

Univerzita Karlova

Filozofická fakulta

Ústav českého jazyka a teorie komunikace

Diplomová práce

Mgr. Anna Dittrichová

Vyhledávání frazémů na základě anomálie v distribuci tvarů

Finding idioms based on anomalous word-form distribution

Praha 2023

Vedoucí práce: Mgr. Dominika Kovářiková, M.A., Ph.D.

Poděkování

Ráda bych tímto poděkovala své vedoucí práce Mgr. Dominice Kovářkové, M.A., Ph.D., za ochotu diplomovou práci vést, mnoho cenných a příjemných konzultací a velmi podrobné připomínky. Dále bych ráda poděkovala celé své rodině, partnerovi a nejbližším přátelům za veškerou podporu během celého studia a za dodávání motivace studium dokončit.

Prohlášení

Prohlašuji, že jsem diplomovou práci vypracovala samostatně, že jsem řádně citovala všechny použité prameny a literaturu a že práce nebyla využita v rámci jiného vysokoškolského studia či k získání jiného nebo stejného titulu.

V Praze dne 18. prosince 2023

Anna Dittrichová

Abstrakt

Diplomová práce se zabývá vyhledáváním, popisem, klasifikací a slovníkovým srovnáním frazémů, a to na základě dat poskytovaných aplikací GramatiKat, která dokáže identifikovat substantivní lemmata s anomální distribucí tvarů. Cílem práce je zjistit, zda gramatické anomálie substantiv (tedy nezvykle vysoká frekvence jednoho či více tvarů v rámci morfologického paradigmatu) ukazují na formálně anomální víceslovné lexémy (tedy frazémy), případně jak se tato souvislost liší v rámci jednotlivých pádů a jaké typy frazémů se v jednotlivých pádech jednotného čísla objevují nejčastěji. Z analýzy, která využívá jazykové korpusy SYN2015 a SYNv11, vyplynulo, že na jednom či více frazémech se podílí celkem 28 % analyzovaných lemmat, přičemž nejčastěji se jedná o akuzativ (až 88 %), nejméně naopak o vokativ (5 %). Ve zkoumaných datech jsou zastoupeny skupiny frazémů se stejnými vlastnostmi, například v dativu se často jednalo o verbální frazémy, ve vokativu o kontaktné frazémy a v lokálu o gramatické frazémy.

Klíčová slova

frazémy, anomálie paradigmatu, distribuce tvarů, GramatiKat, substantiva, korpusová analýza

Abstract

This diploma thesis deals with the search, description, classification, and dictionary comparison of idioms based on data provided by the GramatiKat application, which can identify nominal lemmas with anomalous word-form distribution. The aim of the diploma thesis is to determine whether grammatical anomalies in nouns (unusually high frequencies of one or more forms within a morphological paradigm) indicate formally anomalous multi-word lexemes (idioms) and how this relationship varies across different cases. Additionally, it explores the types of idioms that most commonly appear in individual cases of the singular. The analysis, utilizing SYN2015 and SYNv11 corpora, revealed that 28 % of analyzed lemmas are part of one or more idioms. The most common case is the accusative (88 %), while the least common is the vocative (5 %). The analysis also identified various groups of idioms with similar characteristics. For instance, verbal idioms were frequently observed in the dative, contact idioms predominated in the vocative, and grammatical idioms were prevalent in the locative.

Keywords

idioms, paradigm anomaly, word-form distribution, GramatiKat, nouns, corpus analysis

Obsah

Úvod	11
1 Frazémy a idiomy	13
1.1 Vymezení pojmů frazém a idiom	13
1.2 Rysy frazémů	14
1.3 Typy frazémů	14
1.3.1 Čermákova formálně-strukturní klasifikace	14
1.3.2 Vlastní klasifikace využívaná v analytické části práce	17
1.4 Variabilita frazémů	20
2 Gramatické profily	21
3 Frekvence pádů u substantiv v češtině	23
4 Jazykové korpusy a jejich nástroje a funkce využívané při analýze	27
4.1 Korpus SYN2015	27
4.2 Korpus SYNv11	27
4.3 GramatiKat	27
4.4 FRANTA	32
4.5 Kolokace v korpusu	32
5 Slovníky využívané při analýze a jejich přístup k frazeologii	35
5.1 Slovník české frazeologie a idiomatiky	35
5.2 Akademický slovník současné češtiny	36
6 Metodologie	37
6.1 Cíle analýzy	37
6.2 Použitý materiál	37
6.3 Postup	38
6.4 Popis jednotlivých pádů	40

7	Analýza vybraných anomálních lemmat	41
7.1	Nominativ singuláru	41
7.2	Genitiv singuláru	43
7.3	Dativ singuláru	47
7.4	Akuzativ singuláru	53
7.5	Vokativ singuláru	60
7.6	Lokál singuláru.....	61
7.7	Instrumentál singuláru.....	68
7.8	Srovnání vyhledaných frazémů dle jednotlivých pádů	75
	Závěr	81
	Seznam použité literatury	82
	Přílohy	I

Seznam grafů

Graf 1: Porovnání frekvence pádů u substantiv v procentech dle různých zdrojů	26
---	----

Seznam tabulek

Tabulka 1: Frekvence pádů u substantiv v procentech dle Jelínka, Bečky a Těšitelové (1961, s. 94–95).....	23
Tabulka 2: Frekvence pádů u substantiv v procentech dle Bartoně a kol. (2009, s. 134)	24
Tabulka 3: Frekvence pádů u substantiv v procentech dle Kodýtky (2010, s. 141).....	25
Tabulka 4: Frekvence pádů u substantiv v procentech v SYN2015.....	25
Tabulka 5: Frekvence pádů u substantiv v procentech v SYN2020.....	25
Tabulka 6: Procentuální hodnoty z boxplotů znázorňujících distribuci tvarů (Kováříková 2021b, s. 536)	29
Tabulka 7: Počet horních odlehlých hodnot a spodních extrémů dle jednotlivých tvarů (Kováříková a Kovářík 2021).....	31
Tabulka 8: Celkový počet lemmat anomálních v horních odlehlých hodnotách u jednotlivých pádů a velikost zvoleného vzorku	38
Tabulka 9: Typy vyhledaných frazémů v nominativu singuláru.....	43
Tabulka 10: Typy vyhledaných frazémů v genitivu singuláru.....	47
Tabulka 11: Typy vyhledaných frazémů v dativu singuláru	52
Tabulka 12: Typy vyhledaných frazémů v akuzativu singuláru	59
Tabulka 13: Typy vyhledaných frazémů ve vokativu singuláru	61
Tabulka 14: Typy vyhledaných frazémů v lokálu singuláru	68
Tabulka 15: Typy vyhledaných frazémů v instrumentálu singuláru	75
Tabulka 16: Shrnutí výsledků dle jednotlivých pádů	76
Tabulka 17: Typy frazémů dle jednotlivých pádů.....	76

Seznam obrázků

Obrázek 1: Distribuce tvarů českých substantiv na základě dat z korpusu SYN2015 (Kováříková 2021b, s. 535).....	29
Obrázek 2: Distribuce tvarů lemmatu důsledek (Kováříková a Kovářik 2021).....	30
Obrázek 3: Lemma důsledek, distribuce v singuláru a plurálu (Kováříková a Kovářik 2021)	31

Seznam použitých zkratk

A	adjektivum
ADV	adverbium
ASSČ	Akademický slovník současné češtiny
ČNK	Český národní korpus
FRANTA	Frazémová anotace a textová analýza
int	interjekce
KL	klauze
kon	konjunkce
mono	monokolokabilita
num	numeralie
part	partikule
prep	prepozice
pron	pronomén
S	substantivum
S1	nominativ singuláru
S2	genitiv singuláru
S3	dativ singuláru
S4	akuzativ singuláru
S5	vokativ singuláru
S6	lokál singuláru
S7	instrumentál singuláru
SČFI	Slovník české frazeologie a idiomatiky
syn	synsémantikum
ÚJČ AV ČR	Ústav pro jazyk český Akademie věd České republiky
V	verbum

Úvod

Pouze zlomek lexémů morfologicky bohatého jazyka, jako je čeština, se vyskytuje ve všech paradigmatických formách (Janda a Tyers 2021, s. 109). V některých tvarech pak lexémy často mívají neobvykle vysokou nebo naopak velmi nízkou až nulovou frekvenci. Právě na anomálii v distribuci tvarů českých substantiv je tato práce založena. Jedná se o práci na pomezí korpusové lingvistiky, frazeologie, lexikologie a morfologie.

Cílem práce je zjistit, zda gramatické anomálie substantiv (tj. nezvykle vysoká frekvence jednoho či více tvarů v rámci morfologického paradigmatu) mohou ukazovat na formálně anomální víceslovné lexémy (frazémy). Pozornost je dále soustředěna na to, v jakých pádech se taková souvislost projevuje nejčastěji a jaké typy frazémů se v rámci jednotlivých pádů objevují. V rámci analýzy jsou tedy v korpusu vyhledávána substantiva, která jsou v některém tvaru frekventovanější, než je u českých substantiv běžné, přičemž je zjišťováno, zda je tato anomálie způsobena tím, že jsou součástí frazémů. Vyhledané frazémy jsou následně klasifikovány a rovněž u nich je zaznamenáno, zda jsou slovníkově zachyceny, či nikoli.

Teoretická část práce se nejprve zaměřuje na frazeologii, které se v českém prostředí věnoval zvláště František Čermák, mimo jiné autor Slovníku české frazeologie a idiomatiky, který v 80. letech 20. století vypracoval dosud užívaný teoretický přístup a také přehled typů frazémů (Čermák 2007a, s. 246). Ve druhé kapitole jsou popsány gramatické profily dle konceptu Jandy a Lyashevskayi, které souvisejí s nástrojem GramatiKat, na jehož datech je založena následná analýza. Jelikož gramatické profily slov tvoří frekvenční distribuce tvarů, třetí kapitola se věnuje srovnání frekvence pádů u substantiv v češtině v odborné literatuře od 40. let 20. století do současnosti. Ve čtvrté a páté kapitole jsou popsány jazykové korpusy (SYN2015 a SYNv11), korpusové nástroje a funkce (GramatiKat, FRANTA a kolokace) a slovníky (SČFI a ASSČ) využívané při následné analýze.

V rámci analytické části jsou nejprve v šesté kapitole o metodologii popsány cíle výzkumu, způsob výběru použitého materiálu, postup a podoba popisu jednotlivých pádů. Sedmá kapitola obsahuje samotnou analýzu – v sedmi podkapitolách jsou postupně popsány výsledky dle jednotlivých pádů v singuláru. U každého z nich je vždy uvedeno, kolik lemmat je součástí alespoň jednoho frazému, a dále jsou vypsány všechny vyhledané frazémy (přičemž jsou rozlišeny frazémy slovníkově zachycené a nezachycené), které jsou následně tříděny do devíti předem určených skupin (gramatické frazémy, monokolokabilní frazémy, binomiály, přirovnání, kontaktní frazémy, cizojazyčná spojení, nominální frazémy, verbální frazémy

a propoziční frazémy). Analytická část práce je zakončena srovnáním výsledků dle jednotlivých pádů.

1 Frazémy a idiomy

Frazémy a idiomy, jakožto primární jednotka lingvistické disciplíny frazeologie a idiomatika, spadají pod víceslovné lexikální jednotky, tedy formálně vícečlenná, významově celistvá a ustálená slovní spojení, která lze reprodukovat pouze jako celek a jejichž jednotlivé komponenty nelze nahrazovat synonymy, antonymy či je rozvíjet (Hladká 1996, s. 70; Čermák 2010, s. 225). Čermák (2010, s. 222) víceslovné jednotky rozděluje na:

- a) víceslovné termíny (např. *kyselina sírová*),
- b) víceslovná propria (např. *Ústí nad Labem*),
- c) idiomatické kolokace (např. *nechat někoho na holičkách*),
- d) běžné kolokace uzuální (např. *číst knihu*).

Jejich identifikace je však podle něj velmi obtížná, navíc neexistuje celkový soupis, a tedy ani informace o počtu. Čermák (2010, s. 223) však počty víceslovných lexémů odhaduje na miliony. Tato práce se bude věnovat jen bodu c), tedy idiomatickým kolokacím nazývaným rovněž frazémy a idiomy.

1.1 Vymezení pojmů frazém a idiom

Frazémem/idiomem rozumíme „svou anomálností jedinečnou a ustálenou kombinaci minimálně dvou prvků, z nichž některý/žádný nefunguje stejným způsobem (tj. s obdobnou funkcí) v jiném/jiných kombinacích, popř. se vyskytuje pouze v kombinaci jediné/několika málo“ (Čermák 2010, s. 225). Podle Čermáka (2010, s. 225) je tedy frazém kombinace několika prvků – ať už morfémů, slovních tvarů, nebo kolokací, – ze kterých je alespoň jeden součástí omezeného a uzavřeného paradigmatu. Pokud tedy tento prvek nelze zaměnit za jiný, jedná se o frazém (např. *flintu* ve frazému *házet flintu do žita* nelze nahradit například *puškou*).

V Čermákově pojetí je pak frazém širší pojem než idiom – idiom je tedy vždy frazémem, zatímco ne každý frazém je rovněž idiomem. Frazém totiž zdůrazňuje formální anomálii (např. frazém *křížem krážem* nelze zaměnit za *krážem křížem*), zatímco idiom tu sémantickou (např. idiom *psi počasí*, který byl vytvořen na základě metafor). Právě metafora byla dříve považována za hlavní rys frazémů, dnes je však pojetí frazému širší. Například gramatické frazémy nelze považovat za idiomy, jelikož nemají anomální sémantiku, nýbrž pouze formální anomálii, tj. anomální kolokabilitu svých komponentů. (Čermák 2010, s. 227; Kopřivová 2010, s. 75)

1.2 Rysy frazémů

Podle Čermáka (2010, s. 225) je podstatou frazému a idiomu „anomálie (nepravidelnost) v různém smyslu (syntagmatická a/nebo paradigmatická), vytvářející neopakovatelné a podle žádných pravidel negenerovatelné kombinace, typicky užívané zvláště v mluveném jazyce“. Podle Kopřivové (2010, s. 75) pak lze anomálnost kombinace testovat změnou komponentu daného frazému jiným komponentem, který je ze stejné třídy. Pokud tato výměna není možná, jedná se o frazém. Jako příklad lze uvést frazém *házet flintu do žita*, který není zaměnitelný například za *házet flintu do pšenice*.

Dalším rysem je významová nesoučtovost komponentů frazému – na rozdíl od kolokací totiž není význam z jednotlivých komponentů frazému srozumitelný (Kopřivová 2010, s. 76). Hladká (1996, s. 71) za typické znaky frazémů dále považuje především metaforičnost a expresivnost a upozorňuje, že se ve frazémech často objevují archaismy (např. *kout pikle*) a monokolokabilní prvky (např. *honit bycha*), které se vyskytují pouze v daném frazému. Čechová (2011, s. 66) a Kopřivová (2010, s. 75) vedle již zmíněných rysů uvádějí také ustálenost formy i významu, vícesložkovost a omezenou kolokabilitu a Kochová s Opavskou (2016, s. 167) dodávají denotativní vágnost, významovou neurčitost, hodnotící funkci a idiomaticnost.

1.3 Typy frazémů

Frazémy lze klasifikovat dle různých kritérií, například z hlediska stylistického, sémiotického, diachronního, sémantického či funkčního. Čermák je však k těmto hlediskům kritický a kritéria těchto klasifikací považuje za nejasněná, příliš obecná nebo nedostatečně podložená spolehlivými zdroji. Dále bude proto podrobněji představena Čermákova klasifikace z hlediska formálně-strukturního, „založená na evidenci vstupních komponentů, které sice o frazému mnoho nevypovídají, navenek tu však sledované komponenty jsou většinou povědomé (s výjimkou monokolokabilních slov) a vnějškově shodné se slovy pravidelného jazyka“ (Čermák 2007b, s. 44).

1.3.1 Čermákova formálně-strukturní klasifikace

Čermák (2010, s. 227) tedy frazémy klasifikuje dle povahy vstupních komponentů na tři základní skupiny:

- a) lexikální frazémy,
- b) kolokační frazémy,

c) propoziční frazémy.

Lexikální frazémy

Lexikální frazémy jsou podle Čermáka (2010, s. 227) kombinace morfémů v jednoslovném lexému. Funkčně jsou lexikální frazémy jednoslovné lexémy, jejichž plný popis však v češtině chybí, jelikož se jedná o nejspornější oblast frazeologie. Jako příklad lexikálních frazémů lze uvést například frazémy *kratochvíle*, *tlučhuba*, *každopádně*, *neplecha*, *budiž*, *nač* či *zašít se* (Čermák 2007b, s. 60).

Kolokační frazémy

Kolokační frazémy jsou naopak nejpočetnější skupinou frazeologie, jedná se o kombinace tvarů lexémů s nevětnou funkcí. Na rozdíl od pravidelného jazyka je v rámci frazémů možné kombinovat všechny slovní druhy v libovolném pořadí (například ve frazému *anděl strážný* je přívlastek v postpozici). V rámci kolokačních frazémů Čermák rozděluje 16 základních binárních kombinací dle autosémantických komponentů, 7 základních kombinací synsémantických komponentů a 9 základních kombinací dle autosémantických i synsémantických komponentů. Pod tyto binární typy pak spadají další vícekomponentové typy. (Čermák 2007b, s. 44–48; Čermák 2010, s. 227–230)

Kombinace dle autosémantických komponentů mohou být následující (Čermák 2007b, s. 44–45):

- 1) S–V (např. *pravda vítězí*),
- 2) V–S (např. *chytat lelky*),
- 3) A–S (např. *černá hodinka*),
- 4) S–A (např. *anděl strážný*),
- 5) S–ADV (např. *klobouk dolů*),
- 6) ADV–S (např. *vzhůru nohama*),
- 7) V–A (např. *kápnout božskou*),
- 8) A–V (např. *zmýlená neplatí*),
- 9) V–ADV (např. *dopadnout bledě*),
- 10) ADV–V (např. *zarytě mlčet*),
- 11) A–ADV (např. *malá domů*),
- 12) ADV–A (např. *zbrusu nový*),
- 13) S–S (např. *jablko sváru*),
- 14) V–V (např. *dát se poddat*),

- 15) A–A (např. *neslaný nemastný*),
- 16) ADV–ADV (např. *široko daleko*).

Kolokační frazémy mohou být dále složeny pouze ze synsémantik – Čermák (2007b, s. 46–47) vyděluje 7 základních binárních typů:

- 1) kombinace začínající předložkou (např. *mimo jiné*),
- 2) kombinace začínající spojkou (např. *a sice*),
- 3) kombinace začínající zájmenem (např. *co kdyby*),
- 4) kombinace začínající číslovkou (např. *jeden každý*),
- 5) kombinace začínající citoslovcem (např. *bác ho!*),
- 6) kombinace začínající částicí (např. *no tohle!*),
- 7) kombinace začínající monokolokabilním komponentem (např. *kór když*).

Třetím typem jsou kolokační frazémy, které se skládají jak z autosémantik, tak se synsémantik. Čermák (2007b, s. 47–49) vyděluje 9 základních typů:

- 1) kombinace začínající předložkou (např. *bez okolků*),
- 2) kombinace začínající spojkou (např. *co živ*),
- 3) kombinace začínající zájmenem (např. *co naplat!*),
- 4) kombinace začínající citoslovcem (např. *šup sem!*),
- 5) kombinace začínající částicí (např. *až černo*),
- 6) kombinace začínající monokolokabilním komponentem (např. *baj vočko*),
- 7) kombinace začínající slovesem (např. *být pro*),
- 8) kombinace začínající jménem (např. *křížem krážem, opilý namol*),
- 9) kombinace začínající adverbium (např. *zhola nic*).

Propoziční frazémy

Propoziční frazémy jsou kombinace tvarů lexémů, kolokací i klauzí, a to s větnou funkcí, které Čermák (2007b, s. 52) dále dělí podle počtu zúčastněných mluvčích na monosubjektové a intersubjektové. Propoziční frazémy mívají 2–14 slov, obvykle však 4–5. Čermák (2007b, s. 54) uvádí 10 základních struktur, které lze stejně jako u frazémů kolokačních dále rozšiřovat:

- 1) To je S / A / ADV / mono / prep–S (např. *To je k smíchu.*),
- 2) S–V (např. *Pravda vítězí.*),
- 3) S–V–S (např. *Něco je ve vzduchu.*),
- 4) S–V–A (např. *Vzduch je čistej.*),
- 5) S–V–ADV (např. *To něco pamatuje.*),

- 6) ADV–V / mono–V (např. *Pozdě bycha honit.*),
- 7) S–S (např. *Mladost radost.*),
- 8) prep–S–A–S / S–A / S–ADV / A–ADV aj. (např. *Malá domů.*),
- 9) ADV–ADV (např. *Zkrátka a dobře.*),
- 10) syn–syn (např. *No ne!*).

1.3.2 Vlastní klasifikace využívaná v analytické části práce

Jelikož je Čermákova formálně-strukturní klasifikace frazémů příliš podrobná a dělení dle SČFI na slovesné, neslovesné, větné frazémy a přirovnání zase příliš obecné, byla pro potřeby analytické části práce zvolena vlastní klasifikace. Frazémy byly dle svých vlastností tříděny nejprve do šesti skupin (gramatické frazémy, monokolokabilní frazémy, binomiály, přirovnání, kontaktní frazémy a cizojazyčná spojení). V případě, že nebyly zařazeny do žádné z těchto skupin, rozlišovaly se nominální, verbální a propoziční frazémy. V této podkapitole bude všech těchto devět typů popsáno.

Gramatické frazémy

Gramatické frazémy se vyznačují svou gramatickou funkcí – jsou tvořeny pouze ze syntaktik. Jedná se tedy o víceslovné předložky, spojky, částice, citoslovce, zájmena a číslovky. Dle slovních druhů se tak rozlišuje šest typů gramatických frazémů:

- 1) prepoziční (např. *bez zřetele k*),
- 2) konjukční (např. *i když*),
- 3) partikulární (např. *že jo?*),
- 4) interjekční (např. *Ježíš Marjá!*),
- 5) pronominální (např. *nějaký ten*),
- 6) numerální (např. *první poslední*). (Čermák 2017a)

Monokolokabilní frazémy

Monokolokabilní frazémy jsou ty frazémy, které obsahují komponent, jenž má velmi omezenou kolokabilitu – kombinuje se buď pouze s jedním slovem, nebo s mikrotřídou slov, přičemž Čermák za monokolokabilní považuje výrazy, které se pojí až se sedmi kolokáty (např. *být hin*, *otevřít dokořán*, *tratoliště krve*, *nechat někoho na holičkách*). (Cvrček a Richterová 2015)

Monokolokabilní komponenty se vyskytují zejména ve frazeologii, případně v termínech. Kvůli omezené kolokabilitě jsou jejich lemmata často pouze uměle vykonstruovaná (např.

lemma *holičky*) a zároveň nejsou žádným slovním druhem, protože nefungují samostatně. Tyto komponenty tedy v jazyce nemohou být spojeny s ničím jiným než s touto mikrotřídou slov, a nemají tak vlastní význam. Monokolokabilní tvary se dále mohou vyskytovat i v jednoslovných, tj. lexikálních lexémech (např. *levhart*). (Čermák 2017b)

Binomiály

Struktura binomiálů je založena na „iterativní kombinaci různého druhu, zvl. juxtaponovaném opakování dvou komponentů stejného slovního druhu nebo přímo opakování téže slovní formy“ (Čermák 2017c). Nejčastěji se jedná o substantiva (např. *bez hlavy a paty*), může však jít i o adjektiva (např. *nemastný neslaný*), adverbia (např. *široko daleko*) či slovesa (např. *dohnat a předechnat*). Binomiály často mívají zesilovací funkci (např. *šalba a klam*) či vyjadřují alternativní volbu (např. *peníze, nebo život*). (Čermák 2017c)

Přirovnání

Přirovnání dle Hladké (1996, s. 72) naznačuje větnou strukturu, ale některé pozice v něm se obsazují až v konkrétních kontextech. Přirovnání je přitom používáno při „charakteristice skutečnosti připodobněním jejího znaku k předem danému zobecněnému modelu [a] jeho hlavní funkcí [...] bývá především intenzifikace, zdůraznění a kladné či záporné zhodnocení tohoto znaku“ (Hladká 1996, s. 72). Základní podoba přirovnání má pět složek:

- comparandum (referent určený konkrétním textem, většinou se jedná o označení člověka, např. *Petr*),
- relátor (relační prvek, většinou sloveso, tedy predikát, např. *je*),
- tertium comparationis (společný znak comparanda a comparatu, např. *zvědavý*),
- komparátor (formální znak, většinou *jako*),
- comparatum (zobecněný model, ke kterému se přirovnává, většinou se jedná o zvíře či neživý předmět, např. *opice*). (Čermák 1982, s. 145–146)

Kontaktové frazémy

Do kontaktové frazémy byly v tomto případě zařazeny frazémy s kontaktovou funkcí. Většinou se tedy jedná o pozdravy (např. *Pozdrav pánbůh!*), výrazy díky (např. *Díky bohu!*), kletby a nadávky (např. *A sakra!*) či formule (např. *s dovolením*). (Čermák a kol. 2009, s. 1263)

Cizojazyčná spojení

Cizojazyčná spojení jsou spojení převzatá z cizího textu, nejčastěji z latiny, angličtiny, němčiny či francouzštiny, která se objevují v českém textu (např. *libri prohibiti, by the way*). (Petkevič a kol. 2020, s. 43)

Nominální frazémy

Za nominální (jmenné) frazémy jsou dle SČFI považovány substantivní, adjektivní a adverbialní frazémy, tedy frazémy tvořené z autosémantik. Strukturně se nejčastěji jedná především o tyto typy:

- 1) A–S / S–A (např. *bledá tvář*),
- 2) S–S (např. *tajemství úspěchu*),
- 3) A–A (např. *neslaný nemastný*),
- 4) ADV–A / A–ADV (např. *kříšťálově čistý*). (Čermák, Hronek a kol. 2009a, s. 9–10)

Verbální frazémy

Dle SČFI (Čermák, Hronek a kol. 2009b, s. 9) jsou verbální frazémy ty frazémy, které obsahují sloveso a mají typický slovesný charakter, ve větě zejména komplexního přísudku. Do této skupiny spadají následující kombinace:

- 1) V–S (např. *chytat lelky*),
- 2) V–A (např. *nebýt dnešní*),
- 3) V–ADV (např. *chodit spolu*),
- 4) V–V (např. *umět se narodit*),
- 5) V–KL (např. *myslet si, že snědl všechnu moudrost*),
- 6) V–syn (konkrétně V–pron, V–prep, V–num, V–int, V–part, V–kon, např. *být svůj, být pro, mít všech pět pé, dělat na někoho bububu, ani nezabučet, dát někomu co proto*). (Čermák, Hronek a kol. 2009b, s. 9–11)

Propoziční frazémy

Hlavním rysem propozičních frazémů je jejich větná povaha. Starší úzus sice vytvořil pro propoziční frazémy různá označení (mimo jiné například adagium, aforismus, anekdota, apothegma, axiom, citát, floskule, formule, fráze, gnóma, hádanka, heslo, hlavolam, hříčka, kalambúr, loci communes, maxima, motto, obrat, okřídlená slova, paremie, pozdrav, pořekadlo, poučka, pranostika, průpovídka, přísloví, rčení, rozpočítávadlo, říkadlo, říkanka, sentence, slogan, spojení slov, truismus, úsloví, wellerismus, zaklínadlo, zaříkadlo či zásada), avšak kvůli nedostatku termínů pro jiné výrazy a zároveň kvůli příliš mnoha nejasným výše uvedeným

termínům se v současnosti užívá především jednotný termín, který je založen na společné funkci frazémů, tj. větný frazém. (Čermák a kol. 2009, s. 1241)

1.4 Variabilita frazémů

I přesto, že některé frazémy mají formu, kterou nelze měnit (např. *zuby nehty*), jiné frazémy naopak lze používat v různých obměnách, i když ne zcela libovolných. V této podkapitole bude stanovena používaná terminologie základní formy i těchto různých obměn dle Jelínka, Kopřivové, Petkeviče a Skoumalové (2018, s. 152–153).

Základní slovníková podoba je označována jako invariant. Tato podoba bývá zpravidla tou nejčastěji užívanou (např. *házet flintu do žita*). (Jelínek a kol. 2018, s. 152)

Ve slovnících se rovněž uvádějí i různé varianty frazémů, které mohou být morfologické, lexikální i syntaktické. Nejvíce variant je v češtině zmapováno u slovesných frazémů, kde se mění zpravidla slovesný komponent – nejčastěji se jedná o změnu ve vidu (např. *zůstat/zůstatvat bez povšimnutí*) či záměnu za synonyma (např. *stočit se / svinout se do klubíčka*), nicméně změny mohou probíhat i na úrovni stylu (např. *zamést/zahladit za sebou / po sobě stopy*), regionálnosti (např. *dostat přes hubu/papulu*) či může být frazém rozšiřován (např. *do třetice / do třetice všeho dobrého*). (Jelínek a kol. 2018, s. 152)

Dále autoři uvádějí termín transformace, kdy se frazém užívá jako frazém jiného typu (např. nominalizace: *dosáhnout rekordu – dosažení rekordu*). Transformace někdy může způsobit i významový posun (např. *hrát si s někým jako kočka s myší – hra na kočku a myš*). (Jelínek a kol. 2018, s. 153)

Posledním typem obměny je modifikace, často odkazující na aktuální dění, která vzniká buď záměrně, nebo omylem, když je frazém použit nepřesně (např. *zmizet jako pára nad hrncem – rozplynout se jako pára nad vodou*). Původní frazém i jeho modifikaci je možné identifikovat pouze v případě, že jsou některé formální prvky frazému zachovány. (Jelínek a kol. 2018, s. 153)

2 Gramatické profily

Gramatický profil je termín, který L. Janda a Lyashevskaya definují jako „relativní frekvenční distribuci skloňovaných tvarů slova v korpusu“ (Janda a Lyashevskaya 2011, s. 717). Podle Lázničky umožňují gramatické profily „sledovat distribuční vlastnosti gramatických funkcí na relativně velkém objemu korpusových dat“ (Láznička 2019, s. 28).

Janda a Lyashevskaya ve svém článku zkoumaly na základě korpusových dat vidové dvojice v ruštině; sledovaly, zda jsou tvary tvořeny sufixem a prefixem (tradiční pohled), nebo pouze sufixem (Isačenkova hypotéza – Isačenko 1960), přičemž Isačenkovu hypotézu nepotvrdily. K výzkumu této problematiky jim sloužily právě gramatické profily slov, díky kterým mohly identifikovat slovesa, jež se chovají neobvykle a v rámci jejichž paradigmatu se některé tvary objevují velmi frekventovaně (Janda a Lyashevskaya 2011, s. 717). Ke zkoumání gramatických profilů a neobvyklého chování substantiv slouží nástroj GramatiKat, na jehož datech je tato práce založena.

Janda a Lyashevskaya vycházely z práce Divjak a Gries (2006) a jejich behaviorálních profilů, které používali při výzkumu devíti anglických synonymních sloves s významem *zkusit*. Právě behaviorální profily jim umožňovaly měřit míru podobnosti synonym. V této korpusové studii autoři určili 87 proměnných (ID tagů), tj. morfosyntaktické, syntaktické a sémantické vlastnosti, které tvoří behaviorální profily sloves (Divjak a Gries 2006, s. 23). Dle V. Jandy „jde v podstatě o informaci o tom, jaká slova se vyskytují v blízkosti zkoumaných lemmat, o odpovědi na následující otázky: jakého jsou slovního druhu, do jaké domény patří, jak často se s daným lemmatem pojí“ (Janda 2017, s. 12).

Gramatické profily pak Janda s Lyashevskou považují za podmnožinu behaviorálních profilů. Zatímco Divjak a Gries se zaměřovali na morfologické, syntaktické, sémantické a lexikální aspekty, Janda a Lyashevskaya pracují pouze s morfologickými vlastnostmi, k čemuž je inspirovalo tvrzení lingvisty Šteinfeldta, že se některá slovesa v některých tvarech používají častěji než jiná. (Janda a Lyashevskaya 2011, s. 718; Kuznetsova 2015, s. 15)

Podle Kuznetsové (2011, s. 16–17) jsou behaviorální profily problematické ze dvou důvodů. Za prvé jsou konkrétní proměnné vybírány výzkumníkem, a některá z důležitých proměnných tak nemusí být do výzkumu zahrnuta. Za druhé patří proměnné do různých rovin jazyka (morfologie, syntax, sémantika, lexikologie), které spolu sice úzce souvisí, avšak může docházet k tomu, že proměnným, které spolu korelují, může být přiřazena větší váha než ostatním, a může tak docházet ke zkreslení výsledků. Gramatické profily podle ní tyto výše

zmíněné problematické aspekty behaviorálních profilů řeší – zkoumání různých rovin jazyka nahrazuje zkoumání pouze morfologických vlastností, a výsledek tedy není závislý na výzkumníkovi. Metoda gramatických profilů dle Kuznetsové (2011, s. 28) zkoumá rozdíly ve významu pomocí zkoumání rozdílů ve formě. Právě tak probíhá i analýza v této práci – na základě rozdílů ve formě je zkoumána sémantika.

3 Frekvence pádů u substantiv v češtině

Jak již bylo zmíněno výše, gramatické profily slov tvoří frekvenční distribuce tvarů. Frekvencí pádů v češtině (neboli distribucí tvarů) se poprvé zabývali Jelínek, Bečka a Těšitelová v rámci kvantitativního výzkumu, jehož výstupem byla kniha *Frekvence slov, slovních druhů a tvarů v českém jazyce* (1961). Mluvnice vydané před tímto rokem, tj. *Mluvnice spisovné češtiny* (Trávníček 1948) a *Česká mluvnice* (Havránek a Jedlička 1960; první vydání 1951), frekvenci pádů vůbec nezmiňují. Cílem autorů *Frekvence slov, slovních druhů a tvarů v českém jazyce* bylo pracovat s co nejrozmanitějšími texty, celkem 75 pramenů tedy bylo roztríděno do osmi skupin: beletrie, poezie, literatura pro mládež, dramata, odborná literatura, žurnalistika, vědecká literatura a mluvené projevy uveřejněné tiskem. Použité prameny jsou z let 1926 až 1949, je tedy mapován jazyk meziválečný, válečný a poválečný. (Jelínek, Bečka a Těšitelová 1961, s. 19–22)

Vzorek pro určení frekvence pádů byl složen z 116 807 substantiv. Co se týče čísla, autoři uvádí, že singulár výrazně převažuje nad plurálem (konkrétně 75,81 % : 24,19 %). Tabulka 1 uvádí procentuální frekvenci pádů u všech 14 tvarů, a to ze všech osmi zkoumaných skupin. (Jelínek, Bečka a Těšitelová 1961, s. 86)

S1	S2	S3	S4	S5	S6	S7
22,52 %	16,92 %	3,43 %	15,01 %	1,76 %	8,52 %	7,66 %
P1	P2	P3	P4	P5	P6	P7
5,94 %	7,56 %	1,03 %	4,87 %	0,19 %	2,37 %	2,22 %

Tabulka 1: Frekvence pádů u substantiv v procentech dle Jelínka, Bečky a Těšitelové (1961, s. 94–95)

Frekvence pádů se však liší v rámci jednotlivých skupin. Autoři například uvádějí, že zatímco nominativ je hojně zastoupen v uměleckých textech, genitiv spíše v odborné literatuře, což vysvětlují vysokým počtem neshodných genitivních přívlastků v naučné próze. Stejně tak akuzativ má vysokou frekvenci v uměleckých textech, lokál v odborných a publicistických textech. Celková frekvence vokativu je zkrácena dramaty, v nichž se vyskytuje nadprůměrně. (Jelínek, Bečka a Těšitelová 1961, s. 86)

Frekvenci pádů dále zmiňuje Marie Těšitelová a kolektiv v rámci monografie *Kvantitativní charakteristiky současné češtiny* (Těšitelová a kol. 1985), jež byla doplňkem akademické *Mluvnice češtiny*, avšak pouze zde shrnuje výsledky z výše uvedené práce z 60. let. *Mluvnice češtiny* (Komárek a kol. 1986) ani *Příruční mluvnice češtiny* (Karlík a kol. 1996), tj. mluvnice

vycházející v průběhu 80. a 90. let 20. století, se frekvenci pádů nevěnují, kdežto *Čeština – řeč a jazyk* (první vydání z roku 1996) zmiňuje, že nejfrekventovanějšími pády jsou nominativ, genitiv a akuzativ, zatímco nejnižší frekvenci má vokativ (Čechová a kol. 2011, s. 165).

Pádové distribuci substantiv se podrobněji věnují například *Statistiky češtiny* (Bartoň a kol. 2011), statistická příručka vycházející z korpusu SYN2005, tj. reprezentativního korpusu složeného z publicistických textů z let 2000–2004, z odborné literatury z let 1990–2004 a beletrie. Frekvence pádů je v tomto případě zkoumána na 31 315 440 slovních tvarech substantiv a 529 936 lemmatech. Procentuální rozložení frekvence pádů je zaznamenáno v tabulce 2. Rozložení singuláru a plurálu je téměř shodné s údaji Jelínka, Bečky a Těšitelové – v korpusu SYN2005 (Čermák a kol. 2005) je doloženo 75,2 % singulárových tvarů a 24,8 % plurálových tvarů. (Bartoň a kol. 2011, s. 134)

S1	S2	S3	S4	S5	S6	S7
23,0 %	18,5 %	2,7 %	14,2 %	0,35 %	9,7 %	6,6 %
P1	P2	P3	P4	P5	P6	P7
6,0 %	8,9 %	0,92 %	5,0 %	0,001 %	2,1 %	1,9 %

Tabulka 2: Frekvence pádů u substantiv v procentech dle Bartoně a kol. (2009, s. 134)

Autoři této publikace rovněž uvádějí frekvenci pádů dle jednotlivých typů textů – v beletrii, oborové literatuře a publicistice, avšak rozdílů nekomentují. I z jejich výsledků je však patrné, že genitiv je výrazněji zastoupen v oborové literatuře, v níž má dokonce větší frekvenci než nominativ. (Bartoň a kol. 2011, s. 134)

Frekvenci pádů dále zmiňuje *Mluvnice současné češtiny* (Cvrček a kol. 2010), která vyšla v roce 2010 a podle informací v úvodu popisuje jazyk tak, jak skutečně vypadá, nikoli jak by měl vypadat. Mluvnice je založena na jazykových korpusech ČNK. Tabulka 3 shrnuje informace uvedené v této mluvnici, které jsou oproti jiným zdrojům přibližnější. Z tabulky bohužel není jasné, ze kterého jazykového korpusu byla data čerpána. I v tomto případě je uveden poměr 75 % singulárových tvarů ku 25 % plurálovým. Autoři dále dodávají, že poměry jsou závislé na typu substantiva, což dále rozvádí (například apelativa označující živé bytosti mají nadprůměrně frekventovaný nominativ a podprůměrně frekventovaný lokál, v přátelském hovoru je navíc poměrně častý vokativ; nejfrekventovanějším pádem u apelativ označujících neživé entity je genitiv a akuzativ). (Kodýtek 2010, s. 141)

S1	S2	S3	S4	S5	S6	S7
24 %	20 %	3 %	10 %	< 1 %	10 %	7 %
P1	P2	P3	P4	P5	P6	P7
> 5 %	10 %	1 %	> 5 %	< 0,1 %	2 %	< 2 %

Tabulka 3: Frekvence pádů u substantiv v procentech dle Kodýtky (2010, s. 141)

Od počátku 21. století můžeme frekvenci pádů vyvozovat z velkých jazykových korpusů. Na frekvenci pádů v SYN2015 odkazuje *Velká akademická gramatika spisovné češtiny* (Štícha a kol. 2018), avšak čísla jsou uvedena pro singulár i plurál dohromady. Tabulka 4 tedy čerpá z dat přímo z korpusu SYN2015 (Křen a kol. 2015).

S1	S2	S3	S4	S5	S6	S7
22,17 %	18,04 %	2,80 %	15,61 %	0,30 %	9,70 %	6,77 %
P1	P2	P3	P4	P5	P6	P7
5,37 %	8,68 %	0,90 %	5,60 %	0,01 %	2,11 %	1,93 %

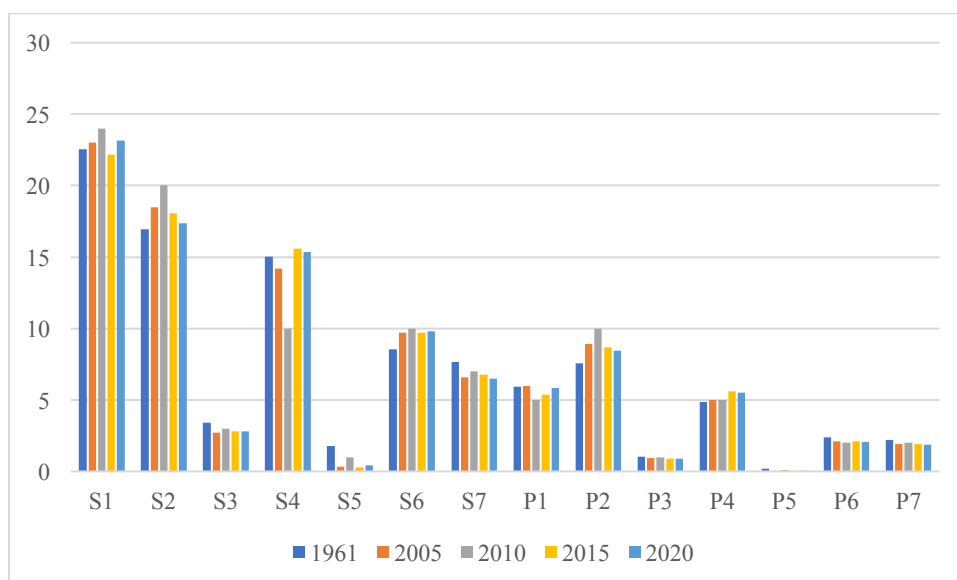
Tabulka 4: Frekvence pádů u substantiv v procentech v SYN2015

V současné době je k dispozici další reprezentativní korpus psané češtiny SYN2020 (Křen a kol. 2020). Frekvence pádů u substantiv v tomto korpuse je zaznamenána v tabulce 5. Rozložení singuláru a plurálu v rámci korpusů SYN2015 a SYN2020 se liší pouze nepatrně a shoduje se s výše uvedenými zdroji (v SYN2015 75,41 % : 24,59 % a v SYN2020 75,35 % : 24,65 %). Zastoupení jednotlivých pádů je v obou korpusech rovněž téměř shodné a odpovídá také informacím ve starších studiích. V singuláru je nejčastějším pádem nominativ, dále genitiv, akuzativ, lokál, instrumentál, dativ a vokativ. V plurálu je zase nejfrekventovanější genitiv, za nímž následuje nominativ, akuzativ, lokál, instrumentál, dativ a vokativ, přičemž genitiv, nominativ a akuzativ plurálu mají vyšší frekvenci než dativ a vokativ singuláru.

S1	S2	S3	S4	S5	S6	S7
23,15 %	17,34 %	2,82 %	15,34 %	0,41 %	9,81 %	6,48 %
P1	P2	P3	P4	P5	P6	P7
5,83 %	8,46 %	0,91 %	5,49 %	0,04 %	2,07 %	1,85 %

Tabulka 5: Frekvence pádů u substantiv v procentech v SYN2020

Graf 1 srovnává všechna výše uvedená data, tedy z *Frekvence slov, slovních druhů a tvarů v českém jazyce (1961)*, ze *Statistik češtiny (SYN2005)*, *Mluvnice současné češtiny (2010)* a jazykových korpusů SYN2015 a SYN2020. Je z něj patrné, že v průběhu let se frekvence jednotlivých tvarů příliš neliší. Pouze *Mluvnice současné češtiny* uvádí menší frekvenci akuzativu v singuláru než ostatní zdroje. Všechny zdroje se dále shodují na čtyřech nejčastějších tvarech: nominativ, genitiv, akuzativ a lokál v singuláru. Dále mezi sebou kolísají instrumentál singuláru a genitiv, nominativ a akuzativ plurálu, po nichž ve všech zdrojích následuje dativ singuláru, lokál plurálu a instrumentál plurálu. Ve všech zdrojích kromě toho nejstaršího dále následuje dativ plurálu (*Frekvence slov, slovních druhů a tvarů v českém jazyce* uvádí vokativ singuláru, který v ostatních zdrojích následuje). Nejméně frekventovaným tvarem je pak vokativ plurálu.



Graf 1: Porovnání frekvence pádů u substantiv v procentech dle různých zdrojů

4 Jazykové korpusy a jejich nástroje a funkce využívané při analýze

V této kapitole budou popsány jazykové korpusy a jejich nástroje a funkce, které jsou použity při analýze.

4.1 Korpus SYN2015

Při analýze jsou využity dva jazykové korpusy – SYN2015 (Křen a kol. 2015) a SYNv11 (Křen a kol. 2022). Korpus SYN2015 byl zvolen kvůli tomu, že na jeho datech je založen nástroj GramatiKat, podle něhož byla vybrána lemmata k analýze. Jedná se o synchronní, reprezentativní korpus psané češtiny, o velikosti 120 milionů tokenů. Je součástí řady SYN, která je vydávána vždy po pěti letech a vždy se zaměřuje na co nejaktuálnější texty; většina textů pochází z období 2010–2014. Jelikož se jedná o korpus reprezentativní, tři hlavní textové makrotypy, tj. beletrie, publicistika a oborová literatura, jsou do korpusu zahrnuty ve stejném poměru, tedy po třetinách. Toto rozložení nereflektuje skutečný poměr výskytu textů, ale zajišťuje rozmanitost textů obsažených v korpusu. Beletristické texty přitom nesmí být starší než 75 let od prvního vydání a 25 let od použitého vydání, oborová literatura musí být poprvé vydána maximálně před 25 lety a publicistika musí pocházet z předchozích pěti let, v tomto případě tedy z let 2010–2014. (Cvrček a Richterová 2022)

4.2 Korpus SYNv11

Kvůli velikosti korpusu se analýza dále opírá o největší dostupný obecný, synchronní korpus – SYNv11, který obsahuje až 6 miliard slov. Tento verzovaný korpus spojuje všechny korpusy řady SYN (tj. SYN2000, SYN2005, SYN2006PUB, SYN2009PUB, SYN2010, SYN2013PUB, SYN2015 a SYN2020) a další dosud nezveřejněné publicistické texty. Vzhledem k těmto textům a k publicistickým korpusům není korpus reprezentativní a v jeho složení výrazně převažuje publicistika (publicistika tvoří až 89 % korpusu oproti 8 % oborové literatury a 3 % beletrie). (Cvrček a Richterová 2023a)

4.3 GramatiKat

GramatiKat, na jehož datech je tato práce založena, je korpusový nástroj pro výzkum gramatických kategorií a gramatických profilů (Kovářiková a Kovářík 2021). Nástroj se řídí principy otevřené vědy, v rámci níž se v posledních letech rozvíjí vývoj nástrojů, které nabízejí předem zpracovaná data, na základě nichž mohou výzkumníci provádět vlastní výzkumy.

Výhodou otevřené vědy je mimo jiné možnost ověřitelnosti výzkumu a možnost navázání na předchozí výzkumy. Tento trend je patrný i v korpusové lingvistice – v posledních pěti letech bylo v rámci ČNK publikováno několik korpusových nástrojů, vedle GramatiKatu například Slovo v kostce (Machálek 2019), Calc (Cvrček 2019) či Mapka (Golášová a kol. 2023). Data jsou tedy v GramatiKatu již zpracována, statisticky vyhodnocena a vizualizována v grafech a tabulkách. (Kováříková 2021b, s. 532–533, s. 542)

Nástroj poskytuje informace o distribuci slovních tvarů jak u celého slovního druhu (v současné době se jedná pouze o substantiva, do budoucna má být rozšířen o adjektiva a verba), tak u konkrétních lemmat, a to v porovnání právě s celou třídou slov stejného slovního druhu. Díky dostupným interaktivním tabulkám je dále možné v aplikaci vyhledávat lemmata s anomální distribucí tvarů, tj. lemmata, která mají neobvykle vysokou frekvenci některého tvaru či tvarů, či naopak velmi nízkou až nulovou frekvenci, přičemž tyto dva jevy spolu mnohdy souvisí (Kováříková a Kovářík 2021, Cvrček a Richterová 2021b).

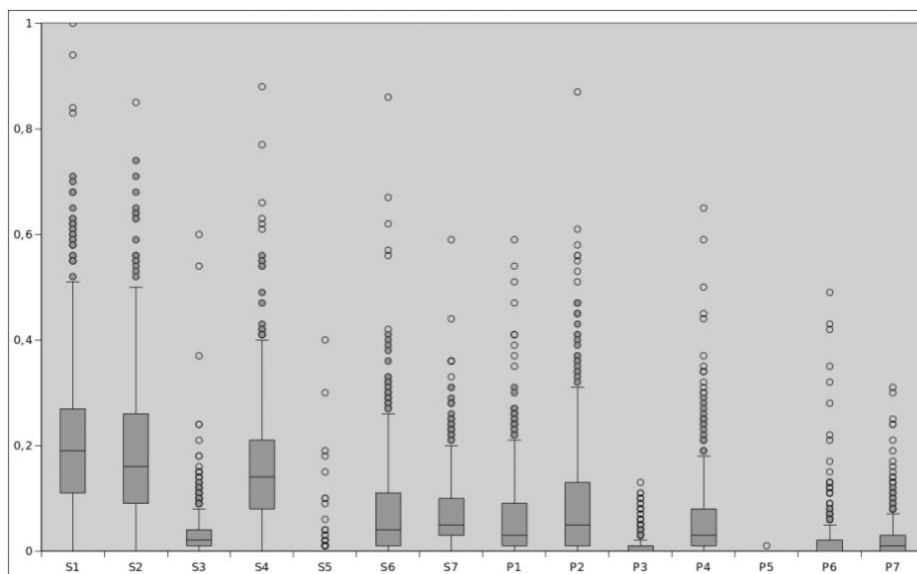
Z dat GramatiKatu vyplývá, že pouze u přibližně 25 % substantiv s frekvencí alespoň 100 je zastoupeno všech 14 tvarů a zároveň žádný z nich není neobvykle frekventovaný. Přibližně u 50 % substantiv je totiž neobvykle frekventovaný alespoň jeden tvar a 50 % substantiv postrádá alespoň jeden tvar paradigmatu, přičemž se tyto skupiny překrývají. (Kováříková 2021b, s. 538) Kováříková (2021b, s. 541) upozorňuje, že jelikož jsou anomálie v distribuci pádů často způsobeny omezenými kolokacemi, jeden z možných dalších výzkumů by bylo vhodné zaměřit na víceslovné jednotky. Právě toto tvrzení inspirovalo téma této diplomové práce.

Nástroj je založen na datech z korpusů ČNK, konkrétně na následujících korpusech:

- SYN2015 (synchronní, reprezentativní korpus psaného jazyka, především s texty z let 2010–2014),
- ORAL verze 4 (korpus neformální mluvené češtiny),
- subkorpusy češtiny a slovenštiny korpusu InterCorp verze 13 (vytvořeny pro účely konference Slovko 2021 pro srovnání českých a slovenských substantiv) (Kováříková 2021b, s. 533–534).

Z těchto korpusů byla vybrána pouze substantiva, jejichž absolutní frekvence byla v daném korpusu alespoň 100. Procentuální zastoupení sedmi pádů ve dvou číslech je založeno na přibližně 15 tisících substantivech z korpusu SYN2015, přičemž při výpočtech má každé lemma stejnou váhu, bez ohledu na frekvenci, a vysoce frekventovaná lemmata tak nezkrslují

výsledky. Právě z těchto dat byl vyvozen krabicový graf znázorňující běžnou distribuci tvarů substantiv, který je u každého hledaného lemmatu totožný (obrázek 1). (Kováříková a Kovářik 2021, Kovářiková 2021a, s. 75)



Obrázek 1: Distribuce tvarů českých substantiv na základě dat z korpusu SYN2015 (Kováříková 2021b, s. 535)

Z tohoto grafu lze vyčíst nejen standardní chování českých substantiv, ale i hodnotu, od které jsou výskyty v tomto tvaru neobvyklé. Tato hodnota je u každého pádu jiná. Hodnoty 25. percentilu, mediánu, 75. percentilu a hranice neobvykle vysokých hodnot jednotlivých tvarů jsou v procentech znázorněny v tabulce 2. (Kováříková 2021b, s. 535)

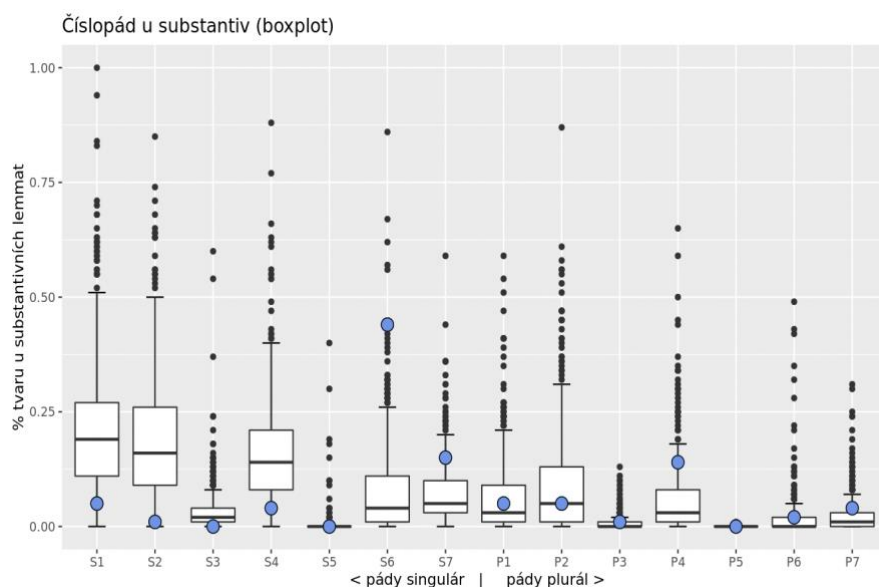
Tvar	S1	S2	S3	S4	S5	S6	S7
25. percentil	12,5 %	7,4 %	0,7 %	7,8 %	0 %	0,9 %	2,8 %
Medián	19,3 %	14,8 %	1,8 %	14,5 %	0 %	3,9 %	5,5 %
75. percentil	28 %	25 %	4 %	22,6 %	0 %	10,5 %	9,8 %
Neobvykle vysoká hodnota	57,4 %	54,9 %	9,3 %	48,7 %	0 %	25,2 %	24,1 %
Tvar	P1	P2	P3	P4	P5	P6	P7
25. percentil	0,5 %	0,5 %	0 %	0,3 %	0 %	0 %	0 %
Medián	3,5 %	4,2 %	0,2 %	2,8 %	0 %	0,4 %	0,9 %
75. percentil	9,8 %	11,8 %	1 %	7,8 %	0 %	1,7 %	3 %
Neobvykle vysoká hodnota	24,1 %	29 %	2,5 %	19,1 %	0 %	4,2 %	7,4 %

Tabulka 6: Procentuální hodnoty z boxplotů znázorňujících distribuci tvarů (Kováříková 2021b, s. 536)

Podle Kovářikové (2021b, s. 534) je však třeba brát v potaz, že „procento hodnot v každé gramatické kategorii může být ovlivněno různými faktory, především velikostí a složením korpusu a četností lemmatu“. Protože je však použitý korpus SYN2015 velký, reprezentativní a vyvážený, zajišťuje poměrně vysokou spolehlivost výsledků.

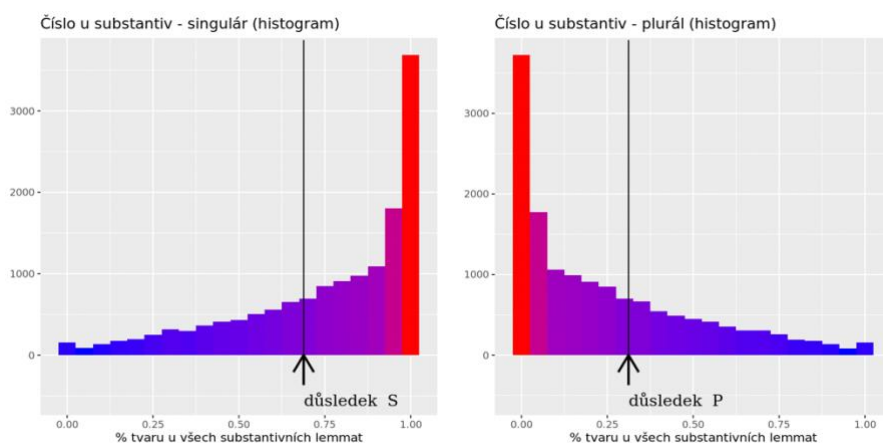
Jelikož jsou zatím k dispozici pouze substantiva, nástroj nabízí data o gramatické kategorii čísla, pádu a jmenného rodu, konkrétně lze zobrazit informace pouze o čísle, dále o čísle a pádu (tzv. číslopád) a o rodě, čísle a pádu (tzv. rodočíslopád). Tyto informace jsou po zadání lemmatu zobrazeny v několika grafech – v krabicových grafech a histogramu. Krabicové grafy znázorňují běžnou distribuci substantiv – střední, krabicovou část ohraničuje 1. a 3. kvartil, přičemž čára uprostřed označuje medián. Dále je vyobrazeno minimum a maximum, tedy nejnižší hodnota pod 1. kvartilem a nejvyšší hodnota nad 3. kvartilem. Odlehlé hodnoty, tj. hodnoty překračující 1,5násobek mezikvartilového rozpětí nad třetím kvartilem a pod prvním kvartilem, označují anomálie. Hodnoty pod prvním kvartilem se však v GramatiKatu nevyskytují, nestandardní jsou tedy nulové výskyty některých tvarů, které jsou způsobeny buď mezerou v paradigmatu daného lemmatu, nebo absencí v korpusu. (Kovářiková a Kovářik 2021, Kovářiková 2021b, s. 534)

Na pozadí tohoto běžného chování českých substantiv se dále zobrazují gramatické profily jednotlivých lemmat. Ty v grafu označují modré tečky. Na příkladu obrázku 2 lze ukázat, že lemma *důsledek* vykazuje neobvykle vysokou frekvenci lokálu singuláru. Tyto grafy lze znázorňovat jak u čísla a číslopádu, tak i u rodočíslopádu. (Kovářiková a Kovářik 2021, Kovářiková 2021b, s. 536)



Obrázek 2: Distribuce tvarů lemmatu *důsledek* (Kovářiková a Kovářik 2021)

V případě gramatické kategorie čísla jsou v GramatiKatu dále dostupné orientační histogramy, které znázorňují absolutní frekvenci lemmat spadajících do dané kategorie a procentuální zastoupení tvarů konkrétního lemmatu (černá šipka). Na příkladu lemmatu *důsledek* na obrázku 3 lze ukázat, že singulárové formy tohoto lemmatu se vyskytují asi v 69 %, zatímco plurálové pouze přibližně ve 31 %. Do této kategorie spadá asi 700 ze zachycených 15 tisíc lemmat. (Kováříková a Kovářik 2021)



Obrázek 3: Lemma *důsledek*, distribuce v singuláru a plurálu (Kováříková a Kovářik 2021)

U čísla a číslopádu lze dále zobrazit interaktivní tabulky, v nichž lze vybírat horní odlehlé hodnoty a spodní extrém, tj. anomálie v distribuci tvarů. Dle jednotlivých tvarů je pak lze třídit. Tabulka znázorňující horní odlehlé hodnoty obsahuje lemmata, jejichž frekvence ve vybraném tvaru je nad 1,5násobkem 75. percentilu, zatímco tabulka znázorňující spodní extrém obsahuje lemmata, jež se v daném tvaru nevyskytují. Tabulka 7 znázorňuje, kolik anomálních lemmat je u každého číslopádu v GramatiKatu dostupných – tedy kolik lemmat má horní odlehlé hodnoty a kolik lemmat se v některém tvaru nevyskytuje vůbec (spodní extrém neboli nulový výskyt). (Kováříková a Kovářik 2021)

Tvar	S1	S2	S3	S4	S5	S6	S7
Horní odlehlé hodnoty	769	361	1 169	321	839	809	852
Nulový výskyt	255	311	2 593	342	13 714	2 448	723
Tvar	P1	P2	P3	P4	P5	P6	P7
Horní odlehlé hodnoty	852	671	1 479	788	20	992	1 002
Nulový výskyt	3 391	3 451	8 842	3 494	14 533	7 499	6 005

Tabulka 7: Počet horních odlehlých hodnot a spodních extrémů dle jednotlivých tvarů (Kováříková a Kovářik 2021)

4.4 FRANTA

Dalším korpusovým nástrojem využívaným při analýze je nástroj FRANTA (Frazémová anotace a textová analýza). Tímto nástrojem jsou frazeologicky označovány některé korpusy ČNK, konkrétně korpusy SYN od verze 4. Jedná se o program, který v korpusu vyhledává víceslovné jednotky (tj. frazémy a ustálené kolokace) z předem daného seznamu, který čítá přibližně 40 tisíc položek a vychází především z Čermákova *Slovníku české frazeologie a idiomatiky*. Nástroj tedy v současné době neumí rozpoznat veškeré víceslovné jednotky, nýbrž jen ty, které jsou zahrnuté v tomto seznamu. Jelikož jsou víceslovné jednotky rozpoznávány automaticky, bez znalosti syntaktické struktury a valence, mohou nastávat problémy s identifikací víceslovných jednotek. Program rovněž neumí rozeznat frazeologický a doslovný význam. Některá spojení se totiž mohou vyskytovat v obou významech (např. *lízat si rány*). (Cvrček a Richterová 2021c)

Víceslovné jednotky jsou označeny pozičními atributy `col_lemma` (collocation lemma, kolokační lemma) a `col_type` (collocation type, kolokační typ). Díky atributu kolokační lemma lze v korpusu vyhledávat konkrétní lemmata (např. lemma *večer*), víceslovné jednotky (např. *ráno_moudřejší_večera*) i víceslovné jednotky dle slovního druhu (např. *N_jako_hrom* vyhledá *díra jako hrom, skandál jako hrom* atd.), případně vyhledávat pomocí zástupných slov (např. *nějaký – v_nějakém_slova_smyslu* vyhledá *v pravém slova smyslu, v dobrém slova smyslu* atd.). Poziční atribut kolokační typ umožňuje vyhledávat víceslovné jednotky na základě jejich typu a závislosti. První písmeno určuje hlavní (H) či závislé (Z) slovo víceslovné jednotky. Hlavní slovo je ve víceslovné jednotce vždy jedno, zpravidla se jedná o poslední slovo. Zbylé jednotky jsou označeny jako závislé (např. *tma [Z] jako [Z] v pekle [H]*). Druhé písmeno pak určuje typ víceslovné jednotky. Jsou rozlišovány tyto typy:

- neslovesné kolokace (K),
- víceslovné spojky (J),
- přísloví, okřídlená rčení, citace apod. (M),
- přirovnání (P),
- větné výrazy (S),
- slovesné kolokace (V). (Cvrček a Richterová 2021c)

4.5 Kolokace v korpusu

Při vyhledávání frazémů je dále nápomocné vyhledávání kolokací v korpusu prostřednictvím korpusového manažeru KonText. Kolokace v korpusové lingvistice označuje

„smysluplné, ustálené, syntagmatické spojení dvou (nebo víc) slovních tvarů (někdy celých lexémů) v blízkém kontextu“ (Cvrček a Richterová 2023b). Ustálenost se pak měří frekvencí v korpusu a mírou komptability členů kolokace s jinými jednotkami v korpusu (Cvrček a Richterová 2023b).

Pro vyhledávání kolokací se používají statistické asociační míry (association measures). Ty usouvztažňují frekvenci jednotek vstupujících do kolokace, frekvenci celého spojení a velikost korpusu. Asociační míry, které lze v korpusech ČNK použít, jsou T-score, MI, MI3, log likelihood, min. sensitivity, logDice, $Mi.log_f$ a relative freq, přičemž každá z nich považuje za významné jiné kolokace, a je proto vhodné výsledky různých asociačních měř porovnávat. Toto automatické vyhledání kolokací vyhledá spíše kolokační kandidáty, výsledky je pak nutné ručně projít a určit, které kandidáty na kolokace lze považovat za skutečné kolokáty. Tento seznam kolokátů je označován jako kolokační profil či paradigma. (Cvrček a Richterová 2021a)

Při výpočtu kolokací se určuje rozsah, ve kterém jsou kolokace hledány, tj. pozice nalevo a napravo od hledaného slova. Pro většinu výzkumů je doporučováno základní nastavení, tedy od -3 do +3 (tři pozice nalevo a tři pozice napravo od hledaného slova). (Cvrček a Richterová 2021a)

Asociační míra MI-score preferuje nízko frekventovaná slova, což mnohdy může vést ke zkreslení výsledků (slova s velmi nízkou frekvencí pak může jako kolokaci označit zcela náhodně), zároveň je však při vyhledávání kolokací možné nastavit spodní hranici frekvence. Naopak na vysoké frekvenci je založena asociační míra T-score, a smysluplné kolokace se v seznamu kolokátů proto objevují až níže. Frekvenci i nenáhodnost výskytu zohledňuje asociační míra logDice, proto je vhodné výsledky třídit primárně podle ní. (Cvrček a Richterová 2021a)

Asociační míra logDice, která je využívána, usouvztažňuje frekvenci jednotlivých slov a celého spojení, nikoli velikost korpusu. K jejímu výpočtu slouží následující vzorec:

$$\logDice(xy) = 14 + \log_2 \frac{2f(xy)}{f(x) + f(y)}$$

Tato rovnice poměří frekvenci celého spojení, $f(xy)$, s průměrem frekvencí jednotlivých slov, $f(x)$, $f(y)$. Asociační míra logDice nabývá hodnot od $-\infty$ do 14. (Cvrček a Richterová 2019)

5 Slovníky využívané při analýze a jejich přístup k frazeologii

V této kapitole budou stručně popsány slovníky, v nichž jsou frazémy vyhledávány, a jejich přístup k frazeologii.

5.1 Slovník české frazeologie a idiomatiky

Hlavním zdrojem je frazeologický slovník, a to Slovník české frazeologie a idiomatiky (SČFI), v němž byla česká frazeologie zpracována nejkomplexněji. Na jeho přípravě se již od 60. let podíleli F. Čermák, J. Hronek, J. Machač, oddělení češtiny pro cizince FF UK a oddělení současné lexikografie ÚJČ AV ČR. SČFI má čtyři svazky – přirovnání (1983), výrazy neslovesné (1988), výrazy slovesné (1994) a výrazy větné (2009). S vydáním čtvrtého dílu byly nově vydány i tři předchozí svazky, při jejichž tvorbě se autoři již mohli opřít o data ČNK. Celkem bylo do slovníku zahrnuto asi 37 tisíc hesel – přibližně 2,4 tisíce přirovnání, 4,8 tisíc nominálních, adverbálních a gramatických výrazů, 9,9 tisíc verbálních frazémů, 10 tisíc verbálních kvazifrazémů a 10 tisíc větných frazémů. V roce 2016 pak byl vydán pátý díl, sjednocující díly předchozí, který nově frazémy prezentuje dle významu, tj. onomaziologický slovník. (Hladká 2017; Hladká 1995; Čermák 2007a, s. 246; Šichová a Šemelík 2017, s. 491–492)

Slovník zachycuje frazeologii a idiomatiku druhé poloviny 20. století, přičemž frazémy a idiomy vymezuje na základě Čermákova pojetí. Kvůli rysu anomální sémantické či formální povahy komponentů tak některé frazémy dokládá zcela poprvé, například ustálená spojení synsémantik. Do slovníku naopak nejsou zahrnuty lexikální frazémy. (Hladká 2017)

Zdrojem dat byly dosud vydané výkladové a frazeologické slovníky, frazeologická kartotéka lexikálního archivu ÚJČ AV ČR, excerpcie literárních děl vydaných po roce 1960, excerpcie záznamů současného mluveného jazyka, ankety a přímý výzkum. (Hladká 1995)

Heslo ve slovníku obvykle obsahuje lemma a zachycení variantnosti, stylovou informaci, gramatickou informaci, strukturní transformaci, valenci, kontext a funkce, význam, exemplifikaci, poznámky, synonyma, opozita, příbuzné výrazy a odkaz na sémantický rejstřík a jinojazyčné ekvivalenty. (Hladká 2017)

SČFI je elektronicky dostupný ve webové aplikaci DEBDict, vyvinuté Adamem Rambouskem z Fakulty informatiky Masarykovy univerzity, která zpřístupňuje slovníky českého jazyka. Právě tato webová aplikace je při následné analýze nápomocná. (Kalivodová nedatováno)

5.2 Akademický slovník současné češtiny

Akademický slovník současné češtiny (ASSČ), jehož cílem je zaznamenat současnou češtinu, vzniká od roku 2012 na oddělení současné lexikologie a lexikografie ÚJČ AV ČR. K prosinci 2023 jsou zveřejněna hesla od písmene A do písmene G. ASSČ je slovník středního rozsahu, tj. po dokončení by měl obsahovat 120–150 tisíc hesel. Na rozdíl od starších slovníků je při jeho tvorbě využíváno poznatků a nástrojů korpusové lingvistiky, primárním zdrojem dat jsou především data z korpusů ČNK (O slovníku 2017–2023). Korpusy se při tvorbě slovníku využívají nejen například k určení stylistické příznakovosti, zjištění synonymie a antonymie či zjištění valence a kolokability, ale také k „identifikaci víceslovných lexikálních jednotek, k rozhodnutí o jejich zařazení a způsobu zpracování, včetně stanovení reprezentativního lemmatu, variantnosti, významu a výběru exemplifikace (platí zejména pro frazémy)“ (Kochová a Opavská 2016, s. 17–18).

ASSČ rozlišuje tři druhy víceslovných jednotek – sousloví, frazémy a víceslovné předložkové, spojkové, částicové a citoslovečné výrazy. Frazémy vymezuje jako „ustálená a reprodukovatelná spojení se sekundární nominativní funkcí, jejichž celkový význam není plně odvoditelný z významu jednotlivých složek“ (Kochová a Opavská 2016, s. 167), jejichž hlavními rysy jsou ustálenost, reprodukovatelnost, denotativní vágnost, významová neurčitost, expresivita, hodnotící funkce, přenesenost, obraznost a idiomaticita.

Do slovníku jsou pak zahrnuty frazémy frekventované a uzualizované, přičemž výběr je založen jak na konfrontaci frazémů v SČFI a SSČ, tak na korpusu SYN, dalších materiálových zdrojích, jako je internet či archiv společnosti Newton Media, a uživatelském povědomí. Zachycovány jsou všechny typy frazémů, tj. přirovnání, jmenné frazémy, slovesné frazémy a větné frazémy. U jednotlivých hesel, tedy v rámci jednoslovného hesla (komponentu), je pro ně vyhrazena frazémová zóna, která zahrnuje lemma, stylistickou charakteristiku, výklad a exemplifikaci. (Kochová a Opavská 2016, s. 174–175)

6 Metodologie

V této kapitole budou nejprve popsány cíle analýzy, následně to, jak byla vybrána použitá data, a nakonec postup analýzy a podoba popisu jednotlivých pádů.

6.1 Cíle analýzy

Cílem této analýzy je především vyhledat frazémy na základě dat poskytovaných online aplikací GramatiKat, která dokáže identifikovat substantivní lemmata s anomální distribucí tvarů, a následně je popsat a klasifikovat. Díky nové metodě vyhledávání frazémů by výsledkem mohly být nejen samotné frazémy, v některých případech dosud nezachycené ve frazeologických slovnících, ale i celé skupiny frazémů s podobnými vlastnostmi. Pro potřeby práce byla zvolena následující hlavní výzkumná otázka:

Ukazují gramatické anomálie substantiv (tj. nezvykle vysoká frekvence jednoho či více tvarů v rámci morfologického paradigmatu) na formálně anomální víceslovné lexémy (frazémy)?

Pokud z analýzy vyplyne kladná odpověď na tuto hlavní výzkumnou otázku, jsou na místě i další dvě vedlejší výzkumné otázky:

V jakých pádech se nejčastěji projevuje souvislost gramatické anomálie substantiv s frazémy?

Jaké typy frazémů se v rámci jednotlivých pádů objevují nejčastěji?

Na tomto místě je třeba zároveň zdůraznit, že cílem práce není data zpracovat v jejich úplnosti. Vzhledem k rozsahu práce to jednak není možné, a jednak je cílem provést pouze sondu a zjistit, zda spolu vůbec gramatická anomálie a anomální povaha frazémů souvisejí. V současné době totiž vzniká druhá verze GramatiKatu, na které by pak mohla být provedena podobná analýza v její úplnosti.

6.2 Použitý materiál

Jelikož se tedy jedná o sondu, byl pro tuto práci vybrán pouze vzorek z dat GramatiKatu. Vzhledem k zadání práce jsou analyzována pouze lemmata anomální v horních odlehlých hodnotách (nikoli lemmata s nulovými výskyty). Data byla dále omezena pouze na tvary v singuláru, a to pouze na část z nich. Kvůli možnosti srovnání jednotlivých pádů byla zvolena hranice prvních 20 % u každého pádu. Jelikož je seznam řazen abecedně, byla do seznamu zařazena i slova za hranicí 20 % se stejnou hodnotou výskytu jako u lemmatu, které se nachází

přesně na zvolené hranici. V tabulce 8 je uvedeno, kolik lemmat anomálních v horních odlehlých hodnotách je u každého pádu celkem přítomno, kolik z nich bylo zařazeno do této analýzy a na kolika procentech byla stanovena spodní hranice pro zařazení lemmatu do vzorku. Do vzorku jsou započítána i chybně označovaná lemmata a lemmata, která jsou v GramatiKatu vlivem chyby uvedena dvakrát (stalo se tak ve třech případech). Celkem je tedy analyzováno 1 102 lemmat, jejich kompletní soupis je uveden v Příloze 1.

	S1	S2	S3	S4	S5	S6	S7	Součet
Celkem	769	361	1 169	321	839	809	852	5 120
Vzorek	162	75	252	66	201	169	177	1 102
Hranice procent	0,72	0,67	0,17	0,7	0,04	0,47	0,39	

Tabulka 8: Celkový počet lemmat anomálních v horních odlehlých hodnotách u jednotlivých pádů a velikost zvoleného vzorku

6.3 Postup

Pro analýzu byl zvolen následující postup: lemmata zařazená do vzorku (o výběru viz podkapitola Použitý materiál) jsou vyhledávána ve dvou korpusech – v SYN2015 a SYNv11. Korpus SYN2015 byl zvolen kvůli tomu, že na jeho datech je založen nástroj GramatiKat, jenž ovlivňoval výběr lemmat pro analýzu. Korpus SYNv11 je zase největším dostupným korpusem, a mohou v něm tedy být případně vyhledány další frazémy, jejichž součástí jsou analyzovaná lemmata. Slova jsou v korpusech vždy vyhledávána podle daného tvaru, ve kterém jsou anomální. Například v případě lemmatu *štych*, které je anomální v singuláru lokálu, se jedná o následující CQL dotaz, který lemma vyhledává jako substantivum v lokálu singuláru:

[lemma="štych" & tag="N..S6.*"]).

Následně je využit nástroj FRANTA, a to jak ve vytvořeném subkorpusu ze SYNv11, konkrétně v korpusu SYN2015, tak v celém SYNv11. Pozičním atributem pro vyhledání je kolokační lemma (se značkou *col_lemma*), přičemž vyhledaná kolokační lemmata jsou seřazena dle frekvence. Výsledky je poté nutné ručně projít a určit, zda se jedná o frazémy, či nikoli. Do výsledků jsou nicméně zařazovány i zkrácené varianty frazémů, většinou slovníkově nezachycené (např. *být na pováženou* i *na pováženou*) – do kolokačních lemmat v tomto případě není zasahováno, o zařazení spojení do kolokačních lemmat by bylo třeba diskutovat s autory nástroje FRANTA. V případě lemmatu *štych* jsou výsledkem frazémy *nechat ve štychu* a *nechávat ve štychu*.

V dalším kroku je vytvořen kolokační profil lemmatu, který může napomoci nalézt některé z frazémů, které nástroj FRANTA neoznačoval. Kolokace jsou v korpusu vyhledávány následujícím způsobem: po zadání výše uvedeného dotazu je vybrána položka menu Kolokace. V kontextu je ponechán základní rozsah od -3 do $+3$, minimální frekvence 3, atribut lemma a třídění podle míry logDice. V případě lemmatu *štych* jsou výsledkem především lemma *nechat* a *v*, z čehož po nahlédnutí do dat vyplývá frazém *nechat ve štychu* (tento frazém je však označován i prostřednictvím nástroje FRANTA).

Jak při procházení kolokací, tak při procházení frekvenčních seznamů kolokačních lemmat je pro zařazení frazému do výsledků stanovena minimální frekvence 3 a výskyt alespoň ve dvou různých zdrojích. V případě menší frekvence či pouze jednoho zdroje by totiž výsledky nebylo možné zobecňovat, může se jednat o autorské frazémy.

Frazémy jsou vyhledávány v těchto dvou krocích (tj. přes kolokace a frazémovou anotaci) z několika důvodů. Jak již bylo zmíněno v kapitole FRANTA, program vyhledává pouze víceslovné jednotky z předem daného seznamu – pokud tedy frazém není v tomto seznamu uveden, nemůže být pomocí tohoto nástroje vyhledán. Dále mohou v některých případech nastat problémy s identifikací víceslovných jednotek, kvůli čemuž nemusí být frazém rovněž vyhledán. Z tohoto důvodu jsou tedy kromě nástroje FRANTA používány také kolokační profily lemmat. Právě tímto způsobem mohou být odhaleny dosud nezachycené frazémy, případně vedle invariantu také jeho varianty, transformace či modifikace.

Nakonec jsou výsledky srovnány s Čermákovým *Slovníkem české frazeologie a idiomatiky* (Čermák a kol. 2009; Čermák, Hronek a kol. 2009a; Čermák, Hronek a kol. 2009b; Čermák, Hronek a kol. 2009c), a to jak s tištěnou verzí, tak s online verzí dostupnou prostřednictvím portálu DEBDict (Horák a Rambousek 2013), a dále také s online *Akademickým slovníkem současné češtiny* (2017–2023), v němž jsou sice v době psaní práce dostupná pouze hesla od A–G, avšak je v něm velmi podrobně zpracována oblast frazeologie. Cílem tohoto srovnání je zjistit, zda jsou vyhledané frazémy slovníkově zachyceny, či nikoli.

Posledním krokem je rozčlenění jednotlivých lemmat dle typu frazémů. Jelikož výsledky analýzy přináší různé skupiny frazémů s podobnými vlastnostmi a klasifikace dle SČFI na slovesné, neslovesné a větné frazémy je pro popsání těchto vlastností nedostatečná, kdežto Čermákova podrobná klasifikace kolokačních a propozičních frazémů dle komponentů naopak příliš obsáhlá, byla pro účely práce zvolena vlastní klasifikace (jednotlivé typy jsou více popsány v kapitole Vlastní klasifikace využívaná v analytické části práce).

Pozornost je věnována především těmto typům frazémů:

- gramatické frazémy (např. *z hlediska*),
- monokolokabilní frazémy (např. *být k máni*),
- binomiály (např. *alfa a omega*),
- přirovnání (např. *žít si jako v bavlnce*),
- kontaktové frazémy (např. *pozdrav pánbůh*),
- cizojazyčná spojení (např. *alma mater*).

Pokud frazém není zařazen ani do jedné z těchto skupin, jsou rozlišeny frazémy:

- nominální (např. *od malička*),
- verbální (např. *hodit zpátečku*),
- propoziční (např. *andělíčku, můj strážníčku, opatruj mi mou dušičku*).

Každé analyzované lemma je pak přiřazeno k tomu typu, jenž je mezi frazémy daného lemmatu nejčastější. V případě, že jsou typy vyrovnané, je lemma zařazeno do dané skupiny podle frazému, který je ve větším korpusu, tj. v SYNv11, frekventovanější (např. lemma *veřejnost* se objevuje ve dvou frazémech: *široké / širší / nejširší veřejnosti* a *vyhýbat se veřejnosti* a v zařazení k nominálním, nikoli verbálním frazémům tedy přispěl fakt, že v SYNv11 má vyšší frekvenci frazém *široké / širší / nejširší veřejnosti*).

6.4 Popis jednotlivých pádů

V kapitolách 7.1 až 7.7 jsou shrnuty výsledky vyhledávání, a to dle jednotlivých pádů. Jednotlivé kapitoly vypadají následovně: Nejprve je uvedeno, kolik lemmat bylo v daném pádu analyzováno a kolik z nich (i procentuálně) vykazuje nějaké frazémy. Dále jsou postupně uvedena všechna lemmata, která se na frazémech podílejí. Lemmata jsou řazena podle toho, jak vysoké procento daného pádu vykazují, nejvýše jsou tedy zařazena lemmata s nejsilnější anomálií (např. i taková, která se vyskytují pouze ve sledovaném pádu). Pod daným lemmatem jsou vždy uvedeny frazémy nalezené v korpusu SYN2015. Pokud je u lemmatu uveden i druhý řádek, byly nalezeny další frazémy ve větších datech (v SYNv11). Tučně jsou přitom označeny ty frazémy, které nezachycuje žádný z používaných slovníků (tedy SČFI a ASSČ).

Pod výpisem všech lemmat, která se podílí na frazémech, se vždy nachází tabulka, jež se zabývá jednotlivými typy frazémů (více viz Postup). V tabulce je vedle uvedení konkrétních lemmat a typů rovněž shrnuto, kolik lemmat bylo k jednotlivým typům zařazeno.

7 Analýza vybraných anomálních lemmat

V této kapitole budou nejprve shrnuty výsledky analýzy vybraných anomálních lemmat (o výběru viz Použitý materiál) dle jednotlivých pádů a nakonec budou jednotlivé pády popsány a srovnány.

7.1 Nominativ singuláru

V nominativu singuláru bylo analyzováno celkem 162 lemmat, z toho 15 z nich je součástí jednoho či více frazémů (9 %).

amen

- s někým (už) je amen/ámen, s něčím (už) je amen/ámen

nehlava

- **hlava nehlava**, mlátit někoho hlava nehlava, bít někoho hlava nehlava,
- **jít do toho hlava nehlava**

alma

- alma mater

macho

- **jako macho**
- **chovat se jako macho**

spolehnutí

- na někoho je/není spolehnutí

pánbůh

- **pánbůh**, být jako **pánbůh**, hrát jako **pánbůh**, byl/stál tam (nahý), jak ho **pánbůh** stvořil, zaplat' (ti to) **pánbůh** / **pánbůh** (ti to) zaplat', pozdrav **pánbůh**, potěš **pánbůh**, **pánbůh** s námi (a zlý pryč), **pánbůh** chraň / chraň **pánbůh**, dej mu **pánbůh** nebe / věčnou slávu; **pánbůh** dal, **pánbůh** vzal; kdyby **pánbůh** dal, (sám) **pánbůh** ví (co), to ví jen **pánbůh**, **pánbůh** mě netrestej, pomáhej **pánbůh**, dejž to **pánbůh**, nedej **pánbůh**, člověk mívá (a/,) **pánbůh** mění
- **dá **pánbůh** zdraví**, **pánbůh** požehnej / požehnej **pánbůh**, pomoz **pánbůh**, **zaplat' **pánbůh** za ty dary**, když **pánbůh** dopustí, i motyka spustí, **aby nás **pánbůh** při zdravém rozumu zachovati ráčil**, **dej mu **pánbůh** lehké odpočinutí**, **zač někoho **pánbůh** trestá**, **pánbůh** ví proč/kdy, **pánbůh** někoho opatruj, dělat něco, až **pánbůh**

brání; **komu pánbůh, tomu všichni svatí; komu v rukou karty šustí, toho pánbůh neopustí**; jako že je pánbůh nade mnou; **člověk přemýšlí, bůh se směje**; koho pánbůh miluje, toho křížkem navštívuje; **dej pánbu dobrýtro**; komu dal bůh úřad, tomu dal i rozum; bůh vysoko, car daleko; umět jako pánbůh, dělat jako pánbůh

alfa

- alfa i/a omega (něčeho, někoho)

mužik

- zelený mužik

ave

- **Ave Maria**
- Ave Caesar

pardon

- tak pardon, to pardon
- **generální pardon**

káča

- točit se jako káča, káča pitomá

grand

- být grand, **zachovat/chovat se jako grand**

kurva

- **kurva fix**

balík

- balík peněz, venkovský/vesnický balík

studio

- sine ira et studio

Typ	Lemmata	Celkem lemmat
monokolokabilní frazémy	nehlava	1
binomiály	alfa	1
přirovnání	macho, káča, grand	3
kontaktové frazémy	pánbůh, pardon, kurva	3
cizojazyčná spojení	alma, ave, studio	3
nominální frazémy	mužík, balík	2
verbální frazémy	amen, spolehnutí	2

Tabulka 9: Typy vyhledaných frazémů v nominativu singuláru

7.2 Genitiv singuláru

V genitivu singuláru bylo analyzováno celkem 75 lemmat, z nichž je 39 součástí jednoho či více frazémů (tj. 52 %).

daleko

- z široka daleka

loci

- **genius loci**

naho

- svléknout do naha

přestání

- bez přestání

vitae

- **curriculum vitae**, **historia magistra vitae**

ztraceno

- **vyšumět do ztracena**, **vyznít do ztracena**

nynějšek

- od nynějška

třetice

- do třetice, do třetice všeho dobrého, **do třetice všeho zlého**

ustání

- bez ustání

nepaměť

- od nepaměti

činění

- mít co do činění s někým/něčím

pradávně

- od pradávna

sebezáchova

- **pud sebezáchovy**

maličko

- od malička

potucha

- nemít (ani nejmenší) potuchy, nebylo po někom (ani) potuchy

hledisko

- z hlediska
- z vyššího hlediska

méněcennost

- **mít komplex méněcennosti**

vínku

- dostat/dostávat něco do vínku, dát/dávat něco do vínku, **vložit něco do vínku**
- **nadělit něco do vínku, získat něco do vínku, vetknout něco do vínku, obdržet něco do vínku**

neznámo

- krok do neznáma, **skok do neznáma**

rozmýšlení

- bez rozmýšlení, **bez velkého rozmýšlení**, říct bez rozmýšlení, **odpovědět bez rozmýšlení**

ražení

- **být nějakého ražení**, být starého ražení
- **být tvrdého ražení**, být hrubého ražení

dřívějšek

- **z dřívějška**

rudo

- rozpálit/rozpálený do ruda

morek

- (až) do morku kostí, **pronikat/proniknout do morku kostí**, nenávidět do morku kostí
- být promrzlý/promrznout do morku kostí

povšimnutí

- **bez povšimnutí**, nechat/nechávat bez povšimnutí, zůstat/zůstatvat bez povšimnutí, **projít bez povšimnutí**, **přejít bez povšimnutí**, **minout/míjet bez povšimnutí**, **ponechat bez povšimnutí**

ptaní

- **bez ptaní**

slonovina

- žít ve věži ze slonoviny

klubíčko

- stočit se do klubíčka, schoulit/choulit se do klubíčka, **stulit se do klubíčka**, svinout se do klubíčka

úhona

- bez úhony, **přežít bez úhony**, **přečkat bez úhony**, vyváznout bez úhony, **přestát bez úhony**, vyjít bez úhony, dojít úhony, **doznat úhony**
- bez úhony na zdraví, utrpět úhony

němota

- **do němoty**, zpít/zpíjet se do němoty, opít/opíjet se do němoty, **ztřískat se do němoty**
- ožrat se do němoty, **setnout se do němoty**, **propít se do němoty**, **zhulákat se do němoty**

slech

- **(není) (ani) vidu ani slechu, ani slechu, nebýt po někom ani vidu ani slechu**

prodlení

- bez prodlení, nebezpečí z prodlení

málo

- jeden z mála

záhrobí

- **hlas ze záhrobí**

váhání

- bez váhání, udělat něco bez váhání

lad

- bez ladu a skladu, být bez ladu a skladu
- ležet bez ladu a skladu

odnětí

- odnětí svobody, trest odnětí svobody

ráže

- dostat se do ráže

století

- na sklonku století, **na přelomu století**

Typ	Lemmata	Celkem lemmat
gramatické frazémy	hledisko	1
monokolokabilní frazémy	naho, ztraceno, třetice, činění, potucha, vínek, rudo	7
binomiály	daleko, slech, lad	3
cizojazyčná spojení	loci, vitae	2
nominální frazémy	přestání, nyníšek, ustání, nepaměť, pradávno, sebezáchova, maličko, neznámo, rozmýšlení, dřívěšek, ptaní, prodlení, málo, záhrobí, váhání, odnětí, století	17
verbální frazémy	méněcennost, ražení, morek, povšimnutí, slonovina, klubíčko, úhona, němota, ráže	9

Tabulka 10: Typy vyhledaných frazémů v genitivu singuláru

7.3 Dativ singuláru

V dativu singuláru bylo analyzováno nejvíce lemmat – celkem 252. Jen 47 z nich je ale součástí jednoho či více frazémů (19 %).

mání

- být k mání

dostání

- být k dostání

nepoznání

- změnit (se) k nepoznání, být k nepoznání, **proměnit se k nepoznání, měnit se k nepoznání**
- být změněný k nepoznání, **proměňovat se k nepoznání**

nerozeznání

- být k nerozeznání od něčeho/někoho

zahození

- nebýt k zahození

snědek

- něco k snědku, **být k snědku, něco dobrého k snědku**

sehnání

- **být k sehnání**

nezaplacení

- být k nezaplacení, být penězi k nezaplacení

dispozice

- **k dispozici, mít k dispozici, dát k dispozici, dostat k dispozici, dávat k dispozici, zůstat k dispozici**

zbláznění

- být k zbláznění

prasknutí

- **nacpaný/nacpat k prasknutí, až k prasknutí, narvaný/narvat k prasknutí, (mít) nervy napjaté/napnuté k prasknutí, nabitý k prasknutí, napnutý / napnout / napínat (až) k prasknutí, naplněný/naplňit k prasknutí, plný k prasknutí, napěchovaný k prasknutí, (být) napjatý k prasknutí, naditý k prasknutí, přežraný k prasknutí**
- **přeplněný k prasknutí, být přecpaný k prasknutí, naplnit až k prasknutí, nacpat se (až) k prasknutí, našponovaný k prasknutí, cpát se k prasknutí, natažený k prasknutí, nažrat se k prasknutí, naplnit se k prasknutí**

uctění

- **k uctění památky**
- **k uctění výročí, k uctění minutou ticha**

vidění

- být k vidění
- mít oči k vidění a uši k slyšení, **k Jiříkovu vidění**

nevíra

- být k nevěře

nepříčetnost

- **dovádět k nepříčetnosti, rozzuřit k nepříčetnosti, dohánět k nepříčetnosti, přivádět k nepříčetnosti**
- **vytáčet k nepříčetnosti, vydráždit k nepříčetnosti, rozlítit k nepříčetnosti, rozběsnit k nepříčetnosti, vytočit k nepříčetnosti**

přesila

- bojovat proti přesile, ubránit se přesile, podlehnout přesile
- bránit se proti přesile, proti drtivé přesile

zamyšlení

- **přimět k zamyšlení, nutit k zamyšlení, vést k zamyšlení, přivést/přivádět k zamyšlení**
- **k zamyšlení se sám nad sebou**

odvrácení

- k odvrácení pozornosti
- k odvrácení pohromy, **k odvrácení pádu**

dokonalost

- dovést/dovedený k dokonalosti, přivést k dokonalosti, **dotáhnout k dokonalosti**

dovršení

- **k dovršení všeho**

čert

- k čertu, jít k čertu, k čertu s tím/ním, poslat k čertu, táhnout k čertu, posílat k čertu, upsat se čertu, **běžet k čertu**
- **čiň čertu dobře (a) peklem se ti odmění, upsat duši čertu, kde k čertu vězí; čiň čertu dobře, peklem se ti odvděčí; podej čertu prst a on ti vezme ruku**

dosažení

- k dosažení (svého) cíle, k dosažení úspěchu, k dosažení výsledku, k dosažení většiny, k dosažení vítězství, k dosažení míru
- k dosažení vrcholu, k dosažení dokonalosti

záhuba

- **být odsouzený / odsoudit k záhubě**
- vyhnout se záhubě

koukání

- **být ke koukání**

nalezení

- nebýt k nalezení, k nalezení řešení, **k nalezení kompromisu**, k nalezení cesty, **k nalezení**, k nalezení smyslu
- k nalezení východiska, k nalezení klidu, k nalezení společné řeči, **k nalezení smyslu života**, k nalezení spravedlnosti, **k nalezení času**, k nalezení odvahy

pokoušení

- odolat/odolávat pokoušení, podlehnout pokoušení

bližní

- **láska k bližnímu**
- **sobě ku cti, bližnímu ku pomoci**

normál

- **vymykát se normálu, vracet / vrátit / návrat k normálu**

uskutečnění

- přispět k uskutečnění, k uskutečnění plánu, k uskutečnění snu
- zabránit uskutečnění, k uskutečnění změny, k uskutečnění nápadu

odvedení

- k odvedení pozornosti
- odvedení na vojnu

šilenství

- **přivádět/přivést k šilenství**, dohánět/dohnat k šilenství, **dovádět/dovést k šilenství**, propadnout šilenství, davovému šilenství
- válečnému šilenství

politování

- být k politování

vábění

- **odolat vábení, podlehnout vábení**
- **odolávat vábení**

pítí

- **dát si k pití**, nebýt k pití, oddávat se pití
- **dát si něco dobrého k pití**, dohnat k pití, svést k pití, **k pití na ex, k pití s rozumem, k pití s mírou**

zastavení

- **nebýt k zastavení**

dobro

- mít k dobru, **k dobru**, dávat/dát k dobru, přičíst/přičítat k dobru, připsat k dobru, získat k dobru, **připočíst k dobru**, k dobru věci

šípek

- poslat někoho k šípku, jít k šípku

spokojenost

- k úplné/plné spokojenosti, **ke všeobecné spokojenosti, k obecné spokojenosti**
- dopadnout ke spokojenosti

udržení

- **nebýt k udržení**

zánik

- odsoudit k zániku, vést k zániku, předurčit k zániku

hrobník

- utéct hrobníkovi z lopaty, **sebrat někoho hrobníkovi z lopaty**
- **sedět hrobníkovi na lopatě, dívat se hrobníkovi na lopatu**

odhalení

- k odhalení tajemství
- k odhalení (pravé) tváře

pata

- od hlavy (až) k patě, prohlédnout/prohlížet si od hlavy (až) k patě, změřit si od hlavy (až) k patě
- k Achillově patě

pochopení

- **k hlubšímu pochopení**, k pochopení smyslu

úhona

- přijít k úhoně, **dojít k úhoně**
- **přicházet k úhoně, docházet k úhoně**

veřejnost

- široké / širší / nejširší veřejnosti
- vyhýbat se veřejnosti

závěr

- dojít k závěru, dospět k závěru, vést k závěru

Typ	Lemmata	Celkem lemmat
monokolokabilní frazémy	mání, snědek, nevíra	3
nominální frazémy	bližní, spokojenost, pata, veřejnost	4
verbální frazémy	dostání, nepoznání, nerozeznání, zahození, sehnání, nezaplacení, dispozice, zbláznění, prasknutí, uctění, vidění, nepřičetnost, přesila, zamyšlení, odvrácení, dokonalost, dovršení, čert, dosažení, záhuba, koukání, nalezení, pokušení, normál, uskutečnění, odvedení, šílenství, politování, vábení, pití, zastavení, dobro, šípek, udržení, zánik, hrobník, odhalení, pochopení, úhona, závěr	40

Tabulka 11: Typy vyhledaných frazémů v dativu singuláru

7.4 Akuzativ singuláru

V akuzativu singuláru bylo analyzováno pouze 66 lemmat, avšak 58 z nich je součástí jednoho či více frazémů (88 %).

bach

- dávat/dát (si) bacha
- dávat/dát si velkého bacha, dávat si bacha na pusu

odiv

- vystavovat/vystavit něco na odiv, dávat/dát něco na odiv, stavět něco na odiv

počest

- **přípitek na počest**
- připít na počest, pít na počest

potaz

- brát v potaz, vzít v potaz

povážená

- být na pováženou, **na pováženou**

roveň

- postavit/stavět něco na roveň s něčím, klást něco na roveň něčemu / s něčím
- položit na roveň něčemu

rozloučená

- na rozloučenou, **dát na rozloučenou, dostat na rozloučenou**

shledaná

- na shledanou

úkor

- dělat něco na úkor něčeho, být na úkor něčeho/někoho, **na úkor**

uvítaná

- na uvítanou

rozmyšlená

- dát někomu čas na rozmyšlenou, vzít si čas na rozmyšlenou, mít čas na rozmyšlenou, nechat/ponechat si čas na rozmyšlenou, dostat čas na rozmyšlenou
- vzít si den na rozmyšlenou, **mít dost času na rozmyšlenou, brát si čas na rozmyšlenou**, nechat si den na rozmyšlenou

poněti

- nemít o něčem (ani nejmenší) poněti, **nemít poněti o čase**

vysvětlená

- **na vysvětlenou**, dodat/dodávat na vysvětlenou, říct/říkat na vysvětlenou, uvést na vysvětlenou
- **uvádět na vysvětlenou**

srozuměná

- dát/dávat někomu na srozuměnou, **přikývnout na srozuměnou**
- dostat od někoho na srozuměnou

omak

- na omak

přítrž

- učinit přítrž, udělat přítrž
- o přítrž

spadnutí

- být na spadnutí

oplátka

- na oplátku, dát někomu něco na oplátku, půjčka na/za oplátku
- vzít si od někoho něco na oplátku, na věčnou oplátku, půjčit si něco na věčnou oplátku, půjčka na věčnou oplátku

neprospěch

- rozhodnout v něčí neprospěch

zálusk

- mít zálusk na něco, dělat si zálusk na něco, dostat zálusk na něco
- **udělat si zálusk, dostávat zálusk, ztratit zálusk**

předvečer

- **v předvečer**

fous

- o fous, **utéci o fous**, minout něco/někoho o fous, **o fous lepší**, uniknout (jen) o fous
- na fous (přesně), být na fous

požadání

- na požádání
- ukázat někomu něco na požádání, dát někomu něco na požádání, udělat něco na požádání, požádání o ruku

rámě

- nabídnout někomu rámě

půl

- **půl na půl**, poslouchat něco/někoho na půl ucha, půl světa, na půl úst, **na půl plynu**, na půl žerdi, **procestovat půl světa**, na půl huby, vlajka na půl žerdi, vyjít někomu na půl cesty vstříc, spát jen na půl oka, odpovídat na půl úst, zastavit se na půl cesty
- být na půl žerdi, **spustit vlajky na půl žerdi**, **mluvit na půl úst**, **svěsit vlajky na půl žerdi**, mít kalhoty na půl žerdi, mluvit na půl huby, říct na půl huby, zůstat stát na půl cesty, odpovídat na půl huby, pozdravit na půl huby

špacír

- pustit/pouštět si pusu na špacír, pouštět/pustit si hubu na špacír, **pustit/pouštět na špacír**, jít na špacír

vrub

- jít na vrub, připsat/připisovat něco na vrub, přičíst/přičítat na vrub, **padat na vrub**, **spadat na vrub**
- připočítat na vrub

potkání

- na setkání, vykládat něco každému na setkání

schovávaná

- hrát/hrávat (si) na schovávanou, hra na schovávanou, **na schovávanou**
- **zahrát si na schovávanou**

vteřinka

- **na vteřinku**

majlant

- stát majlant, **utratit majlant**
- **utrácet majlant**

tušení

- mít tušení, nemít o něčem (ani nejmenší) tušení, **mít neblahé tušení**, mít nejasné tušení, neblahé tušení
- **mít divné tušení**, dostat tušení, divné tušení, pocítit tušení

vybraná

- mít na vybranou, dát/dávat někomu na vybranou, **na vybranou**, dostat na vybranou

píd'

- neustoupit ani o píd'

prospěch

- mít prospěch, **využít/využívat ve svůj prospěch**, **otočit ve svůj prospěch**, rozhodnout v něčí prospěch, získat prospěch, mluvit v něčí prospěch, **obrátit ve svůj prospěch**, ztratit prospěch, **zvrátit ve svůj prospěch**
- **zlomit ve svůj prospěch**, **myslet na svůj/vlastní prospěch**, hledět na svůj prospěch, koukat na svůj prospěch

punc

- dodat/dodávat punc, nést punc někoho/něčeho, **ztratit punc**, **získat punc**, dávat/dát něčemu punc něčeho/někoho

sobota

- na Bílou sobotu
- anglickou sobotu

úterý

- na masopustní úterý
- na Šedivé úterý

čtvrtek

- na Zelený čtvrtek
- na Tučný čtvrtek

pondělí

- na Velikonoční pondělí, na Modré pondělí
- na Svatodušní pondělí, na masopustní pondělí

prim

- hrát někde prim, **držet někde prim**, mít někde prim

šmrnc

- mít šmrnc, dodat šmrnc, dát šmrnc
- ztratit šmrnc

pátek

- na Velký pátek, (už) nějaký (ten) pátek, **v pátek třináctého**, na černý pátek, ještě nějaký ten pátek
- již nějaký ten pátek, **o/na nějaký ten pátek**, mít nějaký ten pátek za sebou, **Velký pátek deštivý dělá rok žízňivý**, pamatovat nějaký (ten) pátek

pozor

- dávat/dát (si) pozor, **dávat/dát (si) velký pozor**, dávat/dát (si) na sebe pozor, dávat/dát (si) dobrý pozor, dávat/dát (si) pozor na jazyk, dávat (si) na někoho pozor, **dávat pozor na cestu, dávat si pozor na pusu**, dávat si pozor na hubu, **dávat na sebe dobrý pozor**
- dávat si pozor na ústa, **dát si pozor na pusu, dávat si velký pozor na pusu, dát si pozor na hubu, dát si pozor na cestu**

smilování

- **pro smilování boží**, čekat na něco/někoho jako na smilování, prosit o smilování
- čekat na něco/někoho jako na smilování boží / boží smilování, neznat smilování, **žadonit o smilování**

úplata

- **za úplatu**

neurčito

- odložit/odkládat něco na neurčito, **odročit na neurčito, odsunout na neurčito**
- **odsouvat na neurčito, odročovat na neurčito**

soustrast

- upřímnou soustrast, vyjádřit/vyjadřovat někomu (svou) soustrast, projevit někomu soustrast, vyslovit někomu soustrast, přijměte mou upřímnou soustrast
- přijměte (mou) upřímnou soustrast, **projevit svou soustrast**, přijměte hlubokou soustrast, přijměte mou nejhlubší soustrast

středa

- na Škaredou středu, na Popeleční středu

peníz

- **za levný peníz**, slušný peníz
- pěkný peníz, nekřesťanský peníz, jidášský peníz, **obracet každý peníz dvakrát**

náskok

- mít náskok, získat náskok, zvýšit náskok, ztratit náskok, dohnat náskok, snížit náskok, dát někomu náskok, **udržovat si náskok**

chvilka

- **za chvilku, každou chvilku**, (ani) chvilku se nezastavit, najít si chvilku (času), chvilku neposedět, **vybrat si slabou chvilku, mít chvilku klidu, chvilku klidu, horkou chvilku**, (mít) slabou chvilku, **(mějte) chvilku strpení**, každý chvilku tahá pilku, světlou chvilku, těžkou chvilku, **mít chvilku pokoj**
- dopřát si chvilku klidu, **požádat/žádat o chvilku strpení, udělat si chvilku času, mít právo na slabou chvilku, dát (na) chvilku pokoj**, poprosit/prosit někoho o chvilku strpení, **mít jasnou chvilku, mít na chvilku pokoj**, nedopřát někomu ani chvilku klidu

hold

- vzdát/vzdávat někomu hold, složit/skládat někomu hold

zpátečka

- **zařadit zpátečku**, hodit zpátečku, **vzít zpátečku, brát zpátečku**, dát zpátečku
- hodit tam zpátečku, dát tam zpátečku

doraz

- jet na doraz, jít na doraz, **dřít na doraz, být na doraz**
- **hrát na doraz, bojovat na doraz, makat na doraz**, dělat na doraz

minutka

- **na minutku, ani na minutku**
- **minutku po minutce**

doživotí

- na doživotí, **dostat (na) doživotí**
- **vyfáknout doživotí, dostat (se) na doživotí za mříže**

nemilost

- upadnout v nemilost, vydat/vydaný (se) někomu/něčemu na milost a nemilost, vzdát se někomu na milost a nemilost
- **upadat v nemilost, vydávat se někomu/něčemu na milost a nemilost, vzdávat se na milost a nemilost**

Typ	Lemmata	Celkem lemmat
monokolokabilní frazémy	bach, odív, potaz, roveň, omak, zálsk, předvečer, špacír, majlant, punc, prim, šmrnc, hold	13
binomiály	nemilost	1
kontaktové frazémy	rozloučená, shledaná, uvítaná, soustrast	4
nominální frazémy	oplátka, setkání, schovávaná, vteřinka, sobota, úterý, čtvrtek, pondělí, pátek, úplata, středa, peníz, minutka	13
verbální frazémy	počest, povážená, úkor, rozmyšlená, ponětí, vysvětlená, srozuměná, přítrž, spadnutí, nepospěch, fous, požádání, rámě, půl, vrub, tušení, vybraná, píd', prospěch, pozor, smilování, neurčito, náskok, chvilka, zpátečka, doraz, doživotí	27

Tabulka 12: Typy vyhledaných frazémů v akuzativu singuláru

7.5 Vokativ singuláru

Ve vokativu singuláru bylo analyzováno celkem 201 lemmat, z nichž pouze 10 je součástí jednoho či více frazémů (5 %).

vůl

- **ty vole**
- starej vole

pán

- pane bože, pane na nebi, **kriste pane**

bůh

- ach bože, (ach) bože můj / můj (ty) bože, pane bože, dobrý bože, **nedej bože**, bože chraň / chraň bože, bože na nebi, bože, smiluj se nad námi
- bože, dej at'/abych, bože, rozbože, bože, chraň mě před mými přáteli, před nepřáteli se ochráním sám, bože, kýž jsem policajtem

hrom

- **a hrome**

prostota

- svatá prostoto

káča

- káčo pitomá, ty káčo

panenka

- panenka skákavá, panenka Maria; sedávej panenka v koutě, budeš-li hodná, najdou tě
- panenka svatá

baba

- (a teď) babo rad'

andělíček

- andělíčku, můj strážníčku, (opatruj mi mou dušičku)

kráva

- **ty krávo**

Typ	Lemmata	Celkem lemmat
kontaktové frazémy	vůl, pán, bůh, hrom, prostota, káča, panenka, kráva	8
verbální frazémy	baba	1
propoziční frazémy	andělíček	1

Tabulka 13: Typy vyhledaných frazémů ve vokativu singuláru

7.6 Lokál singuláru

V lokálu singuláru bylo analyzováno celkem 169 lemmat, přičemž 71 z nich je součástí jednoho či více frazémů (tj. 42 %).

střeh

- být ve střehu

skryt

- v/ve skrytu duše, doufat v skrytu duše
- věřit v skrytu duše, cítit v skrytu duše, v skrytu duše si přát

tichost

- **v tichosti**, ve vši tichosti, **v naprosté tichosti**, udělat něco v tichosti

domnění

- v domnění, že; žít v domnění, že; **v mylném domnění**
- **v bláhovém/blahém domnění**

rámec

- **v rámci**, v rámci (daných) možností
- **v rámci možnosti**

sklonek

- na sklonku roku, na sklonku života, na sklonku dne, na sklonku léta, **na sklonku války**, na sklonku století, na sklonku týdne
- na sklonku měsíce

nedohledno

- být v nedohlednu

soulad

- být v souladu, **v souladu**, **v souladu s**

mžik

- v mžiku

pokraj

- **balancovat na pokraji, potácet se na pokraji, být na pokraji propasti**, být na pokraji něčeho, stát na pokraji něčeho
- **stát na pokraji propasti**

rozmezí

- **pohybovat se v rozmezí**, být v rozmezí, ležet v rozmezí

přepočet

- v přepočtu

duch

- v duchu, říkat/říct si něco v duchu, **nést se v duchu**, pomyslet/myslet si něco v duchu, v duchu se smát, **v tomto duchu**, klesat / poklesnout / klesnout na duchu, na těle i na duchu / na duchu i na těle, v duchu doby, o duchu svatém, být v duchu s někým, v dobrém/nejlepším duchu, **nést se v přátelském duchu**
- **v bojovném duchu, tvůrčím duchu**, v obchodním duchu, o zlém duchu, chudý na duchu, prospívat na těle i na duchu, o duchu času, prospívat na duchu, **v mdlém duchu**

štych

- nechat/nechávat někoho ve štychu

výslužba

- **být ve výslužbě**

náručí

- **držet v náručí**, sevřít/svírat v náručí, být v náručí, mít v náručí, tisknout v náručí, **skončit v náručí**
- **končit v náručí**

pomezí

- **na pomezí**

vzduchoprázdno

- **žít ve vzduchoprázdnu**

mezičas

- v mezičase

mezidobí

- v mezidobí

současnost

- v současnosti

krátkost

- v krátkosti

pořádek

- být v pořádku, **v pořádku**, to je v pořádku, **nebýt (úplně) v pořádku**, všechno je v pořádku, **mít v pořádku**, (být) na denním pořádku, **být úplně v pořádku**, **být v naprostém pořádku**, být v nejlepším pořádku, mít to v hlavě v pořádku, **vrátit se v pořádku**, **proběhnout v pořádku**, **v naprostém pořádku**, udržovat něco v pořádku, **dojet v pořádku**, pěkně po pořádku, **v zasedacím pořádku**, **probíhat v pořádku**, vzít (to) po pořádku, o veřejném pořádku
- mít něco na denním pořádku, **v ústavním pořádku**, **vracet se v pořádku**, **(být) na běžném pořádku**, **hezky po pořádku**, **vzít to pěkně po pořádku**, držet v pořádku, na tanečním pořádku, brát (to) (pěkně) po pořádku, v pěkném pořádku, tak pěkně po pořádku, dát po pořádku

porovnání

- v porovnání s

obležení

- být v obležení

přelom

- **na přelomu**, **na přelomu roku**, **na přelomu století**

výsluní

- **na výsluní**, být na výsluní, **hřát/vyhřívát se na výsluní**
- na výsluní slávy, na výsluní přízně, **hřát/vyhřívát se na výsluní slávy**

stručnost

- ve stručnosti

špice

- být na špici něčeho, stát na špici něčeho

krajíček

- mít slzy na krajíčku, mít na krajíčku, **se slzami na krajíčku**, mít pláč na krajíčku, **s pláčem na krajíčku**

podstata

- v podstatě, ve své podstatě, **v samé podstatě**, o podstatě věci/věci, o pravé podstatě
- **v pravé podstatě věci**

půle

- v půli cesty, zůstat na/v půli cesty, zastavit se na/v půli cesty
- uváznout v půli cesty

bavlnka

- mít (se) jako v bavlnce, žít si jako v bavlnce, **cítit se jako v bavlnce**, chovat někoho jako v bavlnce
- **žít v bavlnce, hýčkat jako v bavlnce**

průběh

- **v průběhu**

srovnání

- ve srovnání s
- srovnání se zemí, srovnání s krokem, srovnání kroku

civil

- být v civilu, chodit v civilu

hloub

- v hloubi duše, v hloubi duše cítit, v hloubi srdce cítit

podpaží

- vzít někoho v podpaží

rozpuk

- být v rozpuku, v rozpuku mládí

pozadí

- být v pozadí, stát v pozadí
- **sedět na pozadí**

souvislost

- v souvislosti s

báze

- na bázi

případ

- **v tomto/tomhle případě**, v každém případě, v žádném případě, **v takovém případě**, **v tom případě**, **v něčem případě**, v případě potřeby, **v lepším případě**, v nejlepším případě, v daném případě, v krajním případě, v nejhorším případě, **v případě nouze**, v kladném případě, **v nutném případě**, **v případě nejvyšší nouze**, **v některém případě**
- v záporném případě, o zajímavém případě, o odstrašujícím případě, **o beznadějném případě**

rozcestí

- být na rozcestí, stát na rozcestí, **ocitnout se na rozcestí**
- **octnout se na rozcestí**

mizina

- být na mizině
- **octnout se na mizině**

propadliště

- zmizet v propadlišti (dějin), skončit v propadlišti (dějin), **v propadlišti dějin**

tajnost

- držet / udržet / udržovat něco v tajnosti, **v tajnosti**, uchovat/uchovávat něco v tajnosti, **zůstat v tajnosti**, ve vší tajnosti, **(uchovat) v naprosté tajnosti**, zachovat něco v tajnosti, **držet v přísné tajnosti**
- **držet v naprosté tajnosti**, podržet něco v tajnosti, chovat něco v tajnosti, **udržet v naprosté tajnosti**, zachovat **v naprosté tajnosti**, **udržet v přísné tajnosti**, uchovat **v přísné tajnosti**, zachovat/zachovávat **v přísné tajnosti**

zajetí

- být v zajetí, držet v zajetí, **žít v zajetí**, **ocitnout se v zajetí**, **chovat v zajetí**
- mít v zajetí

návaznost

- **v návaznosti na**

pořadí

- být v pořadí

základ

- na základě

zavěťří

- být v zavěťří

paluba

- vítěj na palubě

pohoda

- být v pohodě, **v pohodě**, to je v pohodě, **v klidu a pohodě**
- **dožít v klidu a pohodě**

rozpor

- být v rozporu, **být v rozporu s něčím**, **v příkrém rozporu s něčím**
- být v rozporu se skutečností, **v příkrém rozporu**, být v rozporu s pravdou

výše

- **být ve výši**, být na výši
- být na výši doby, stát na výši doby

chvost

- být na chvostě, **držet se na chvostě**
- **vlát na chvostě**, **vézt se na chvostě**

kontext

- **v tomto kontextu**, v dobovém kontextu

ústraní

- v **ústraní**, žít v ústraní, **někde v ústraní**, **život v ústraní**, **držet se v ústraní**, být v ústraní

podoba

- v **podobě**, v lidské podobě, v **čisté podobě**, v pravé podobě
- ukázat se ve své pravé podobě

rauš

- **být v rauši**

léto

- v zimě v létě, v babím létě
- až se zima zeptá, cos dělal v létě; v milostivém létě

područí

- být v područí někoho/něčeho

průměr

- v **průměru**

úmysl

- mít něco v úmyslu, v dobrém úmyslu, ve zlém úmyslu
- přicházet v dobrém úmyslu

funus

- s křížkem po funuse, přijít s křížkem po funuse
- **přicházet s křížkem po funuse**, **přijet s křížkem po funuse**, být jako na/o funuse, tvářit se jako na funuse

popředí

- být v popředí, stát v popředí, **být v popředí zájmu**, **stát v popředí zájmu**, **držet se v popředí**

tíseň

- být v tísní, (být) v časové tísní, (octnout se / být) ve finanční tísní

defenzíva

- být v defenzívě

rovinka

- v cílové rovině

samota

- **o samotě**, nechat/nechávat někoho o samotě, být o samotě, **mluvit o samotě**, zůstat o samotě, octnout se o samotě, žít o samotě

Typ	Lemmata	Celkem lemmat
gramatické frazémy	rámec, soulad, pomezí, porovnání, přelom, průběh, srovnání, souvislost, báze, návaznost, základ	11
monokolokabilní frazémy	střeh, skryt, nedohledno, mžik, štych, špice, hloub, propadliště, chvost, rauš	10
binomiály	pohoda	1
přirovnání	bavlnka	1
nominální frazémy	tichost, domnění, sklonek, přepočet, duch, mezičas, mezidobí, současnost, krátkost, výsluní, stručnost, podstata, půle, případ, kontext, podoba, léto, průměr, rovinka	19
verbální frazémy	pokraj, rozmezí, výslužba, náručí, vzduchoprázdno, pořádek, obležení, krajíček, civil, podpaží, rozpuk, pozadí, rozcestí, mizina, tajnost, zjetí, pořadí, závětrí, rozpor, výše, ústraní, područí, úmysl, funus, popředí, tíseň, defenzíva, samota	28
propoziční frazémy	paluba	1

Tabulka 14: Typy vyhledaných frazémů v lokálu singuláru

7.7 Instrumentál singuláru

V instrumentálu singuláru bylo analyzováno celkem 177 lemmat, z nichž 66 je součástí jednoho či více frazémů (tedy 37 %).

povděk

- kvitovat něco s povděkem, přijmout něco s povděkem
- **přijímat s povděkem**

přihlédnutí

- s přihlédnutím k

nedávno

- před nedávnem

mávnutí

- **jako/jakoby mávnutím kouzelného proutku, odbýt mávnutím ruky**, jako mávnutím proutku, přejít něco mávnutím ruky, zmizet jako mávnutím (kouzelného) proutku
- **odbývat mávnutím ruky**, přecházet něco mávnutím ruky, jako mávnutím čarovného proutku, **jako mávnutím kouzelným proutkem**, zmizet jako mávnutím čarovného proutku

úkos

- **podívat se úkosem, pohlédnout úkosem, dívat se úkosem, hledět úkosem, pohlížet úkosem**
- **sledovat úkosem**

coul

- každým coulem, být každým coulem

drobnohled

- být pod drobnohledem, **pod drobnohledem**, mít něco/někoho pod drobnohledem, ocitnout se pod drobnohledem
- sledovat něco/někoho pod drobnohledem, zkoumat něco/někoho pod drobnohledem, pozorovat něco/někoho pod drobnohledem

taktovka

- pod taktovkou někoho/něčeho

verva

- s vervou, s takovou vervou, s plnou vervou, s velkou vervou, se vší vervou, dělat něco s vervou, s novou vervou
- s vervou sobě/jemu vlastní

lub

- mít něco za lubem

podiv

- být s podivem, s podivem
- přijmout něco s podivem

součást

- být nedílnou součástí, stát se nedílnou součástí

klepnutí

- klepnutím přes prsty

oklika

- vzít to oklikou, jet oklikou, jít na to oklikou
- dovědět se oklikou

poklepání

- poklepáním na rameno

brva

- nehnout brvou, hnout brvou, pohnout brvou
- hýbat brvou

grácie

- s grácií

pompa

- s velkou pompou

despekt

- pohlížet s despektem (na něco/někoho), vyjadřovat se s despektem (o něčem/někom), hledět s despektem (na něco/někoho)
- shlížet/nahlížet s despektem (na něco/někoho), dívat se s despektem (na něco/někoho)

záštita

- **pod záštitou někoho/něčeho**
- záštitou míru, bezpečnou záštitou

ztělesnění

- být ztělesněním klidu, **být ztělesněním něčeho**

pokus

- poklusem klus
- poklusem v klus

výjimka

- být výjimkou, čestnou výjimkou, stát se výjimkou, **výjimkou potvrzující pravidlo, světlou výjimkou**

zosobnění

- **stát se zosobněním něčeho, být zosobněním něčeho**

rozmysl

- **s rozmyslem**, jednat s rozmyslem, dělat něco s rozmyslem

směr

- **směrem do, směrem k/ke, směrem na, směrem od, směrem proti**, jít směrem, ubírat se směrem, **(první) krok správným směrem**

terč

- **být terčem kritiky**, být terčem posměchu, stát se terčem posměchu, **stát se terčem kritiky**, stát se terčem vtipů, **stát se terčem výsměchu**
- být terčem vtipů, **živým terčem**, stát se terčem pomluv/pomluvy, být terčem pomluv

pohrůžka

- **pod pohrůžkou**

synonymum

- **stát se synonymem pro něco, být synonymem pro něco**

vypětí

- s vypětím všech sil, s vypětím sil
- s krajním vypětím sil

nedočkavost

- hořet nedočkavostí
- umírat nedočkavostí

plášťík

- pod plášťíkem, **pod plášťíkem noci**
- **pod plášťíkem tmy**

trysk

- **rozběhnout se tryskem, říít se tryskem, vyrazit tryskem**
- **pádit tryskem, uhánět tryskem**

chlouba

- **být chloubou někoho/něčeho, stát se chloubou někoho/něčeho**

způsob

- svým způsobem, **nějakým způsobem, takovým způsobem, určitým způsobem**, tím způsobem, žádným způsobem, nejlepším možným způsobem

poklička

- držet/udržet něco pod pokličkou, zůstat pod pokličkou, být pod pokličkou
- bublat pod pokličkou

záminka

- být záminkou, stát se záminkou, pod průhlednou záminkou, pod záminkou
- vítanou záminkou

novota

- **zářit novotou**

přítěž

- **stávat/stát se přítěží, být přítěží**

rozednění

- **před rozedněním, s rozedněním**

tahák

- **stát se tahákem, být tahákem, být hlavním tahákem**

třešnička

- být třešničkou na dortu

leitmotiv

- být leitmotivem, stát se leitmotivem

nadhled

- s nadhledem, brát s nadhledem

poukaz

- s poukazem na

dovolení

- s dovozením

pravděpodobnost

- s velkou pravděpodobností, se vši pravděpodobností, počítat s pravděpodobností

fofr

- vzít to fofrem

lákadlo

- být lákadlem

lehkost

- s lehkostí

úsměv

- s úsměvem, s širokým úsměvem, s úsměvem na rtech, s úsměvem na tváři, se zářivým úsměvem, s pobaveným úsměvem, s potutelným úsměvem, se šibalským úsměvem, s úsměvem od ucha k uchu, blýsknout úsměvem, s trpkým úsměvem, s blaženým úsměvem, se sladkým úsměvem, zářit úsměvem, přejít úsměvem, s hořkým úsměvem, s pohrdavým úsměvem, s kyselým úsměvem, s úsměvem jde všechno líp, se strojeným úsměvem
- říkat/říct s úsměvem na rtech, s cukerinovým úsměvem

volant

- sedět za volantem, usnout za volantem

předzvěst

- stát se předzvěstí něčeho, být předzvěstí něčeho

skřípění

- se skřípěním zubů
- s pláčem a skřípěním zubů

vynaložení

- **s vynaložením úsilí**, s vynaložením energie, s vynaložením námahy

gusto

- s gustem, **s nějakým gustem**
- udělat něco s gustem

lupa

- prohlížet něco pod lupou, **hledat lupou**
- zkoumat něco pod lupou

prásknutí

- prásknutím dveřmi
- prásknutím bičem

příčinění

- vlastním příčiněním

bič

- práskat bičem, jako když bičem mrská, **jako šlehnutí bičem, moci se převlékat za bičem, jako švihnutí bičem**
- s bičem božím, **jako když bičem práskne**, moci se svlíkat za bičem

cíl

- jít (tvrdě) za svým cílem, minout/míjet se cílem, vydat se za cílem, **vytčeným cílem**
- **tajit se cílem**

fiasko

- **skončit/končit fiaskem**

nadsázka

- **s nadsázkou, říci/říkat s nadsázkou**

prominutí

- **s prominutím**
- prominutím trestu

rouška

- pod rouškou tmy, pod rouškou noci, pod rouškou tajemství
- zůstávat/zůstat zahalen rouškou tajemství, být zahalen rouškou tajemství

zenit

- být za zenitem

Typ	Lemmata	Celkem lemmat
gramatické frazémy	přihlédnutí, směr, pohrůžka	3
monokolokabilní frazémy	coul, verva, lub, brva, novota, fofr, fiasko, zenit	8
přirovnání	mávnutí, bič	2
kontaktové frazémy	dovolení, prominutí	2
nominální frazémy	nedávno, taktovka, grácie, pompa, záštita, vypětí, pláštík, způsob, záminka, rozednění, nadhled, poukaz, pravděpodobnost, lehkost, úsměv, gusto, přičinění, nadsázka, rouška	19
verbální frazémy	povděk, úkos, drobnohled, podiv, součást, klepnutí, oklika, poklepání, despekt, ztělesnění, výjimka, zosobnění, rozmysl, terč, synonymum, nedočkavost, trysk, chlouba, poklička, přítěž, tahák, třešnička, leitmotiv, lákadlo, volant, předzvěst, skřípění, vynaložení, lupa, prásknutí, cíl	31
propoziční frazémy	poklus	1

Tabulka 15: Typy vyhledaných frazémů v instrumentálu singuláru

7.8 Srovnání vyhledaných frazémů dle jednotlivých pádů

V této analýze bylo analyzováno celkem 1 102 lemmat, z nichž se 306, tj. 28 %, podílí na jednom či více frazémech. Tabulka 16 shrnuje, kolik bylo u jednotlivých pádů analyzováno celkem lemmat, kolik z nich tvoří frazémy a o jaký se jedná procentuální podíl.

	S1	S2	S3	S4	S5	S6	S7	Celkem
Počet analyzovaných lemmat	162	75	252	66	201	169	177	1 102
Počet lemmat podílejících se na frazémeh	15	39	47	58	10	71	66	306
Procentuální podíl	9 %	52 %	19 %	88 %	5 %	42 %	37 %	28 %

Tabulka 16: Shrnutí výsledků dle jednotlivých pádů

Z tabulky 16 je zřetelné, že výrazně největší podíl lemmat, která tvoří frazémy, se vyskytuje ve vybraném vzorku v akuzativu singuláru (88 %). Více než poloviční podíl (52 %) má rovněž genitiv singuláru. Následuje lokál singuláru (42 %), instrumentál singuláru (37 %) a dativ singuláru (19 %). Nejmenší podíl tvoří dle očekávání nominativ singuláru (9 %) a vokativ singuláru (pouze 5 %). Z výsledků dle jednotlivých pádů navíc vyvstaly různé skupiny frazémů se stejnými vlastnostmi, které budou popsány dále. Tabulka 17 shrnuje, kolik lemmat bylo k typům v jednotlivých pádech zařazeno.

Typ	S1	S2	S3	S4	S5	S6	S7	Celkem
gramatické frazémy	0	1	0	0	0	11	3	15
monokolokabilní frazémy	1	7	3	13	0	10	8	42
binomiály	1	3	0	1	0	1	0	6
přirovnání	3	0	0	0	0	1	2	6
kontaktové frazémy	3	0	0	4	8	0	2	17
cizojazyčná spojení	3	2	0	0	0	0	0	5
nominální frazémy	2	17	4	13	0	19	19	74
verbální frazémy	2	9	40	27	1	28	31	138
propoziční frazémy	0	0	0	0	1	1	1	3

Tabulka 17: Typy frazémů dle jednotlivých pádů

V nominativu singuláru bylo analyzováno celkem 162 lemmat, většinou se jednalo o vlastní jména, případně cizojazyčná slova. Pouze u 15 lemmat byl tedy vyhledán alespoň jeden frazém, přičemž vyhledané frazémy byly velmi různorodé. V nominativu bylo nalezeno nejvíce přirovnání (*chovat se jako macho, točit se jako káča, zachovat se jako grand*) a cizojazyčných

spojení – ve všech případech se jednalo o latinská úsloví (*alma mater, Ave Maria, Ave Caesar, sine ira et studio*). Vedle běžných verbálních a nominálních frazémů byly rovněž v nominativu nalezeny binomiály (*alfa a omega*), monokolokabilní frazémy (*hlava nehlava*) a kontaktové frazémy (např. *pozdrav pánbůh, to pardon, kurva fix*). Lemma *pánbůh* je v nominativu součástí nejvíce frazémů (celkem 20 v SYN2015 a dalších 21 v SYNv11), přičemž v nalezených frazémech byla vedle kontaktové frazémy také například přísloví (např. *komu v rukou karty šustí, toho pánbůh neopustí*) či přirovnání (např. *hrát jako pánbůh*). Většina nalezených frazémů je již slovníkově zachycena, případně se jedná pouze o nezachycené varianty (např. *hlava nehlava* vedle *mlátit někoho hlava nehlava*). Slovníkově nezachycené frazémy a jejich varianty jsou znázorněny v popisu pádu, nicméně na tomto místě uvedme pouze lemmata, u nichž není ani jeden frazém slovníkově zachycen. V nominativu se jedná o lemmata *macho* (*chovat se jako macho*) a *kurva* (*kurva fix*).

Analyzovaná lemmata genitivu singuláru byla součástí alespoň jednoho frazému ve více než polovině případů, konkrétně frazémy tvoří 39 lemmat z celkových 75. Jednalo se přitom spíše o lemmata z první poloviny vzorku, například prvních dvanáct lemmat vykazovalo znaky frazému. V genitivu se ze všech pádů nejčastěji objevovaly binomiály (*z široka daleka, ani vidu, ani slechu, bez ladu a skladu*). Stejně jako v nominativu, i zde byla nalezena cizojazyčná spojení, přičemž se jednalo o latinská úsloví (*genius loci, historia magistra vitae, curriculum vitae*). V genitivu byl taktéž nalezen jeden gramatický frazém (víceslovná sekundární předložka *z hlediska*) a větší množství frazémů obsahujících monokolokabilní komponenty (*svléknout se do naha, do třetice všeho dobrého, mít co do činění s někým/něčím, nemít ani potuchy, dostat něco do vínku, rozpálit do ruda*). Nominální frazémy se v případě genitivu velmi často pojily s předložkou *bez* (*bez přestání, bez ustání, bez rozmýšlení, bez povšimnutí, bez ptaní, bez úhony, bez prodlení, bez váhání*) či *od* (*od nynějška, od nepaměti, od pradávna, od malička*), u verbálních se zase často opakovala předložka *do* (*vyšumět do ztracena, nenávidět do morku kostí, stočit se do klubička, zpít se do němoty, dostat se do ráže*). I v tomto případě byla již většina frazémů slovníkově zachycena, nezachyceny jsou některé varianty (např. *nadělit něco do vínku* vedle *dát něco do vínku*) či modifikace (*do třetice všeho zlého*). Ani jeden frazém není slovníkově zachycen u 8 lemmat: *loci* (*genius loci*), *ztraceno* (*vyšumět/vyznít do ztracena*), *sebezáchova* (*pud sebezáchovy*), *méněcennost* (*mít komplex méněcennosti*), *dřívějšek* (*z dřívějška*), *ptaní* (*bez ptaní*), *slech* (*není ani vidu ani slechu; ani slechu; nebýt po někom ani vidu ani slechu*) a *záhrobí* (*hlas ze záhrobí*).

V dativu singuláru bylo analyzováno nejvíce lemmat, a to 252, přičemž ve většině případů se jednalo o deverbální substantiva (např. *politování, nezaplacení*), která se používají především v administrativním stylu. Na frazémeh se z analyzovaných lemmat podílí celkem 47 lemmat. Jednalo se především o první lemmata v seznamu (tedy ta nejanomálnější) – například z prvních sedmnácti analyzovaných lemmat je čtrnáct součástí alespoň jednoho frazému, přičemž ze zbylých tří lemmat jsou navíc dvě chybně označovaná (*rosa* a *dceř*). Tři z vyhledaných frazémů obsahovaly monokolokabilní komponenty (*být k mání, něco k snědku, být k nevíře*) a jednou byl nalezen gramatický frazém (*k dispozici*). V dativu se rovněž neobjevovaly téměř žádné nominální frazémy (výjimku tvoří frazémy *láska k bližnímu, k úplné spokojenosti, od hlavy k patě, široké veřejnosti*). Většina nalezených frazémů tedy byly verbální frazémy. Častějším typem je spojení se slovesem (např. *být k dostání, být k nepoznání, změnit se k nerozeznání*), méně často se jedná o deverbální substantiva, která však byla ve slovnících hledána pod verbálními frazémy – jde tedy o transformace frazémů (např. *k odvrácení pozornosti* místo *odvrátit pozornost*). Téměř všechny frazémy se navíc v případě dativu pojí s předložkou *k* (např. *k dispozici, nacpaný k prasknutí, jít k čertu*). Velmi různorodé bylo lemma *čert* (v SYN2015 bylo nalezeno 8 frazémů a v SYNv11 dalších 5), které se objevovalo v kontaktoých frazémeh (*k čertu*), verbálních frazémeh (*jít k čertu*) i v přísloví (*čiň čertu dobře a peklem se ti odmění*). Použité slovníky pak v žádném frazému nezachycují lemmata *sehnání* (*být k sehnání*), *dispozice* (*k dispozici, mít / dát / dostat / dávat / zůstat k dispozici*), *nepříčetnosti* (*dovádět / rozzuřit / dohánět / přivádět / vytáčet / vydráždit / rozlítit / rozběsnit / vytočit k nepřičetnosti*), *dovršení* (*k dovršení všeho*), *koukání* (*být ke koukání*), *bližní* (*láska k bližnímu; sobě ku cti, bližnímu ku pomoci*), *normál* (*vymykat se normálu, vracet / vrátit / návrat k normálu*), *vábení* (*odolat / podlehnout / odolávat vábení*) a *zastavení* (*nebýt k zastavení*).

V akuzativu singuláru se ze všech pádů podílí na alespoň jednom frazému nejvíce lemmat, a to až 88 % z nich (tedy 58 z 66). Opakovaně se zde objevovala substantivizovaná adjektiva (např. *rozmyšlená, uvítaná, rozloučená*) či názvy dnů. V tomto pádu se vyskytuje nejvíce monokolokabilních frazémů (*dávat (si) bacha, vystavovat něco na odív, brát v potaz, postavit něco na roveň s něčím, na omak, mít zálsk na něco, v předvečer, pustit si pusu na špacír, stát majlant, dodat punc, hrát někde prim, mít šmrnc, vzdát někomu hold*). Rovněž se zde opakovaly některé kontaktové frazémy, ve většině případů pojící se s předložkou *na* (*na rozloučenou, na shledanou, na uvítanou, upřímnou soustrast, pro smilování boží*). V akuzativu se ojediněle objevovaly též binomiály (*na milost a nemilost*), gramatické frazémy (*na úkor*), pranostiky

(*Velký pátek deštivý dělá rok žízňivý*) či přirovnání (*čekat na něco/někoho jako na smilování*). Součástí nominálních frazémů byly často názvy dnů (např. *na Bílou sobotu, na Zelený čtvrtek*, ale i *ještě nějaký ten pátek*), ve verbálních frazémech se zase často objevovala již zmiňovaná substantivizovaná adjektiva (např. *být na pováženou, dát někomu čas na rozmyšlenou, dodat na vysvětlenou*). Verbální frazémy se zde navíc opakovaně pojí se slovesem *dát/dávat* (např. *dávat si bacha, dát někomu čas na rozmyšlenou, dát někomu na srozuměnou, dát někomu něco na požádání, dát něčemu punc něčeho/někoho*). Téměř všechny frazémy se pak pojí s předložkou *na* (např. *na Šedivé úterý, odložit něco na neurčito, jet na doraz, na minutku, na doživotí*). Ani jeden frazém není navíc slovníkově zachycen pouze u čtyř lemmat: *předvečer* (v *předvečer*), *vteřinka* (na *vteřinku*), *úplata* (za *úplatu*), *minutka* (na *minutku, ani na minutku, minutku po minutce*).

V případě vokativu singuláru bylo z celkem 201 analyzovaných lemmat pouze deset součástí frazému či frazémů. Ostatní lemmata byla povětšinou vlastní jména, funkce či tituly osob, případně jednoslovné vulgarismy. V případě 8 lemmat z 10 pak šlo o kontaktové frazémy (*ty vole, pane bože, a hrome, svatá prostoto, káčo pitomá, panenka skákavá, ty krávo*), v jednom případě o propoziční frazém (*andělíčku, můj strážníčku, opatruj mi mou dušičku*) a v jednom případě o verbální frazém (*babo rad*). V případě lemmatu *panenka* bylo zachyceno i jedno přísloví (*sedávej panenka v koutě, budeš-li hodná, najdou tě*). Slovníkově nezachycena jsou přitom pouze lemmata *hrom* (a *hrome*) a *kráva* (*ty krávo*).

Analyzovaná lemmata lokálu singuláru byla součástí alespoň jednoho frazému v téměř polovině případů, konkrétně tvoří frazém 71 z celkových 169 lemmat. Stejně jako v případě dativu, i zde se často jednalo o deverbální substantiva (např. *skončení, porovnání*). V tomto pádu se rovněž vyskytovalo mnoho frazémů obsahujících monokolokabilní komponent (*být ve střehu, ve skrytu duše, být v nedohlednu, v mžiku, nechat někoho ve štychu, být na špici něčeho, v hloubi duše, zmizet v propadlišti dějin, být na chvostě, být v rauši*). Lokál se však vyznačuje zejména největším počtem gramatických frazémů (celkem 11), především pojících se s předložkou *v*, případně *na* (v *rámci, v souladu (s), na pomezí, v porovnání s, na přelomu, v průběhu, ve srovnání s, v souvislosti s, na bázi, v návaznosti na, na základě*). S těmito předložkami se pojila i většina všech frazémů, které tvořila lemmata v lokálu. Dále byly mezi frazémy nalezeny rovněž binomiály (v *klidu a pohodě, na těle i na duchu*), přirovnání (*mít se jako v bavlnce, tvářit se jako na funuse*), kontaktové frazémy (v *pohodě*) či propoziční frazémy (*vítej na palubě, to je v pořádku*). V lokálu byl zároveň nalezen velký počet lemmat, u nichž není žádný frazém slovníkově zachycen (celkem 17). Jednalo se o lemmata *mžik* (v *mžiku*),

přepočet (v přepočtu), výslužba (být ve výslužbě), pomezí (na pomezí), vzduchoprázdno (žít ve vzduchoprázdnu), mezičas (v mezičase), mezidobí (v mezidobí), současnost (v současnosti), krátkost (v krátkosti), porovnání (v porovnání s), obležení (být v obležení), přelom (na přelomu (roku/století)), stručnost (ve stručnosti), průběh (v průběhu), návaznost (v návaznosti na), rauš (být v rauši) a průměr (v průměru).

V instrumentálu singuláru bylo z celkem 177 analyzovaných lemmat 66 součástí alespoň jednoho frazému, a to především první lemmata v seznamu (například z prvních 15 lemmat vykazovalo znaky frazému celkem 11 z nich). I v instrumentálu se objevilo několik gramatických frazémů (*s přihlédnutím k, směrem do/k/na/od/proti, pod záštitou, pod pláštěm, pod zámkou*), dvě přirovnání (*jako mávnutím kouzelného proutku, jako když bičem mrská*), dva kontaktní frazémy (*s dovolením, s prominutím*) či propoziční frazém (*poklusem klus*). Nalezeno bylo rovněž několik monokolokabilních frazémů (*každým coulem, s vervou, mít něco za lubem, nehnout brvou, zářit novotou, končit fiaskem, být za zenitem*). Většina nominálních i verbálních frazémů se pak pojila s předložkou *s* (např. *s vervou, dělat něco s rozmyslem*), případně *pod* (např. *pod taktovkou něčeho, držet něco pod pokličkou*). V instrumentálu bylo zároveň nalezeno nejvíce lemmat, u nichž není žádný frazém slovníkově zachycen (celkem 23): *úkos (podívat se / pohlédnout / dívat se / hledět / pohlížet / sledovat úkosem), součást (být / stát se nedílnou součástí), grácie (s grácií), pompa (s velkou pompou), despekt (pohlížet / vyjadřovat se / hledět / shlížet / nahlížet / dívat se s despektem), zosobnění (stát se / být zosobněním něčeho), pohružka (pod pohružkou), synonymum (stát se / být synonymem pro něco), trysk (rozběhnout se / řítit se / vyrazit / pádit / uhánět tryskem), chloubka (být / stát se chloubou někoho/něčeho), novota (zářit novotou), přítěž (stávat se / stát se / být přítěží), rozednění (před/s rozedněním), tahák (stát se / být (hlavním) tahákem), nadhled (s nadhledem, brát s nadhledem), poukaz (s poukazem na), fofr (fofrem, vzít to fofrem), lákadlo (být lákadlem), lehkost (s lehkostí), volant (sedět/usnout za volantem), předzvěst (stát se / být předzvěstí něčeho), fiasko (skončit/končit fiaskem) a nadsázka (s nadsázkou, říci/říkat s nadsázkou).*

Závěr

Hlavním cílem práce bylo zjistit, zda gramatické anomálie substantiv (tj. nezvykle vysoká frekvence jednoho či více tvarů v rámci morfologického paradigmatu) mohou ukazovat na formálně anomální víceslovné lexémy (frazémy). Tato souvislost z analýzy skutečně vyplynula – celkem 28 % analyzovaných lemmat (306 z celkových 1 102) je součástí alespoň jednoho frazému. Výsledné číslo však velmi snižují lemmata s anomálií v nominativu a vokativu, která jsou součástí frazémů jen z 9 % a 5 %, což je dáno především vlastními jmény v seznamu analyzovaných lemmat. Ve zbylých pěti pádech však analyzovaná lemmata tvoří frazémy z 19–88 %, z čehož již souvislost mezi gramatickou anomálií a frazémů vyplývá.

Další výzkumnou otázkou bylo, zda se v materiálu projeví rozdíly mezi jednotlivými pády jednotného čísla. Gramatické anomálie a formálně anomální víceslovné lexémy spolu nejvíce souvisejí v akuzativu singuláru (lemmata tvoří frazémy až z 88 %), dále v genitivu singuláru (52 %), lokálu singuláru (42 %), instrumentálu singuláru (37 %) a dativu singuláru (19 %). Naopak téměř vůbec spolu nesouvisejí v nominativu singuláru (9 %) a vokativu singuláru (5 %).

Z výsledků analýzy dále vyplynulo, že výrazně nejčastějšími typy frazémů souvisejících s gramatickou anomálií jsou verbální frazémy (např. *nebýt k zahození*), které se objevují ve všech pádech, nejtypičtější jsou však pro dativ, v němž se neobjevovaly téměř žádné jiné typy frazémů. Ve všech pádech se rovněž vyskytovaly nominální frazémy (např. *nebezpečí z prodlení*) a frazémy s monokolokabilním komponentem, kterých bylo nalezeno nejvíce v akuzativu (např. *dávat bacha*). Gramatické frazémy (konkrétně víceslovné předložky) se objevovaly především v lokálu (např. *v souvislosti s*), kontaktové frazémy ve vokativu (např. *a hrome*), cizojazyčná spojení pouze v nominativu a akuzativu (např. *alma mater*) a propoziční frazémy byly nalezeny pouze dva, a to v lokálu a instrumentálu (např. *vítej na palubě*). Ojediněle byla mezi frazémy nalezena rovněž přirovnání, nejvíce v nominativu (např. *být jako pánbůh*), a binomiály, především v genitivu (např. *bez ladu a skladu*).

Přínosem práce tedy je potvrzení smysluplnosti systematického výzkumu frazémů na základě anomálie v distribuci tvarů lemmatu. Jelikož se v tomto případě jednalo pouze o sondu, do budoucna by se nabízelo provést výzkum komplexněji, na větším množství dat, ale především i na substantivech v plurálu. Vzhledem k plánovanému rozšíření GramatiKatu se rovněž nabízí podobný výzkum adjektiv a verb.

Seznam použité literatury

Akademický slovník současné češtiny [online]. Praha: Ústav pro jazyk český AV ČR, v. v. i, 2017–2023. Dostupný z: <https://slovníkcestiny.cz>.

BARTOŇ, Tomáš, CVRČEK, Václav, ČERMÁK, František, JELÍNEK, Tomáš a PETKEVIČ, Vladimír. *Statistiky češtiny*. Praha: Nakladatelství Lidové noviny, 2009. ISBN 978-80-7106-594-4.

CVRČEK, Václav a kol. *Mluvnice současné češtiny 1: Jak se píše a jak se mluví*. Praha: Karolinum, 2010. ISBN 978-80-246-1743-5.

CVRČEK, Václav a RICHTEROVÁ, Olga (ed.). cnk:syn:verze11. *Příručka ČNK* [online]. 2023a [cit. 28. 11. 2023]. Dostupné z: <https://wiki.korpus.cz/doku.php/cnk:syn:verze11>.

CVRČEK, Václav a RICHTEROVÁ, Olga (ed.). cnk:syn2015. *Příručka ČNK* [online]. 2022 [cit. 16. 5. 2023]. Dostupné z: <https://wiki.korpus.cz/doku.php/cnk:syn2015>.

CVRČEK, Václav a RICHTEROVÁ, Olga (ed.). kurz:hledani_kolokaci. *Příručka ČNK* [online]. 2021a [cit. 17. 5. 2023]. Dostupné z: https://wiki.korpus.cz/doku.php/kurz:hledani_kolokaci.

CVRČEK, Václav a RICHTEROVÁ, Olga (ed.). manualy:gramatikat. *Příručka ČNK* [online]. 2021b [cit. 16. 5. 2023]. Dostupné z: <https://wiki.korpus.cz/doku.php/manualy:gramatikat>.

CVRČEK, Václav a RICHTEROVÁ, Olga (ed.). pojmy:kolokace. *Příručka ČNK* [online]. 2023b [cit. 17. 5. 2023]. Dostupné z: <https://wiki.korpus.cz/doku.php/pojmy:kolokace>.

CVRČEK, Václav a RICHTEROVÁ, Olga (ed.). pojmy:monokolokabilita. *Příručka ČNK* [online]. 2015 [cit. 10. 12. 2023]. Dostupné z: <https://wiki.korpus.cz/doku.php/pojmy:monokolokabilita>.

CVRČEK, Václav a RICHTEROVÁ, Olga (ed.). pojmy:asociacni_miry. *Příručka ČNK* [online]. 2019 [cit. 17. 5. 2023]. Dostupné z: https://wiki.korpus.cz/doku.php/pojmy:asociacni_miry.

CVRČEK, Václav a RICHTEROVÁ, Olga (ed.). seznamy:frazemy. *Příručka ČNK* [online]. 2021c [cit. 16. 5. 2023]. Dostupné z: <https://wiki.korpus.cz/doku.php/seznamy:frazemy>.

ČECHOVÁ, Marie, DOKULIL, Miloš, HLAVSA, Zdeněk, HRBÁČEK, Josef a HRUŠKOVÁ, Zdeňka. *Čeština – řeč a jazyk*. Praha: SPN, 2011. ISBN 978-80-7235-413-9.

ČERMÁK, František a kol. *Slovník české frazeologie a idiomatiky 4: Výrazy větné*. Praha: Leda, 2009. ISBN 978-80-7335-219-6.

ČERMÁK, František, HRONEK, Jiří a kol. *Slovník české frazeologie a idiomatiky 2: Výrazy neslovesné*. Praha: Leda, 2009a. ISBN 978-80-7335-217-2.

ČERMÁK, František, HRONEK, Jiří a kol. *Slovník české frazeologie a idiomatiky 3: Výrazy slovesné*. Praha: Leda, 2009b. ISBN 978-80-7335-218-9.

- ČERMÁK, František, HRONEK, Jiří a kol. *Slovník české frazeologie a idiomatiky 1: Přirovnání*. Praha: Leda, 2009c. ISBN 978-80-7335-216-5.
- ČERMÁK, František. Binomiál. In: KARLÍK, Petr, NEKULA, Marek a PLESKALOVÁ, Jana (ed). *CzechEncy – Nový encyklopedický slovník češtiny* [online]. Brno: 2017c [cit. 10. 12. 2023]. Dostupné z: <https://www.czechency.org/slovník/BINOMIÁL>.
- ČERMÁK, František. *Frazeologie a idiomatika: česká a obecná = Czech and General Phraseology*. Praha: Karolinum, 2007b. ISBN 978-80-246-1371-0.
- ČERMÁK, František. Frazeologie a idiomatika. In: PLESKALOVÁ, Jana a kol. *Kapitoly z dějin české jazykové bohemistiky*. Praha: Academia, 2007a, s. 243–248. ISBN 978-80-200-1523-5.
- ČERMÁK, František. Gramatický frazém. In: KARLÍK, Petr, NEKULA, Marek a PLESKALOVÁ, Jana (ed). *CzechEncy – Nový encyklopedický slovník češtiny* [online]. Brno: 2017a [cit. 10. 12. 2023]. Dostupné z: https://www.czechency.org/slovník/GRAMATICKÝ_FRAZÉM.
- ČERMÁK, František. *Idiomatika a frazeologie češtiny*. Praha: Univerzita Karlova, 1982.
- ČERMÁK, František. *Lexikon a sémantika*. Praha: Nakladatelství Lidové noviny, 2010. ISBN 978-80-7422-020-3.
- ČERMÁK, František. Monokolokabilita. In: KARLÍK, Petr, NEKULA, Marek a PLESKALOVÁ, Jana (ed). *CzechEncy – Nový encyklopedický slovník češtiny* [online]. Brno: 2017b [cit. 10. 12. 2023]. Dostupné z: <https://www.czechency.org/slovník/MONOKOLOKABILITA>.
- DIVJAK, Dagmar a GRIES, Stefan Th. Ways of trying in Russian: clustering behavioral profiles. *Corpus linguistics and linguistics theory*. De Gruyter. 2006, 2(1), 23–60.
- HAVRÁNEK, Bohuslav a JEDLIČKA, Alois. *Český mluvnické*. Praha: Státní pedagogické nakladatelství, 1960.
- HLADKÁ, Zdeňka. FRAZEOLOGICKÝ SLOVNÍK. In: KARLÍK, Petr, NEKULA, Marek a PLESKALOVÁ, Jana (ed.). *CzechEncy – Nový encyklopedický slovník češtiny* [online]. Brno: Masarykova univerzita, 2017 [cit. 17. 5. 2023]. Dostupné z: https://www.czechency.org/slovník/FRAZEOLOGICKÝ_SLOVNÍK.
- HLADKÁ, Zdeňka. Lexikologie. In: KARLÍK, Petr, NEKULA, Marek, RUSÍNOVÁ, Zdenka a kol. *Příruční mluvnické češtiny*. Praha: Nakladatelství Lidové noviny, 1996, s. 65–108. ISBN 80-7106-303-7.
- HLADKÁ, Zdeňka. Slovník české frazeologie a idiomatiky. *Slovo a slovesnost* [online]. 1995, 56(4), s. 285–291. Dostupné z: <http://sas.ujc.cas.cz/archiv.php?art=3649>.
- ISAČENKO, Alexandr Vasilijevič. *Grammatičeskij stroj russkogo jazyka v sopostavlenii s slovackim. Čast vtoraja: morfologija*. Bratislava: Izdatel'stvo akademii nauk, 1960.

JANDA, Laura a LYASHEVSKAYA, Olga. Grammatical profiles and the interaction of the lexicon with aspect, tense, and mood in Russian. *Cognitive linguistics* 22. 2011, (4), 717–762. DOI 10.1515/cogl.2011.027

JANDA, Laura a TYERS, M. Francis. Less is more: why all paradigms are defective, and why that is a good thing. *Corpus linguistics and linguistic theory*. De Gruyter. 2021, 17(1), 109–141.

JANDA, Vojtěch. *Frekvenční distribuce nominální flexe v češtině*. Praha, 2017. Bakalářská práce. Univerzita Karlova, Filozofická fakulta, Ústav obecné lingvistiky. Vedoucí práce Křivan, Jan.

JELÍNEK, Jaroslav, BEČKA, Josef Václav a TĚŠITELOVÁ, Marie. *Frekvence slov, slovních druhů a tvarů v českém jazyce*. Praha: Státní pedagogické nakladatelství, 1961.

JELÍNEK, Tomáš, KOPŘIVOVÁ, Marie, PETKEVIČ, Vladimír a SKOUMALOVÁ, Hana. *Variabilita českých frazémů v úzu*. Časopis pro moderní filologii. 2018, 100(2), 151–175.

KALIVODOVÁ, Veronika. Informace o slovnících a podrobný návod. *DEB: Prohlížeč a editor slovníků* [online]. Brno: Centrum zpracování přirozeného jazyka MU [cit. 11. 12. 2023]. Dostupné z: https://deb.fi.muni.cz/download/debdict_manual.pdf.

KARLÍK, Petr, NEKULA, Marek, RUSÍNOVÁ, Zdenka a GREPL, Miroslav. *Příruční mluvnice češtiny*. Praha: Nakladatelství Lidové noviny, 2012. ISBN 978-80-7106-624-8.

KODÝTEK, Vilém. Podstatná jména (substantiva). In: CVRČEK, Václav a kol. *Mluvnice současné češtiny 1: Jak se píše a jak se mluví*. Praha: Karolinum, 2010, s. 134–195. ISBN 978-80-246-1743-5.

KOCHOVÁ, Pavla a OPAVSKÁ, Zdeňka (ed.). *Kapitoly z koncepce Akademického slovníku současné češtiny*. Praha: Ústav pro jazyk český AV ČR, 2016. ISBN 978-80-86496-89-4.

KOMÁREK, Miroslav, KOŘENSKÝ, Jan, PETR, Jan a VESELKOVÁ, Jarmila. *Mluvnice češtiny 2: Tvarosloví*. Praha: Academia, 1986.

KOPŘIVOVÁ, Marie. Lexikologie. In: CVRČEK, Václav a kol. *Mluvnice současné češtiny 1: Jak se píše a jak se mluví*. Praha: Karolinum, 2010, s. 65–79. ISBN 978-80-246-1743-5.

KOVÁŘÍKOVÁ, Dominika. Co je v ČNK nového XI. *Korpus – gramatika – axiologie*. 2021a, 24(1), 74–76.

KOVÁŘÍKOVÁ, Dominika. *Sharing data through specialized corpus-based tools: the case of GramatiKat*. Jazykovedný časopis. 2021b, 72(2), 531–544.

KUZNETSOVA, Julia. *Linguistic Profiles: Going from Form to Meaning via Statistics*. Berlin: De Gruyter Mouton, 2015. ISBN 3-11-039348-4.

LÁZNIČKA, Michal. Jmenná flexe v češtině: životnost, rod a gramatické profily. In: *Sborník abstraktů CALC-1: 1. konference České asociace kognitivní lingvistiky*. Praha: Česká asociace kognitivní lingvistiky, 2019.

O slovníku. *Akademický slovník současné češtiny* [online]. Praha: Ústav pro jazyk český AV ČR, v. v. i, 2017–2023. Dostupný z: <https://slovníkcestiny.cz>.

PETKEVIČ, Vladimír, KOPŘIVOVÁ, Marie, HNÁTKOVÁ, Milena, JELÍNEK, Tomáš, KOPŘIVA, Pavel, ROSEN, Alexandr, SKOUMALOVÁ, Hana a VONDŘIČKA, Pavel. *Typologie víceslovných jednotek v češtině a frekvenční zastoupení jejich hlavních vlastností v žánrově vyváženém korpusu*. Studie z aplikované lingvistiky. 2020, 11(2), 37–62.

ŠICHOVÁ, Kateřina a ŠEMELÍK, Martin. Slovník české frazeologie a idiomatiky. Jazykovedný časopis. 2017, 68(3), 491–499.

ŠTÍCHA, František a kol. *Velká akademická gramatika spisovné češtiny*. Praha: Academia, 2018. ISBN 978-80-200-2719-1.

TĚŠITELOVÁ, Marie a kol. *Kvantitativní charakteristiky současné češtiny*. Praha: Academia, 1985.

TRÁVNÍČEK, František. *Mluvnice spisovné češtiny, část 1: Hláskosloví, tvoření slov, tvarosloví*. Praha: Melantrich, 1948.

Jazykové korpusy a nástroje

CVRČEK, Václav. *Calc: Korpusová kalkulačka*. Praha: Ústav Českého národního korpusu FF UK, 2019. Dostupný z: <https://www.korpus.cz>.

ČERMÁK, František, DOLEŽALOVÁ-SPOUSTOVÁ, Drahomíra, HLAVÁČOVÁ, Jaroslava, HNÁTKOVÁ, Milena, JELÍNEK, Tomáš, KOCEK, Jan, KOPŘIVOVÁ, Marie, KŘEN, Michal, NOVOTNÁ, Renata, PETKEVIČ, Vladimír, SCHMIEDTOVÁ, Věra, SKOUMALOVÁ, Hana, ŠULC, Michal a VELÍŠEK, Zdeněk. *SYN2005: žánrově vyvážený korpus psané češtiny*. Praha: Ústav Českého národního korpusu FF UK, 2005. Dostupný z: <https://korpus.cz>.

GOLÁŇOVÁ, Hana, WACLAWIČOVÁ, Martina, PEJCHA, Jakub, ČAPKA, Tomáš a BENEŠOVÁ, Lucie. *Mapka: mapová aplikace pro korpusy mluvené češtiny. Verze 2.0*. Praha: Ústav Českého národního korpusu FF UK, 2023. Dostupný z: <https://www.korpus.cz>.

HORÁK, Aleš a RAMBOUSEK, Adam. PRALED – A New Kind of Lexicographic Workstation. In: PRZEPIÓRKOWSKI, Adam, PIASECKI, Maciej, JASSEM, Krzysztof, FUGLEWICZ, Piotr, ed. *Studies in Computational Linguistic: Applications*. Berlín: Springer, 2013, s. 131–141. ISBN 978-3-642-34398-8.

KOVÁŘÍKOVÁ, Dominika a KOVÁŘÍK, Oleg. *GramatiKat: Nástroj pro výzkum gramatických kategorií a gramatických profilů*. Praha: Ústav Českého národního korpusu FF UK, 2021. Dostupný z: <https://korpus.cz/gramatikat>.

KŘEN, Michal, CVRČEK, Václav, ČAPKA, Tomáš, ČERMÁKOVÁ, Anna, HNÁTKOVÁ, Milena, CHLUMSKÁ, Lucie, JELÍNEK, Tomáš, KOVÁŘÍKOVÁ, Dominika, PETKEVIČ, Vladimír, PROCHÁZKA, Pavel, SKOUMALOVÁ, Hana, ŠKRABAL, Michal, TRUNEČEK, Petr, VONDŘIČKA, Pavel a ZASINA, Adrian. *SYN2015: reprezentativní korpus psané češtiny*. Praha: Ústav Českého národního korpusu FF UK, 2015. Dostupný z: <https://korpus.cz>.

KŘEN, Michal, CVRČEK, Václav, HENYŠ, Jan, HNÁTKOVÁ, Milena, JELÍNEK, Tomáš, KOCEK, Jan, KOVÁŘÍKOVÁ, Dominika, KŘIVAN, Jan, MILIČKA, Jiří, PETKEVIČ, Vladimír, PROCHÁZKA, Pavel, SKOUMALOVÁ, Hana, ŠINDLEROVÁ, Jana a ŠKRABAL, Michal. *SYN2020: reprezentativní korpus psané češtiny*. Praha: Ústav Českého národního korpusu FF UK, 2020. Dostupný z: <https://www.korpus.cz>.

KŘEN, Michal, CVRČEK, Václav, HNÁTKOVÁ, Milena, JELÍNEK, Tomáš, KOCEK, Jan, KOVÁŘÍKOVÁ, Dominika, KŘIVAN, Jan, MILIČKA, Jiří, PETKEVIČ, Vladimír, PROCHÁZKA, Pavel, SKOUMALOVÁ, Hana, ŠINDLEROVÁ, Jana a ŠKRABAL, Michal. *Korpus SYN, verze 11*. Praha: Ústav Českého národního korpusu FF UK, 2022. Dostupný z: <https://www.korpus.cz>.

MACHÁLEK, Tomáš. *Slovo v kostce – agregátor slovních profilů*. Praha: Ústav Českého národního korpusu FF UK, 2019. Dostupný z: <https://www.korpus.cz>.

Přílohy

Příloha 1: Soupis všech analyzovaných lemmat z GramatiKatu

	S1	S2	S3	S4	S5	S6	S7
1	amen	daleko	mání	bach	mylord	česko	povděk
2	min	loci	dostání	odiv	děvenka	relativno	polohlas
3	nehlava	naho	nepoznání	počest	vůl	neskutečno	šept
4	orchestra	přestání	nerozeznání	potaz	pán	střeh	prostřednictví
5	dodo	vitae	zahození	povážená	signor	odkvět	přihlednutí
6	back	ztraceno	snědek	roveň	drahoušek	skryt	nedávno
7	bionomie	nynějšek	sehnání	rozloučená	viceprezidentka	tichost	mávnutí
8	judita	třetice	nezaplacení	shledaná	hajzl	domnění	úkos
9	graham	ustání	dceř	úkor	brach	posledek	coul
10	konrád	nepaměť	dispozice	uvítaná	kámoš	upečení	drobnohled
11	allen	činění	zbláznění	rozmyšlená	zajda	skončení	taktovka
12	lambda	pradávnost	prasknutí	ponětí	čubka	rámec	trnutí
13	čon	minulo	uctění	vysvětlená	slečinka	tuzemsko	verva

14	briefing	studeno	rosa	srozuměná	beruška	sklonek	lub
15	fokus	pasus	šlehání	omak	sir	nedohledno	podiv
16	sedláček	diabet	vidění	přítrž	bača	soulad	usnutí
17	čížek	sebezáchova	nevíra	spadnutí	inspektorka	zaschnutí	prizma
18	dvořák	maličko	poškrábání	oplátka	miláček	mžik	g
19	ken	myokard	šarka	neprospěch	píča	pokraj	určitost
20	lenka	potucha	hapsání	zálsk	lhářka	rozmezí	kývnutí
21	magda	beztíže	nepřičetnost	předvečer	zmrđ	reál	součást
22	siemens	hledisko	přesila	chvilinka	soudružka	přepočet	krempa
23	veronika	méněcennost	drahoušek	fous	trouba	duch	nedopatření
24	adéla	vínek	zamyšlení	požadání	buzerant	předklon	klepnutí
25	běda	vnitro	brouček	rámě	šmejđ	štych	oklika
26	bobby	neznámo	prošetření	půl	matička	výslužba	vitro
27	dom	rozmýšlení	zapamatování	špacír	nána	náručí	žuchnutí
28	rumunsko	ražení	loňsko	vrub	slečna	pomezí	sedma
29	blank	proveditelnost	tao	opodstatnění	prdelka	vzduchoprázdno	středobod

30	kanada	dřívějšek	stažení	potkání	brácha	mezičas	poklepání
31	linda	rudo	nadřízení	schovávaná	bůh	mezidobí	brva
32	alma	vycházení	nahlédnutí	vnadidlo	tribun	přestrojení	leknutí
33	sára	žábronožka	ochutnání	vteřinka	dceruška	současnost	grácie
34	info	morek	rozluštění	chvilíčka	kretén	krátkost	předstih
35	weber	povšimnutí	koroze	majlant	mizera	palubovka	pompa
36	copyright	tělovýchova	odvrácení	vyžádání	rotmistr	pořádek	tonik
37	durman	představenstvo	chasa	tušení	kočička	porovnání	despekt
38	havel	ptaní	dokonalost	vybraná	komoří	vychladnutí	úklona
39	stefan	provenience	otka	píď	palatin	obležení	pokrčení
40	zuzka	slonovina	vniknutí	prospěch	sestřenka	přelom	postižení
41	čína	všehomír	dovršení	punc	blb	podřep	znevýhodnění
42	hradec	klubíčko	miláček	sobota	mrcha	polospánek	samozřejmost
43	iris	promile	předkládání	úterý	starouš	uplynutí	sebezapření
44	odrážka	úhona	přehodnocení	čtvrtek	strejda	výsluní	záštita
45	ukrajina	němota	přehřátí	pondělí	vrchní	zámotek	levačka

46	benjamín	slech	vyzvednutí	prim	blbeček	odečtení	ó
47	hana	terra	znehodnocení	šmrnc	centurion	sousedství	započetí
48	koukej	trula	zodpovězení	pátek	coura	stručnost	ztělesnění
49	op	joint	čert	pozor	debil	obdržení	izolepa
50	alan	prodlení	naturel	smilování	doktorka	otka	poklus
51	journal	málo	navození	úplata	hrom	pleso	povzdech
52	krysař	stalinismus	projednání	neurčito	parchant	pološero	svépomoc
53	macho	státnost	dosažení	soustrast	píka	špice	devíza
54	matěj	blízko	kapuca	středa	signora	krajíček	spolubydlící
55	sparta	relativita	poražení	rov	tetička	podstata	chodeckost
56	beta	skot	var	peníz	brouček	půle	oběšení
57	egypt	venek	záhuba	náskok	chudinka	bavlnka	stisknutí
58	real	záhrobí	zviditelnění	chvilka	doktor	průběh	tematika
59	ben	č	koukání	hold	excelence	setmění	výjimka
60	klid'as	váhání	moloch	pláně	inspektor	srovnání	zosobnění
61	lodivod	zeměměřictví	nelibost	zpátečka	jonathan	blízkost	pečlivost

62	mama	konečník	ořešák	doraz	major	ich-forma	rozmysl
63	man	školství	přehřívání	minutka	prefekt	absolvování	směr
64	mars	donucení	vyjasnění	doživotí	taťka	civil	terč
65	spolehnutí	epopeje	nalezení	fort	tupec	fidlovačka	kliknutí
66	zvonek	lad	objasnění	nemilost	astra	úpatí	podtitul
67	cab	nerez	obohacení		bogomil	dál	pohružka
68	donna	odnětí	pokoušení		bráška	foyer	pseudonym
69	editorial	rádium	prokázání		chudáček	hloub	synonymum
70	henry	červeno	utlumení		děda	mezivrství	vypětí
71	jonathan	metro	vzteklina		děvka	předjaří	nedočkavost
72	kargo	promile	missa		pitomec	prostudování	škubnutí
73	mluvčí	ráže	následování		seržant	podpaží	vatelín
74	nexus	šestinedělí	starostová		sestřička	řad	kopnutí
75	zeman	století	vykonání		srab	rozpuk	kapuce
76	dakota		výročí		abbé	pozadí	kurzíva
77	kulič		zkvalitnění		dakota	přečtení	mávání

78	romana		apartheid		dron	přízemí	pojítka
79	drive		aréna		grázl	souvislost	plášťík
80	engineering		bližní		holčička	báze	sudí
81	jack		hnízdění		hrabě	poklid	trysk
82	kuba		kolorit		kadet	případ	chlouba
83	shift		kondenzace		magistra	rozcestí	čumák
84	special		normál		milá	mizina	příspění
85	union		poškození		navážka	plac	úsmešek
86	august		překonání		neřád	propadliště	způsob
87	blanka		prosazení		šikovatel	tajnost	poklička
88	guide		prvák		sráč	zajetí	postrach
89	pánbůh		rozpoznání		babi	gesce	záminka
90	analytička		synek		chlapec	návaznost	novota
91	rabi		uskutečnění		flip	pořadí	petrolej
92	rodačka		zabránění		hoch	požití	pravačka
93	sun		zamezení		hospodin	předehrávka	příložení

94	alfa		zhoršení		kapitán	čtvrtletí	přítež
95	amerika		zhoršování		kráska	předsálí	přízvuk
96	epilog		zklidnění		markýz	přílet	rozednění
97	lelek		zmírnění		mars	rozstřel	štěteček
98	line		zničení		prevít	základ	tahák
99	logos		establishment		prostota	závětrí	třešnička
100	magazine		narušení		reverend	antuka	úšklebek
101	mužik		páchání		velekněžka	ozdobení	věřící
102	ara		políbení		bastard	paluba	dosazení
103	ave		přezimování		blbec	pohoda	leitmotiv
104	board		prohloubení		detektiv	rozpor	lemování
105	court		prosvětlení		duce	umytí	m
106	development		soudruh		kníže	výše	mačeta
107	jiřina		ujišťování		madam	bezvědomí	nadhled
108	primárka		usazování		milý	budoucnost	poukaz
109	air		vyvolání		ničema	chvost	přidání

110	chanel		zefektivnění		poručík	kontext	přípona
111	dlouhán		brach		pusinka	podpalubí	razance
112	golem		devastace		schovanka	protisměr	ukazováček
113	hospital		dopadení		soukmenovec	střížna	vyznačení
114	kivi		hromadění		struska	ústraní	dějiště
115	kolář		mha		synek	vyrušení	dovolení
116	magna		naplnění		buchtíčka	hrazda	dychtivost
117	místostarostka		narušování		čarodějka	podoba	garant
118	pardon		odvedení		chytrák	předpokoj	milimetr
119	píd'alka		opotřebení		elfka	přítmí	nenucenost
120	rebus		převrácení		generál	rauš	poklice
121	transit		šílenství		id	sklopení	pravděpodobnost
122	book		upevnění		idiot	frita	ulehčení
123	freeware		upevňování		káča	jaro	db
124	jojo		vyřešení		kačenka	souš	fofr
125	jr		vyzkoušení		kamarád	znění	funění

126	káča		zpochybnění		karotka	cizina	hadřík
127	koordinátorka		byte		kastelán	léto	lákadlo
128	plus		chudáček		mademoiselle	područí	lehkost
129	reb		degradace		mamá	rožeň	patronka
130	servis		eskalace		mim	autonehoda	půl
131	sos		jubileum		monsieur	branže	šlehačka
132	starostka		namáhání		panenka	hantýrka	úsměv
133	abbé		politování		plukovník	kanape	volant
134	blond'ák		potěcha		praporčík	kóma	brčko
135	director		přemnožení		satan	lotyština	deviza
136	ester		rozbití		svině	mezipatro	dojetí
137	ferdinand		sblížení		táta	náměstí	doveda
138	františek		uveřejnění		víla	podpaží	hybatel
139	grand		vábení		vojín	průměr	klus
140	haló		zlepšení		arbitr	prvák	odlet
141	job		celulitida		baba	sestavování	patronát

142	kouč		dosahování		bratříček	úmysl	potisk
143	kurva		ionizace		čtenářka	voziček	předzvěst
144	lincoln		odmocnina		čurák	zámoří	prudkost
145	mellitus		páření		dobrodinec	funus	rekuperace
146	oštěpařka		potlačení		dom	ilegalita	skřípění
147	psycholožka		prozkoumání		frajer	jeviště	ukazovák
148	vesna		sepsání		guvernér	nábřeží	vařečka
149	android		uhrazení		kaprál	nadloží	vazelína
150	balík		urychlení		kurva	podhůří	vynaložení
151	bill		usmrcení		lord	pohraničí	zaujetí
152	čahoun		uštknutí		luna	pololetí	artikl
153	datum		vymizení		pater	popředí	gusto
154	hall		vyschnutí		pilates	sečtení	lupa
155	juan		vytvoření		pokrytec	suterén	prásknutí
156	leader		zadlužování		polda	tíseň	příčinění
157	mam		zatracení		potvora	absolutorium	vstříkování

158	místostarosta		adopce		primitiv	aula	apex
159	rychlobruslařka		bratřiček		princezna	defenzíva	autismus
160	studio		disociace		rada	opilost	azbuka
161	violka		encefalitida		sob	poschodí	bič
162	zastupitelka		kněžství		synáček	předpolí	cíl
163			nastartování		zbabělec	příšeří	dominanta
164			obezřetnost		zmetek	půlnoc	fiasko
165			pítí		almara	rovinka	idnes
166			pobavení		andělíček	samota	mísení
167			poškození		andílek	vyjmutí	nadsázka
168			překračování		archivář	zákulisí	nevýhoda
169			přetížení		chlapeček	zapnutí	oddenek
170			pronikání		darebák		odůvodnění
171			proplacení		děvečka		podšívka
172			rozvinutí		doňa		pousmání
173			schválení		fantom		přeplňování

174			snížení		filip		prominutí
175			snižování		fízl		rouška
176			uskladnění		hulvát		úhlopříčka
177			usnadnění		ignorant		zenit
178			utrácení		imperátor		
179			vyplutí		kráva		
180			zastavení		květinka		
181			zděšení		lenoch		
182			zpomalení		lump		
183			bojkot		maminka		
184			dobro		negr		
185			formulování		nimrod		
186			hniloba		panáček		
187			kamelot		peptid		
188			korupce		pisálek		
189			lidskost		plac		

190			neplacení		pobuda		
191			novinařina		profesor		
192			odeslání		stefan		
193			posílení		strážmistr		
194			praporčík		surovec		
195			rabování		švagr		
196			rekombinace		svobodník		
197			rover		teta		
198			rozštěpení		úchyl		
199			šípek		um		
200			spokojenost		výpravčí		
201			stíhání		výsost		
202			terorismus				
203			udržení				
204			vyhledání				
205			vyplnění				

206			vystřízlivění				
207			zablokování				
208			zánik				
209			zavlažování				
210			zdokonalení				
211			zhlédnutí				
212			získání				
213			zpřesnění				
214			zvýšení				
215			demolice				
216			disko				
217			dřívějšek				
218			exmanželka				
219			fašismus				
220			hrobník				
221			ježíšek				

222			kontaminace				
223			kůrovec				
224			odhalení				
225			opuštění				
226			ovládnutí				
227			oxidace				
228			part				
229			pata				
230			pochopení				
231			potírání				
232			přepracování				
233			přespání				
234			recyklace				
235			rezignace				
236			rozjímání				
237			sebepoznání				

238			stabilizace				
239			uchování				
240			úhona				
241			uvolňování				
242			veřejnost				
243			vstřebávání				
244			vykonávání				
245			vykreslení				
246			zakrytí				
247			zastrašování				
248			závěr				
249			zdůraznění				
250			zotavení				
251			zpochybňování				
252			zvýraznění				