázal, že základní podoby rodných jmen se uplatňují převážně v oblastech politiky a státních institucí. Jejich hypokoristické varianty, které se nevyznačují vysokými frekvencemi, jsou u žen spojovány s odkazem na mladistvé, u mužů pak s oblastí zábavy, umění a sportu.

1. Úvod

Předmětem zájmu předkládané studie je korpusová diachronní analýza českých rodných jmen, která mají v České republice k příslušnému roku (období 1993–2013) největší počet žijících nositelů. Jedná se o tři ženská jména a jejich hypokoristické podoby a tři mužská jména a jejich hypokoristické podoby. Z vymezeného období jsme se zaměřili na konkrétní roky, jimiž jsou 1993, 1998, 2003, 2008 a 2013.² Vybraná antroponyma jsme podrobili kolokační analýze na publicistických textech *Českého národního korpusu* (dále jen ČNK), *SYN verze II* (Křen a kol. 2022; dále jen SYN v11). Vytyčili jsme si přitom dva základní cíle, a to analýzu společensko-kulturního kontextu pojícího se s těmito proprii a srovnání tematických oblastí publicistiky, v nichž se nejčastěji používá základní podoba rodného jména se situacemi, kdy je preferována jeho hypokoristická forma. Předpokládali jsme, že nejčastějšími kolokáty u základních podob zkoumaných antroponym budou výrazy odkazující na politiku, oficiální funkce a instituce. U hypokoristických variant jsme očekávali nejčastější kolokace v prostředí zábavy, popř. kultury obecně.

2. Vlastní jména a korpusové badání

Obecně lze konstatovat, že korpus nabízí zcela nové možnosti lingvistického zkoumání. Můžeme díky němu pracovat se značným množstvím dat, ta je možné měřit, komparovat a intersubjektivně ana-

² Zkoumaný materiál jsme omezili podle dat v příslušných letech s tím, že jsme jako počáteční zvolili rok 1993, který se váže ke vzniku České republiky a následně jsme pokračovali vždy po pětiletí. Obdržený statistický materiál končí rokem 2015.

lyzovat. Nejen v zahraničí, ale i v českém prostředí je korpus využíván stále více i pro analýzu proprii. V současné české onomastice se lingvisté zaměřují především na čtyři oblasti výzkumu vlastních jmen (viz David, Klemensová, Místecký 2022; přehledově David, Místecký 2023). Jedná se o analýzu proprii pomocí korpusových nástrojů – kolokace, frekvence, klíčová slova, tematická slova atd. (např. Drkošová, Místecký 2022; David, Klemensová, Místecký (eds.) 2022, s. 126–143), dále o popis gramatických charakteristik proprii (např. Tušková 2011) či jejich detekci a desambiguaci (např. Osolsobě, Žižková 2021) a nejnověji jejich stylistické a stylometrické vlastnosti (např. Cvrček a kol. 2020). Největší překážkou ve výzkumu vlastních jmen v českých korpusech je to, že ve většině z nich propria nejsou anotována, což souvisí rovněž s obtížnou formální klasifikací samotného vlastního jména jako lexikální skupiny či jeho druhů (otázky transonymizace a homonymie, srov. Tušková 2019).

3. Východiska a metodologie zkoumání

Co se týká východisek zkoumání, využili jsme data pocházející z databázi Českého statistického úřadu (dále jen ČSÚ) a Ministerstva vnitra České republiky. Jelikož seznamy nejčastějších rodných jmen nejsou v dnešní době již veřejně dostupné, požádali jsme o ně Jiřího Trexlera, pracovníka ČSÚ, který v roce 2016 provedl statistický výzkum zaměřený na počty nositelů rodných jmen v jednotlivých letech. Z tohoto seznamu jsme vybrali rodná jména, jež měla nejvyšší počet žijících nositelů v příslušných letech (1993, 1998, 2003, 2008, 2013). V případě ženských jmen se jedná o *Janu*, *Marii*, *Evu* a *Hanu*, u mužských jmen pak o *Jiřího*, *Petra* a *Jana*. Hypokoristické podoby těchto jmen jsme čerpali z publikace Miloslavy Knappové *Jak se bude vaše dítě jmenovat?* (2017), která ovšem nezohledňuje frekventovanost daných forem. Pro každé rodné jméno jsme proto vybrali jeho nejčastější hypokoristikum podle frekvenčního průzkumu v ČNK. Pro ženská jména to jsou podoby *Janička*, *Maruška*, *Evička* a *Hanka* a pro mužská jména *Jirka*, *Pěta*, *Honza*. Zmíněná hypokoristika

jsme zadávali pomocí dotazu v jazyce CQL typu [lemma="Maruška"&tag="..F.*"].

Pomocí korpusového rozhraní *KonText* jsme provedli kolokační analýzu vybraných rodných jmen a jejich hypokoristických forem. Kolokační analýza se zaměřuje na ustálená sousloví, tj. slova, která se nejčastěji pojí se zkoumaným výrazem (viz Brezina, McEnery, Wattam 2015, s. 139–173). Zvolili jsme tento korpus, protože jsme se chtěli zaměřit na publicistiku,³ která nejvíce odráží živý jazyk a je pravděpodobně nejvíce čtena (Křen a kol. 2022). Byli jsme si zároveň vědomi, že se v tomto korpusu vyskytují různé frekvence, pokud se jedná o srovnání základních podob jmen s příslušnými domácími formami. Například jméno *Jiří* v roce 1993 má 7187 výskytů (viz tab. 9), ale hypokoristikum *Jirka* má pouze 171 výskytů (viz tab. 10), což způsobuje zároveň to, že u hypokoristických variant se objevují slabší kolokační vazby. Uvědomujeme si také, že složení zvoleného subkorpusu obsahuje různé typy periodik, která ovlivňují výskyt jmen, což neodráží fungování jmen v celé populaci. Nelze proto stanovit závazné závěry – v našem výzkumu jsme se chtěli spíše zaměřit na trendy a příznaky spojené se zkoumanými jmény, a to i v diachronní perspektivě, tedy ve sledovaném dvacetiletí.

Při analyzování příslušného jména jsme pracovali s dotazem typu [lemma="Marie"&tag="..F.*"]. V rámci kolokační analýzy jsme pracovali s kolokačním oknem od -5 do 5 (z důvodu požadovaného širšího kontextu) a zohledňovali jsme pouze kolokáty a kolokace, které se objevily s minimální frekvencí 3, přičemž jsme se zaměřili na pět nejsilnějších kolokátů. Jako asociační míru jsme zvolili *logDice* (dále jen ID), která vyhledává exkluzivní kolokace nezávislé na frekvenci; zároveň není obtížná pro lingvistickou interpretaci (Brezina 2018; Cvrček 2019; Drkošová, Místecký 2022, s. 102).

³ Jedná se o psanou publicistiku, což na jedné straně omezuje používání hypokoristik typických pro děti a mládí (Pleskalová 2011, Knappová 2017), ale současně nabízí možnost sledovat případy, kdy jsou tyto příznakové formy použity.

4. Kolokační analýza – ženská rodná jména

V následující části se soustředíme na kolokační analýzu tří nejfrekventovanějších ženských jmen v letech 1993, 1998, 2003, 2008 a 2013. Jedná se o základní podoby osobních jmen a jejich hypokoristické formy, jejichž relativní korpusovou frekvencí⁴ prezentují tabulky č. 1 a 2; jména jsou přitom řazena dle dostupných dat ČSÚ. Jednotlivé výsledky kolokační analýzy v korpusu SYN v11 jsou shrnuty v tabulkách č. 3 až 10.

Tabulka 1. Tři nejčastější ženská jména – základní podoba

SYN v11, publicistika									
1993		1998		2003		2008		2013	
Jméno	i.p.m.	Jméno	i.p.m.	Jméno	i.p.m.	Jméno	i.p.m.	Jméno	i.p.m.
Marie	62,82	Jana	70,89	Jana	102,93	Jana	37,46	Jana	147,42
Jana	52,86	Marie	90,6	Marie	107,67	Marie	101,86	Marie	110,75
Eva	44,26	Eva	65,12	Eva	83,05	Hana	67,76	Hana	69,14

Tabulka 2. Tři nejčastější ženská jména – hypokoristika

SYN v11, publicistika									
1993		1998		2003		2008		2013	
Jméno	i.p.m.	Jméno	i.p.m.	Jméno	i.p.m.	Jméno	i.p.m.	Jméno	i.p.m.
Maruška	1,9	Janička	0,4	Janička	0,38	Janička	0,59	Janička	0,94
Janička	0,06	Maruška	1,23	Maruška	1,97	Maruška	2,3	Maruška	3,79
Evička	1,01	Evička	0,25	Evička	0,15	Hanka	6,26	Hanka	7,93

4.1 Kolokační analýza prvního nejčastějšího ženského jména

4.1.1 Základní podoba rodných jmen

Analýza provedená v publicistických textech korpusu SYN v11 ukázala, že pět prvních kolokátů zvolených ženských jmen jsou

⁴ Zvolili jsme jednotku i.p.m., která uvádí počet výskytů na milion pozic (Kovaříková 2021).

příjmení. Základní podoby nejfrekventovanějších proprií v letech 1993 (*Marie*), 2008 (*Jana*) a 2013 (*Jana*) se očekávatelně pojí nejčastěji s příjmeními osob zaměstnaných v různých státních institucích, např. *Decreux* (ředitelka Francouzského ústavu pro výzkum ve společenských vědách, tab. 3: 1993), *Ponděličková* (mluvčí havířovského magistrátu, tab. 3: 2008), *Kroutilová* (jihlavská policejní mluvčí, tab. 3: 2013). V roce 1998 se rodné jméno *Jana* vyskytuje nejčastěji v kontextu sportovkyň, např. *Novotná* (tenistka, tab. 3: 1998), *Šaldová* (běžkyně na lyžích, tab. 3: 1998). Výrazy, které obklopují nejfrekventovanější rodné ženské jméno v roce 2003 (*Jana*), zase odkazují na oblast publicistiky, např. *Havliigerová* (hospodářská novinářka, tab. 3: 2003), *Kasalová* (programová ředitelka TV Prima, tab. 3: 2003). V jednom případě, tj. u kolokátu *Andretti* vyskytujícího se v roce 1993 (tab. 3: 1993), se jedná o chybné korpusové značkování – výskyty se týkají italsko-amerického pilota Formule 1 *Maria Andrettiho*. V korpuse se totiž příjmení *Andretti* pojí s lemmatem ženského rodného jména *Marie*, správně by mělo být označeno jako substantivum mužského rodu (podle jména *Mario*).

Tabulka 3. Kolokační analýza prvního nejčastějšího ženského jména – základní podoby

SYN v11, publicistika									
1993 (<i>Marie</i>)		1998 (<i>Jana</i>)		2003 (<i>Jana</i>)		2008 (<i>Jana</i>)		2013 (<i>Jana</i>)	
Kolokát	ID	Kolokát	ID	Kolokát	ID	Kolokát	ID	Kolokát	ID
<i>Stiborová</i>	8,29	<i>Vargová</i>	8,96	<i>Havliigerová</i>	8,78	<i>Bartíková</i>	9,14	<i>Nagyová</i>	10,27
<i>Strubová</i>	7,51	<i>Novotná</i>	8,94	<i>Hybášková</i>	8,61	<i>Ponděličková</i>	8,69	<i>Bobošíková</i>	10,05
<i>Brůčková</i>	7,37	<i>Klusáková</i>	8,39	<i>Kasalová</i>	8,14	<i>Tomková</i>	8,51	<i>Kroutilová</i>	9,18
<i>Ducreux</i>	7,23	<i>Marcová</i>	8,21	<i>Kameníková</i>	8,02	<i>Bobošíková</i>	8,40	<i>Fischerová</i>	8,49
<i>Andretti</i>	7,10	<i>Šaldová</i>	7,90	<i>Bobošíková</i>	7,97	<i>Šípková</i>	8,30	<i>Macalíková</i>	8,19

4.1.2 Hypokoristika

V letech 1993, 2003, 2008 a 2013 je v publicistice korpusu SYN v11 dominantní spojení sledovaných hypokoristických forem *Maruška* a *Janička* s výrazy, které odkazují na děti nebo lidi charakteristické svým mládím, např. *Kudeřiková* (mladá protinacistická odbojářka, tab. 4: 1993), *holčička* (devítiletá oběť násilí, 2003), *prdelka* (přezdívka dítěte, tab. 4: 2003), *Císařová* (účastnice dětské soutěže „Princezna Berounského deníku“, tab. 4: 2013). V roce 1998 je nejsilnějším kolokátem příjmení *Kopecká*, které patří české herečce (tab. 4: 1998). Rovněž další kolokáty jmen pocházejících z tohoto období se vážou ke kontextu filmového umění – *vyprávět a moc* (tab. 4: 1998).

Tabulka 4. Kolokační analýza prvního nejčastějšího ženského jména – hypokoristika

SYN v11, publicistika									
1993 (<i>Maruška</i>)		1998 (<i>Janička</i>)		2003 (<i>Janička</i>)		2008 (<i>Janička</i>)		2013 (<i>Janička</i>)	
Kolokát	ID	Kolokát	ID	Kolokát	ID	Kolokát	ID	Kolokát	ID
<i>Kudeřiková</i>	6,11	<i>Kopecká</i>	2,73	<i>devítiletý</i>	1,57	<i>prdelka</i>	6,62	<i>Štochllová</i>	7,23
<i>valčík</i>	3,18	<i>vyprávět</i>	-2,15	<i>holčička</i>	-0,13	<i>řečnění</i>	5,24	<i>Císařová</i>	7,01
<i>Maruška</i>	2,69	<i>moc</i>	-4,58	<i>maminka</i>	-2,1	<i>pošeptat</i>	4,38	<i>Vildová</i>	6,56
<i>někdo</i>	-4,25	!	-5,08	<i>smrt</i>	-3,08	<i>visící</i>	3,64	<i>Švestková</i>	6,29
...	-4,76	<i>malý</i>	-5,37	<i>vidět</i>	-5,01	<i>Žilková</i>	3,57	<i>Daník</i>	6,18

4.2 Kolokační analýza druhého nejčastějšího ženského jména

4.2.1 Základní podoba rodných jmen

V případě základních podob druhého nejfrekventovanějšího ženského jména (*Jana*, *Marie*) analýza ukázala, že nejčastějšími kolokáty jsou opět příjmení, ale také související substantiva. Ze zmíněných substantiv jsou to *panna* a *nanebevzeť*, která se neodmyslitelně pojí s náboženskou tematikou, stejně jako samotné proprium *Marie* (tab. 5: 1998, 2003, 2008, 2013). Dále jsou druhá nejčastější antroponyma v letech 1993 (*Jana*), 2003 (*Marie*), 2008 (*Marie*) a 2013 (*Ma-*

rie) příjmení, která odkazují na osoby, jež působí ve státních institucích, např. *Foltinová* (ředitelka odboru hygieny Ministerstva vnitra, tab. 5: 1993), *Bohatá* (zastupující generální ředitelka Eurostatu, tab. 5: 2003), *Štrbáková* (olomoucká policejní mluvčí, tab. 5: 2013). V neposlední řadě se také rodná jména pojí s oblastí umění a literatury. Jsou to například kolokáty *Štroblová* (česká překladatelka a spisovatelka, tab. 5: 1993), *Uchytlová* (sochařka, tab. 5: 1998) a *Rottrová* (zpěvačka, tab. 5: 2008, 2013). Značné zastoupení mají rovněž kolokáty poukazující na politickou oblast. Jsou to: *Noveská* (poslankyně, tab. 5: 1998), *Součková* (ministryně zdravotnictví, tab. 5: 2003) a *Benešová* (státní zástupkyně, tab. 5: 2013). V minimálních počtech se pak vyskytovaly kolokáty pojící se s oblastmi žurnalistika, historie a zábava.

Tabulka 5. Kolokační analýza druhého nejčastějšího ženského jména – základní podoby

SYN v11, publicistika									
1993 (Jana)		1998 (Marie)		2003 (Marie)		2008 (Marie)		2013 (Marie)	
Kolokát	ID	Kolokát	ID	Kolokát	ID	Kolokát	ID	Kolokát	ID
<i>Klusáková</i>	8,13	<i>panna</i>	8,96	<i>Součková</i>	10,56	<i>panna</i>	9,72	<i>Štrbáková</i>	10,14
<i>Rybská</i>	7,77	<i>nanebevzetí</i>	8,94	<i>nanebevzetí</i>	8,9	<i>nanebevzetí</i>	9,62	<i>panna</i>	9,78
<i>Štroblová</i>	7,48	<i>Uchytlová</i>	8,39	<i>panna</i>	8,7	<i>Masaříková</i>	8,35	<i>nanebevzetí</i>	9,52
<i>Ramešová</i>	7,46	<i>Plotěná</i>	8,21	<i>Bohatá</i>	8,54	<i>Rottrová</i>	8,21	<i>Rottrová</i>	8,56
<i>Foltinová</i>	7,46	<i>Noveská</i>	7,90	<i>Masaříková</i>	8,39	<i>Terezie</i>	7,93	<i>Benešová</i>	8,56

4.2.2 Hypokoristika

V případě hypokoristických podob druhého nejčastějšího ženského jména nebylo možné provést kolokační analýzu propria *Janička* (1993), jelikož tato podoba měla v korpusu pouze jeden výskyt. Analyzovali jsme proto jen hypokoristikum *Maruška*, jehož nejčastější kolokáty (1998, 2003, 2008 a 2013) odkazovaly na děti a mladistvé,

např. *dvojčátko* (jedno z dvojčat zpěvačky Báry Basikové, tab. 6: 2003), *Klumparová* (dítě mateřské školy, tab. 6: 2013). Dalším oblastí bylo umění a literatura, např. *měsíčky* (pohádka *O Dvanácti měsíčkách*, zde postava Marušky, tab. 6: 2013). Objevil se také odkaz na sport, a to v podobě kolokátů *RK0201* (sportovní výsledky *Marušky Jánské*, tab. 6: 2008) a *RK0197* (sportovní výsledky *Marušky Foglové*, tab. 6: 2008). I v případě této analýzy jsme objevili chybné tagování, a to u kolokátů *Knietel*, *Lafata* (tab. 6: 1998), *Vyskoč* a *Kordík* (tab. 6: 2003). Proprium *Maruška* v korpusu bylo chybně považováno za podobu ženského jména, jedná se ovšem o mužské příjmení (*Pavel Maruška* – fotbalista). Uvedené kolokáty jsou pak příjmení jeho spoluhráčů.

Tabulka 6. Kolokační analýza druhého nejčastějšího ženského jména – hypokoristika

SYN v11, publicistika									
1993 (Janička)		1998 (Maruška)		2003 (Maruška)		2008 (Maruška)		2013 (Maruška)	
Kolokát	ID	Kolokát	ID	Kolokát	ID	Kolokát	ID	Kolokát	ID
–	–	<i>Knietel</i>	7,7	<i>Kudeříková</i>	8,47	<i>Kudeříková</i>	9,01	<i>Kudeříková</i>	7,35
–	–	<i>Kudeříková</i>	6,42	<i>hrošice</i>	8,28	<i>Zachovalová</i>	7,81	<i>Kovalecká</i>	6,58
–	–	<i>Rottrová</i>	3,45	<i>Vyskoč</i>	7,34	<i>RK0201</i>	7,65	<i>Klumparová</i>	6,44
–	–	<i>Lafata</i>	2,42	<i>Kordík</i>	5,2	<i>RK0197</i>	7,65	<i>Trtková</i>	6,26
–	–	<i>stařenka</i>	1,94	<i>dvojčátko</i>	4,06	<i>Hustáková</i>	7,42	<i>měsíčky</i>	6,17

4.3 Kolokační analýza třetího nejčastějšího ženského jména

4.3.1 Základní podoba rodných jmen

Třetí nejčastější jméno, které jsme v korpusu SYN v11 podrobili analýze, bylo v letech 1993, 1998 a 2003 jméno *Eva* a v letech 2008 a 2013 jméno *Hana*. Analýzou jsme opět potvrdili, že nejčastějšími kolokáty pojícími se se základními formami jmen, jsou příjmení. Základní podoba výše uvedených rodných jmen se nejčastěji v roce

1993 objevovala v oblastech žurnalistiky, např. kolokáty *Bobůrková* (novinářka, tab. 7: 1993) nebo *VY MY* (sekce časopisu *Reflex*, redaktorka *Eva* Valentová, tab. 7: 1993). V roce 1998 dominovalo sportovní prostředí, např. *Háková* (reprezentantka v biatlonu, tab. 7: 1998) nebo *Celbová* (volejbalistka, tab. 7: 1998). Pro rok 2003, 2008 a 2013 byly charakteristické odkazy na zaměstnance státních institucí, např. *Miklíková* (policejní mluvčí, tab. 7: 2003) nebo *Kroftová* (policejní mluvčí, tab. 7: 2013).

Tabulka 7. Kolokační analýza třetího nejčastějšího ženského jména – základní podoby

SYN v11, publicistika									
1993 (<i>Eva</i>)		1998 (<i>Eva</i>)		2003 (<i>Eva</i>)		2008 (<i>Hana</i>)		2013 (<i>Hana</i>)	
Kolokát	ID	Kolokát	ID	Kolokát	ID	Kolokát	ID	Kolokát	ID
<i>VY</i>	9,95	<i>Háková</i>	8,89	<i>Miklíková</i>	9,83	<i>Bašová</i>	9,08	<i>Burianová</i>	9,14
<i>MY</i>	9,16	<i>Celbová</i>	8,47	<i>Plíhalová</i>	9,22	<i>Kotková</i>	9,06	<i>Zagorová</i>	8,95
<i>Štěpančíková</i>	8,23	<i>Urbanová</i>	8,45	<i>Tomšovicová</i>	8,72	<i>Hegerová</i>	9,02	<i>Szotkowská</i>	8,92
<i>Bobůrková</i>	8,19	<i>Perónová</i>	8,45	<i>Holubová</i>	8,51	<i>Moudrá</i>	8,87	<i>Moudrá</i>	8,82
<i>Olmerová</i>	8,14	<i>Brodská</i>	8,44	<i>Kmentová</i>	8,40	<i>Pohanová</i>	8,86	<i>Kroftová</i>	8,68

4.3.2 Hypokoristika

V letech 1993, 1998 a 2003 se hypokoristická varianta *Evička* objevila zejména s odkazem na děti nebo mladistvé, např. *dcera* (studentka umělecké průmyslovky, tab. 8: 1998). Asociace tohoto jména a kolokátu byla velmi slabá. V letech 2008 i 2013 dominovala oblast zábavy, např. *Mašliková* (modelka, tab. 8: 2008), *Kynychová* (provozní fitness klubu, tab. 8: 2008, 2013) a *Panka* (travesti skupina Hany Panky, tab. 8: 2013). Dále se v těchto obdobích objevil kolokát *Rabenhauptová* směřující do oblasti politiky (místostarostka města Prahatice, tab. 8: 2008, 2013). Vyskytl se zde velmi zajímavý rozdíl mezi kolokáty jmen *Evička* a *Hanka*, což zřejmě souvisí s jejich rozdílným chápáním. Hypokoristikum *Evička*, je asociováno s dětmi,

blízkými, příbuznými; *Hanka* je běžnou fungující zdobnělinou. U jména *Evička* se neobjevují žádná příjmení, ale spíše odkazy na rodinné vztahy (*dcera*), a osobní identitu (*já*).

Tabulka 8. Kolokační analýza třetího nejčastějšího ženského jména – hypokoristika

SYN v11, publicistika									
1993 (<i>Evička</i>)		1998 (<i>Evička</i>)		2003 (<i>Evička</i>)		2008 (<i>Hanka</i>)		2013 (<i>Hanka</i>)	
Kolokát	ID	Kolokát	ID	Kolokát	ID	Kolokát	ID	Kolokát	ID
<i>rád</i>	-4,68	<i>dcera</i>	-2,42	<i>já</i>	-6,97	<i>Rabenhauptová</i>	9,13	<i>Panka</i>	10,11
<i>já</i>	-6,97	<i>řici</i>	-6,23	<i>už</i>	-7,21	<i>Kynychová</i>	8,95	<i>Kynychová</i>	9,12
:	-7,45	<i>nový</i>	-6,68	<i>mít</i>	-8,15	<i>Mašliková</i>	8,77	<i>Tonča</i>	9,02
<i>on</i>	-8,07	<i>když</i>	-6,69	<i>o</i>	-8,37	<i>Smržová</i>	8,38	<i>Michopulu</i>	8,01
<i>že</i>	-8,68	<i>já</i>	-6,97	<i>ten</i>	-8,67	<i>Pressburg</i>	7,46	<i>Rabenhauptová</i>	7,84

5 Kolokační analýza – mužská jména

Ve druhé části výzkumu jsme se soustředili na kolokáty tří nejčastějších mužských jmen pocházejících z let 1993, 1998, 2003, 2008 a 2013 (viz tabulky č. 9 až 18). Připomeňme, že kolokační analýzu základních podob zkoumaných antroponym a jejich hypokoristik jsme provedli pomocí nástroje *KonText*, a to v publicistice korpusu *SYN v11*. Přehled nejfrekventovanějších rodných jmen a jejich absolutní korpusovou frekvenci představují tabulky č. 9 a 10.

Tabulka 9. Tři nejčastější mužská jména – základní podoba

SYN v11, publicistika									
1993		1998		2003		2008		2013	
Jméno	i.p.m.	Jméno	i.p.m.	Jméno	i.p.m.	Jméno	i.p.m.	Jméno	i.p.m.
<i>Jiří</i>	426,36	<i>Jiří</i>	459,31	<i>Jiří</i>	456,18	<i>Jiří</i>	486,33	<i>Petr</i>	444,77
<i>Petr</i>	293,36	<i>Petr</i>	398,64	<i>Petr</i>	428,31	<i>Petr</i>	495,56	<i>Jiří</i>	501,99
<i>Jan</i>	466,23	<i>Jan</i>	506,07	<i>Jan</i>	444,27	<i>Jan</i>	460,42	<i>Jan</i>	478,08

Tabulka 10. Tři nejčastější mužská jména – hypokoristika

SYN v11, publicistika									
1993		1998		2003		2008		2013	
Jméno	i.p.m.	Jméno	i.p.m.	Jméno	i.p.m.	Jméno	i.p.m.	Jméno	i.p.m.
<i>Jirka</i>	10,14	<i>Jirka</i>	11,11	<i>Jirka</i>	17,29	<i>Jirka</i>	19,59	<i>Pěťa</i>	2,81
<i>Pěťa</i>	0,71	<i>Pěťa</i>	0,77	<i>Pěťa</i>	1,12	<i>Pěťa</i>	2,18	<i>Jirka</i>	23,96
<i>Honza</i>	10,62	<i>Honza</i>	11,96	<i>Honza</i>	17,68	<i>Honza</i>	22,79	<i>Honza</i>	31,71

5.1 Kolokační analýza prvního nejčastějšího mužského jména

5.1.1 Základní podoba rodných jmen

V roce 1993 patří mezi kolokáty s nejvyšší dosaženou hodnotou dle zvolených parametrů výrazy odkazující na nositele daného jména jako na osobu, která plní vysokou státní funkci, např. *Šetina* (generální prokurátor, tab. 11: 1993), *Hronovský* (ředitel Generálního ředitelství cel ČR, tab. 11: 1993). Pro roky 1998, 2008 a 2013 jsou charakteristické kolokáty ve formě příjmení, které se pojí s tehdejší významnou politickou osobností, např. *Skalický* (politik, tab. 11: 1998), *Paroubek* (bývalý premiér, tab. 11: 2008), *Fiala* (ministr, tab. 11: 2013). Výraznou složku pro rok 2003 tvoří jednotky označující sportovce, např. *Kotrba* (fotbalista, tab. 11: 2003), *Hudler* (hokejista, tab. 11: 2003). Je zajímavé, že se ve zkoumaných obdobích objevují rovněž odkazy spojené s historickou postavou (kolokát *Poděbrady* fungující nejčastěji jako složka toponyma či chrématonyma vzniklého transonymizací, např. král *Jiří z Poděbrad* náměstí *Jiřího z Poděbrad*, tab. 11: 2008) a uměním a literaturou, např. *Bezruč* (básník, tab. 11: 2013).

Tabulka 11. Kolokační analýza prvního nejčastějšího mužského jména – základní podoby

SYN v11, publicistika									
1993 (<i>Jiří</i>)		1998 (<i>Jiří</i>)		2003 (<i>Jiří</i>)		2008 (<i>Jiří</i>)		2013 (<i>Petr</i>)	
Kolokát	ID	Kolokát	ID	Kolokát	ID	Kolokát	ID	Kolokát	ID
<i>Šetina</i>	10,15	<i>Skalický</i>	8,73	<i>Kotrba</i>	8,66	<i>Čunek</i>	9,74	<i>Nečas</i>	9,83

<i>Rulf</i>	7,58	<i>Pehe</i>	7,95	<i>Hudler</i>	7,91	<i>Paroubek</i>	9,51	<i>Fiala</i>	7,81
<i>Hronovský</i>	7,45	<i>Menzel</i>	7,83	<i>Jarošík</i>	7,81	<i>Šulc</i>	7,81	<i>Bezruč</i>	7,65
<i>prokurátor</i>	7,30	<i>Křovák</i>	7,82	<i>Rusnok</i>	7,67	<i>Pospíšil</i>	7,51	<i>Bendl</i>	7,58
<i>Honajzer</i>	7,15	<i>Kotrba</i>	7,80	<i>Šneberger</i>	7,62	<i>Poděbrady</i>	7,27	<i>Gazdik</i>	7,57

5.1.2 Hypokoristika

Hypokoristické formy sledovaných mužských rodných jmen se v publicistice nejčastěji vyskytují v kontextu umění, literatury a sportu. Rok 1993 je vyváženě zastoupen kolokáty, které odkazují na oblast filmovou, divadelní a sportovní, např. *Parma* (skokan na lyžích, tab. 12: 1993), *Jiřík* (herec, který vytvořil roli *ovčáka Jirky*, tab. 12: 1993). Charakteristickým rysem dalších zkoumaných období, tj. 1998 a 2003, je odkaz na sportovní prostředí, např. *Hřebec* (tenista, tab. 12: 1998), *Vitásek* (veslař tab. 12: 2003). V roce 2008 do popředí vystupují dva typy vazeb – kolokáty se pojí s prostředím zábavy, např. *Zonyga* (vítěz show X Factor, tab. 12: 2008), nebo s výrazy charakterizujícími povahu člověka, např. *floutek* (lehkomyslný, rozmařilý mladík, tab. 12: 2008). Výraznou složku posledního zkoumaného roku (2013) tvoří výskyty, které lze vnímat jako kolokáty spojené s kulturní oblastí, např. *Prokofjevův*, *Prokofjev* (autor symfonické pohádky *Pěťa a Vlk*, tab. 12: 2013).

Tabulka 12. Kolokační analýza prvního nejčastějšího mužského jména – hypokoristika

SYN v11, publicistika									
1993 (<i>Jirka</i>)		1998 (<i>Jirka</i>)		2003 (<i>Jirka</i>)		2008 (<i>Jirka</i>)		2013 (<i>Pěťa</i>)	
Kolokát	ID	Kolokát	ID	Kolokát	ID	Kolokát	ID	Kolokát	ID
<i>Spalek</i>	8,77	<i>Vochoška</i>	7,76	<i>Koběřský</i>	6,73	<i>Zonyga</i>	7,99	<i>Razím</i>	7,49
<i>Parma</i>	3,78	<i>kvestor</i>	7,01	<i>Vitásek</i>	6,52	<i>Floutek</i>	6,86	<i>Prokofjevův</i>	6,51
<i>Jiřík</i>	2,82	<i>Hřebec</i>	6,97	<i>Jarošík</i>	6,39	<i>Mádl</i>	6,60	<i>Štochlová</i>	6,36
<i>Lábus</i>	2,75	<i>Barbořík</i>	6,78	<i>kvestor</i>	6,31	<i>Podvodníček</i>	6,60	<i>Bajza</i>	6,31
<i>Kotrba</i>	2,38	<i>euro-poučka</i>	6,34	<i>párový</i>	6,22	<i>Doleček</i>	6,30	<i>Prokofjev</i>	6,17

5.2 Kolokační analýza druhého nejčastějšího mužského jména

5.2.1 Základní podoba rodných jmen

Kolokační analýza základní podoby druhého nejčastějšího mužského jména (*Petr, Jiří*) odhalila, že v letech 1998, 2003, 2008, 2013 převládají odkazy na osoby působící v politické sféře. Jsou to například kolokáty *Pithart* (předseda Senátu ČR, tab. 13: 1998, 2003), *Nečas* (ministr, tab. 13: 2008), *Dienstbier* (politik, syn politika Jiřího Dienstbiera staršího, tab. 13: 2013). Další častou oblastí bylo umění a literatura – *Bezruč* (básník, tab. 13: 1998, 2003), *Zelenka* (scénárista a režisér, tab. 13: 2008), *Menzel* (režisér, tab. 13: 2013). Celkem 3 kolokáty poukazují na sportovní oblast, jsou to: *Korda* (tenista, tab. 13: 1998), *Uličný* (fotbalista, tab. 13: 1998, 2003) a *Rada* (fotbalista, tab. 13: 2008). Pouze pro rok 1993 je typická oblast žurnalistiky, zde např. *Jedinák* (fotograf časopisu Reflex, tab. 13: 1993), *Janyška* (novinář, diplomat, tab. 13: 1993).

Tabulka 13. Kolokační analýza druhého nejčastějšího mužského jména – základní podoby

SYN v11, publicistika									
1993 (<i>Petr</i>)		1998 (<i>Petr</i>)		2003 (<i>Petr</i>)		2008 (<i>Petr</i>)		2013 (<i>Jiří</i>)	
Kolokát	ID	Kolokát	ID	Kolokát	ID	Kolokát	ID	Kolokát	ID
<i>Jedinák</i>	9,84	<i>Korda</i>	9,36	<i>Pithart</i>	8,86	<i>Rada</i>	8,63	<i>Rusnok</i>	9,88
<i>Pit'ha</i>	8,92	<i>Pithart</i>	8,89	<i>Bezruč</i>	8,53	<i>Bendl</i>	8,52	<i>Dienstbier</i>	9,27
<i>Janyška</i>	8,13	<i>Bezruč</i>	8,88	<i>Mareš</i>	8,08	<i>Gandalovič</i>	8,19	<i>Zimola</i>	8,05
<i>Wohl-hmut</i>	7,56	<i>Moos</i>	8,37	<i>Uličný</i>	7,80	<i>Zelenka</i>	7,98	<i>Menzel</i>	7,83
<i>Broulík</i>	7,34	<i>Uličný</i>	8,13	<i>Ibl</i>	7,67	<i>Nečas</i>	7,88	<i>Pospíšil</i>	7,52

5.2.2 Hypokoristika

U hypokoristických podob druhého nejčastějšího mužského jména jsme do podrobnější analýzy nezahrnuli rok 1993, jelikož tyto výsledky nebyly pro náš výzkum relevantní, jedná se totiž o interpunkční

znaménka, předložky a spojky. Co se tedy týče zbylých let, výrazné zastoupení zde měly kolokáty vztahující se k oblasti zábavy – např. *golet* (název rockového klubu, tab. 14: 1998), *tygřík* (loutková pohádka, tab. 14: 2003), *Jardův* (postava přítele z filmové komedie *Poslední plavky*, tab. 14: 2008), *Hadaš* (dětský bavič, kouzelník, tab. 14: 2013) – a také umění a literatury, např. *Bajza* (protagonista knihy *Bylo nás pět*, tab. 14: 1998), *vlk* (součást názvu symfonické pohádky Sergeje Prokofjeva *Pěť a vlk*, tab. 14: 2003), *Tázok* (slovenský rapper, tab. 14: 2008). Pro rok 2013 byly typické kolokáty vztahující se ke sportu, např. *Štajner* (fotbalista, tab. 14: 2013), *Magál* (běžec na lyžích, tab. 14: 2013), *Lalák* (hokejista, tab. 14: 2013).

Tabulka 14. Kolokační analýza druhého nejčastějšího mužského jména – hypokoristika

SYN v11, publicistika									
1993 (<i>Pěť a</i>)		1998 (<i>Pěť a</i>)		2003 (<i>Pěť a</i>)		2008 (<i>Pěť a</i>)		2013 (<i>Jirka</i>)	
Kolokát	ID	Kolokát	ID	Kolokát	ID	Kolokát	ID	Kolokát	ID
(	-7,52	<i>Neználek</i>	7,65	<i>tygřík</i>	6,21	<i>Tázok</i>	8,21	<i>Hadaš</i>	7,36
)	-7,55	<i>golet</i>	6,28	<i>Budišov</i>	3,71	<i>Prokofjevův</i>	7,35	<i>Štajner</i>	6,51
s	-8,91	<i>rockotěka</i>	6,11	<i>krycí</i>	1,81	<i>Jardův</i>	7,3	<i>Magál</i>	6,29
a	-9,66	<i>Bajza</i>	5,88	<i>symbolický</i>	1,04	<i>Korman</i>	7,23	<i>Lalák</i>	6,24
se	-9,93	<i>Crazy</i>	4,1	<i>vlk</i>	1,04	<i>Řičář</i>	7,07	<i>Povr</i>	6,17

5.3 Kolokační analýza třetího nejčastějšího mužského jména

5.3.1 Základní podoba rodných jmen

Třetím nejčastějším jménem bylo ve všech pěti zkoumaných letech jméno *Jan*. V roce 1993 se nejvíce objevovalo v kolokacích spojených se žurnalistikou, např. *Šibík* (fotoreportér, tab. 15: 1993) nebo *Foll* (filmový publicista, tab. 15: 1993). V letech 1998, 2003 a 2013 dominovaly kolokace spojené s politickou oblastí, např. *Kavan* (politik a diplomat, tab. 15: 1998), *Zahradil* (poslanec Poslaneckého sněmovny Parlamentu ČR, tab. 15: 2003) nebo *Fischer* (statistik a politik, tab. 15: 2013). V roce 2008 a 2013 byla taktéž výrazná historicko-náboženská sféra, např. *Nepomucký* (světec, jeden z českých

zemských patronů, tab. 15: 2013) nebo *Palach* (student, který se v roce 1969 upálil na protest proti srpnové okupaci, tab. 15: 2013).

Tabulka 15. Kolokační analýza třetího nejčastějšího mužského jména – základní podoby

SYN v11, publicistika									
1993 (Jan)		1998 (Jan)		2003 (Jan)		2008 (Jan)		2013 (Jan)	
Kolokát	ID	Kolokát	ID	Kolokát	ID	Kolokát	ID	Kolokát	ID
<i>Mihaliček</i>	9,35	<i>Ruml</i>	10,49	<i>Koller</i>	8,49	<i>Švejnar</i>	9,93	<i>Fischer</i>	9,1
<i>Šibík</i>	8,9	<i>Koukal</i>	9,41	<i>Hřebejk</i>	8,49	<i>Nepomucký</i>	8,23	<i>Palach</i>	8,3
<i>Kryčer</i>	8,59	<i>Kavan</i>	9,02	<i>Werich</i>	8,46	<i>Koller</i>	7,97	<i>Nepomucký</i>	8,15
<i>Foll</i>	8,58	<i>Kasal</i>	8,53	<i>Kavan</i>	8,38	<i>Křtitel</i>	7,7	<i>Křtitel</i>	7,53
<i>Petránek</i>	8,43	<i>Fencl</i>	7,88	<i>Zahradil</i>	8,04	<i>Palach</i>	7,66	<i>Burian</i>	7,46

5.3.2 Hypokoristika

Hypokoristikum *Honza*, tj. varianta třetího nejfrekventovanějšího jména, se nejčastěji objevovalo v oblasti umění a sportu. Ve všech zvolených letech kromě roku 2013 dominuje odkaz na umělecké zaměření dané osoby, např. *Svěrák* (režisér a scenárista, tab. 16: 1993), *Vyčítal* (hudebník, kreslíř-karikaturista, tab. 16: 1998) nebo *Muchow* (skladatel, hudební producent, tab. 16: 2003). V těchto letech je výrazné rovněž spojení se zábavní oblastí, např. *Dědek* (novinář a hudební publicista, na TV Prima vysílá od roku 2012 pořad *7 pádů Honzy Dědka*, tab. 16: 2013), *křupan* (Křupani – volné sdružení muzikantů, které vede *Honza Křupan*, tab. 16: 1998). V roce 2013 se nejčastěji vyskytuje odkaz na sportovní prostředí, např. *Nezmar* (fotbalista, tab. 16: 2013).

Tabulka 16. Kolokační analýza třetího nejčastějšího mužského jména – hypokoristika

Syn v11, publicistika									
1993 (<i>Honza</i>)		1998 (<i>Honza</i>)		2003 (<i>Honza</i>)		2008 (<i>Honza</i>)		2013 (<i>Honza</i>)	
Kolokát	ID	Kolokát	ID	Kolokát	ID	Kolokát	ID	Kolokát	ID
<i>Zábrana</i>	5,33	<i>Vyčítal</i>	8,09	<i>PALDĚ</i>	7,74	<i>Vyčítal</i>	7,06	<i>Dědek</i>	6,45

<i>Malíř</i>	3,94	<i>křupan</i>	7,37	<i>Vyčítal</i>	7,23	<i>Lušovský</i>	6,43	<i>Nezmar</i>	6,43
<i>Hajný</i>	3,16	<i>Drdův</i>	6,92	<i>Kultra</i>	6,92	<i>Topinka</i>	6,24	<i>Vyčítal</i>	6,25
<i>kameraman</i>	1,4	<i>Carolina</i>	6,64	<i>Muchow</i>	6,37	<i>floutek</i>	6,12	<i>Strmiska</i>	6,23
<i>Svěrák</i>	1,38	<i>BMG</i>	6,37	<i>Kutra</i>	6,36	<i>zajetý</i>	5,88	<i>Trousil</i>	6,21

Závěry

Pomocí korpusového nástroje KonText jsme provedli kolokační analýzu vybraných ženských a mužských rodných jmen a jejich hypokoristických podob v publicistice korpusu SYN v11 v období 1993 až 2013. Cílem bylo analyzovat společensko-kulturní kontext a vyexcerpovat nejčastější oblasti, se kterými se daná jména pojí.

Jak již bylo řečeno výše, analýze jsme podrobili tři ženská jména a tři mužská jména, která měla ve vybraných letech (1993, 1998, 2003, 2008 a 2013) nejvyšší počet žijících nositelů (*Marie, Jana, Eva, Hana a Jiří, Petr, Jan*). Rovněž jsme analyzovali jejich hypokoristické varianty (*Maruška, Janička, Evička, Hanka a Jirka, Pěťa, Honza*). Nejčastějšími kolokáty základních podob rodných jmen ženských i mužských byla dle našich očekávání příjmení. Co se týče převažujících tematických oblastí výskytu, byla to zejména sféra politiky, státních institucí a žurnalistiky. V politice se častěji vyskytovala mužská jména, ale naopak v oblasti státních institucí dominovala ženská jména. V prostředí žurnalistiky bylo zastoupení proprií (podle rodu jména) rovnoměrné. Hypokoristické varianty jmen ženských i mužských se nejčastěji objevovaly ve sféře umění a zábavy, kde měla hypokoristika rovnoměrné zastoupení jak mezi ženami, tak mezi muži. Další oblastí byl sport, kde převládaly mužské varianty jmen, a odkaz na děti a mládí, kde měla majoritní zastoupení ženská hypokoristika.

Pokud se na výsledky zkoumání podíváme z diachronní perspektivy, zjistíme, že v letech 1993, 2003 a 2013 významně převažují ženská propria v oblastech s odkazem na zaměstnance státních institucí. Roky 1998 a 2008 se vyznačují různorodostí, jde např. O prostředí sportu, náboženství, umění, ale i politiky. Mužská propria se v roce

1993 nejčastěji objevila v oblasti žurnalistiky, ale od roku 1998 až po poslední zkoumaný rok (2013) dominuje sféra politiky a sportu.

Hypokoristické varianty se v průběhu let nevyznačovaly velkými frekvencemi. Co se týče oblastí výskytu, u ženských forem převažoval odkaz na mladistvé, u mužských hypokoristik pak spojení s uměním a zábavní oblastí.

Hypotézy předpokládající výskyt základních podob rodných jmen s kolokáty souvisejícími s politickou sférou a státními institucemi lze považovat za udržitelné. Standardizované podoby rodných jmen se pak občasné objevily v prostředí sportu a zábavy. Stejně jsme ověřily předpoklad výskytu hypokoristik v prostředí zábavy. Výjimečně jsme pak zaznamenali domácké varianty rodných jmen ve sféře politiky a státních institucí. Do budoucna by bylo záhodno prozkoumat vedle nejčastějších kolokací také variabilitu jmen, jelikož se sledovaná jména mohou lišit v tom, jak variabilní je jejich okolí.

Literatura

- Brezina V., McEnery T., Wattam S., 2015, *Collocations in Context. A New Perspective on Collocation Networks*. „International Journal of Corpus Linguistics“ 20(2), s. 139–173.
- Brezina V., *Statistics in Corpus Linguistics. A Practical Guide*. Cambridge: Cambridge University Press.
- Cvrček V., 2019, *Asociační (kolokační) míry. Wiki – Český národní korpus*. Online: https://wiki.korpus.cz/doku.php/pojmy:asociacni_miry [cit. 14. 11. 2023].
- Cvrček V. a kol., 2020, *Registry v češtině*. Praha: Nakladatelství Lidové noviny.
- David J., Klemensová T., Místecký M., 2022, *Od etymologie ke krajině: onomastika pro 21. století*. Brno: Host / Ostrava: Ostravská univerzita.
- Knappová M., 2017, *Jak se bude vaše dítě jmenovat?* Praha: Academia.
- Kováříková D., 2021, *i.p.m.* Praha: Ústav Českého národního korpusu FF UK. Online: <https://wiki.korpus.cz/doku.php/pojmy:ipm> [cit. 14. 11. 2023].
- Křen M., Cvrček V., Hnátková M., Jelínek T., Kocek J., Kovářiková D., Křivan J., Milička J., Petkevič V., Procházka P., Skoumalová H., Šindlerová J., Škrabal M., 2022, *Korpus SYN, verze 11*. Praha: Ústav Českého národního korpusu FF UK. Online: <https://www.korpus.cz> [cit. 14. 11. 2023].

- Místecký M., Drkošková V., 2022, *Od patvaru k superstar: frekvenční a kolokační analýza názvu Česko v korpusové publicistice z let 1990–2018*. „Acta onomastica“ 63(1), s. 99–115.
- Osolsobě K., Žižková H., 2021, *Homonymie mezi oikonymy a antroponymy za-končenými na -slav/-slava jako problém automatické morfologické analýzy*. „Acta onomastica“ 62 (2), s. 377–403.
- Pleskalová J., 2011, *Vývoj vlastních jmen osobních v českých zemích v letech 1000–2010*. Brno: Masarykova univerzita.
- Tušková J. M., 2019, *Korpusy češtiny a výzkum (gramatických vlastností) proprií*. In: A. Chomová, J. Krško, I. Valentová (eds.). *Konvergence a divergence v propriální sféře*. Bratislava: Veda, s. 126–137.
- Tušková J. M., 2011, *Deklinací systém femininních oikonym v češtině. Synchronní pohled na základě Českého národního korpusu*. Praha: Nakladatelství Lidové noviny.