

Univerzita Karlova
Filozofická fakulta
Ústav českého jazyka a teorie komunikace
Filologie – Český jazyk

Disertační práce

Mgr. Karolína Vyskočilová

Analytické predikáty s kategoriálním slovesem a jejich využití pro morfologickou anotaci

*Light verb constructions
and their exploitation for morphological annotation*

Vedoucí práce: doc. RNDr. Vladimír Petkevič, CSc.

2022

Poděkování

Poděkování patří mému milému školiteli, doc. RNDr. Vladimíru Petkevičovi, CSc., a to nejen za cenné rady a vedení disertační práce, ale i za laskavou podporu a pochopení, se kterým mne provázel doktorským studiem. Děkuji také své partnerce Evě, která mi byla po celou dobu neskutečnou oporou i pozorným prvním čtenářem. Dále děkuji svým prarodičům, babičce za všechnu důvěru a dědovi za to, že mi ukázal krásu vědy.

Prohlášení

Prohlašuji, že jsem disertační práci napsala samostatně s využitím pouze uvedených a řádně citovaných pramenů a literatury a že práce nebyla využita v rámci jiného vysokoškolského studia či k získání jiného nebo stejného titulu.

V Praze dne 25. 9. 2022

Mgr. Karolína Vyskočilová, v. r.

Abstrakt

Disertační práce se zaměřuje na analytický predikát s kategoriálním slovesem (dále APKS), tedy spojení jako *provádět kontrolu* nebo *chovat úctu*. Ilustruje, jak zúročit teoretické poznání těchto spojení v praxi, vytěžit ho při morfologické disambiguaci, a tím potenciálně ovlivnit i syntaktickou analýzu.

Teoretická část práce pokrývá tři oblasti – APKS, značkování korpusů a pravidlový disambiguační systém LanGr. APKS charakterizuje včetně kritérií pro jejich určování, popisuje současný stav jejich poznání a usouvztažňuje příspěvky k tématu, které vyšly v průběhu posledních patnácti let; důraz je přitom kladen zejména na syntaktický pohled na tyto konstrukce. Zároveň práce mapuje existující soupisy těchto spojení. Následuje popis procesu značkování psaných korpusů Českého národního korpusu, s nímž úzce souvisí i systém automatické morfologické disambiguace LanGr, kterému je věnována pozornost jak na teoretické úrovni tvorby pravidel, tak na úrovni jejich zápisu ve stejnojmenném programovacím jazyce.

Praktická část práce se zaměřuje na konkrétní výsek morfologické homonymie akuzativních tvarů a zlepšuje morfologickou anotaci na základě nově vytvořených pravidel pro systém LanGr. Nejprve v první případové studii na datech z korpusu SYNv10 identifikuje pravidelnosti APKS s bezpředložkovým akuzativem, následně zjištění ověřuje v druhé případové studii. Výsledná kritéria pro vyhledávání APKS jsou přetavena v pravidla pro systém LanGr, jenž díky tomu nově dovede úspěšně disambiguovat konstrukce, které doposud nebyly pravidly pokryté. Součástí práce je i série skriptů v programovacím jazyce Python pro jednodušší anotaci korpusových dat a vyhodnocování výsledků.

Klíčová slova

analytické predikáty s kategoriálním slovesem, jazykový korpus, morfologická anotace, disambiguace, LanGr

Abstract

This Ph.D. thesis deals with light verb constructions (LVCs), such as *provádět kontrolu* (to perform a check) or *chovat úctu* (to show respect). It demonstrates how to apply theoretical knowledge of these constructions into practice, exploiting it during morphological disambiguation and thus potentially improving syntactic analysis.

The theoretical part of the thesis covers three areas: light verb constructions, corpus annotation, and LanGr rule-based morphological disambiguation tagging. At first, LVCs are characterized, including their identification criteria, followed by a description of the current state of research on LVCs and a summary of papers published on the topic over the last fifteen years, with a particular emphasis on the syntactic approach to these constructions. A compilation of existing LVC inventories is also provided. Furthermore, Czech National Corpus written corpora tagging process is outlined, as it is closely related to the LanGr tool. Finally, LanGr rule creation and code implementation are covered.

The practical part of the thesis addresses nominative-accusative case homonymy. New rules for the LanGr system are also developed to improve morphological annotation. In a case study, the most frequent forms of direct object LVCs are retrieved using data from the SYNv10 corpus. Subsequently, the findings are verified in a second case study. The resulting LVCs retrieval criteria are translated into rules for the LanGr system, which is, from now on, able to successfully disambiguate constructions not previously covered. The thesis concludes with a series of Python scripts for corpus data annotation and result evaluation.

Key words

light verb constructions, language corpus, morphological annotation, disambiguation, LanGr

Obsah

Poděkování	i
Prohlášení	ii
Abstrakt	iii
Klíčová slova.....	iii
Abstract	iv
Key words	iv
1 Úvod	1
1.1 Motivace a přínos práce.....	1
1.2 Cíl práce.....	2
1.3 Struktura práce.....	2
2 Analytický predikát s kategoriálním slovesem	5
2.1 K definici APKS.....	6
2.1.1 Rozdílná pojmenování APKS	6
2.1.2 Kritéria a testy pro určení APKS	7
2.1.2.1 Redukční test	8
2.1.2.2 Test koreference argumentů	11
2.1.2.3 Další pomocná kritéria	13
2.1.2.4 Kritéria pro anotaci APKS v PDT	14
2.2 Problematika APKS z pohledu lingvistických disciplín a rozcestník odborné literatury	15
2.2.1 Stylistická a komunikační rovina.....	18
2.2.2 Lexikologie, lexikografie a frazeologie	20
2.2.2.1 APKS jako ustálená jednotka	21
2.2.2.2 Původ a tvorba APKS.....	22
2.2.2.3 Sémantika spojení.....	23
2.2.3 Syntax	24
2.2.3.1 Dvourovinná valenční syntax	24
2.2.3.2 Lexikogramatická škola	25
2.2.3.3 Funkční generativní popis	26
2.2.3.3.1 Značení APKS v PDT.....	26
2.2.3.3.2 Příspěvky k APKS (zejm. po roce 2010).....	29
2.2.4 Diachronní výzkum.....	34
2.3 Predikativní jméno.....	35

2.3.1	Deverbativa	37
2.3.2	Valence	39
2.3.3	Rozvíjení predikativního jména	41
2.4	Kategoriální sloveso	42
2.4.1	Variabilita kategoriálních sloves	43
2.4.2	Typologie	44
2.4.3	Desémantizace	46
2.4.4	Valence	49
2.4.4.1	Koreference a mapování valenčních struktur	49
2.4.4.2	Sémantická vs. syntaktická řízenost APKS	51
2.4.4.3	Pozice predikativního jména a jeho pád	52
2.4.4.4	Třetí valenční doplnění	53
2.5	APKS jako analytický predikát	55
2.6	Seznamy českých APKS	58
2.6.1	Přílohy ke studiím	58
2.6.2	VALLEX	58
2.6.3	ParaDi	60
3	Značkování korpusů	61
3.1	Tokenizace, větná segmentace a morfologická analýza	62
3.2	Přidání verbtagů	65
3.3	Morfologická disambiguace	65
3.3.1	Statistické (stochastické) metody	67
3.3.2	Pravidlové metody	68
3.3.3	Hybridní metody	70
3.4	Syntaktická anotace	71
3.5	Dodatečné úpravy a finalizace	71
4	System LanGr	72
4.1	Pravidla	72
4.2	Průběh disambiguace	74
4.3	Implementace pravidel LanGr	74
4.3.1	Negativní bigram	76
4.3.2	Sjednocení sekvence s omezením	77
5	Praktická část	78
5.1	Motivace a předpoklady	78
5.2	Data a jejich zpracování	80

5.3	APKS s predikativním jménem v akuzativu.....	82
5.3.1	Případová studie: <i>chovat/chovávat</i>	82
5.3.1.1	Získání seznamu potenciálně predikativních jmen	83
5.3.1.2	Anotace vět obsahujících hledané sloveso i podstatné jméno.....	85
5.3.1.3	Analýza dat.....	86
5.3.1.3.1	Obecná charakteristika dat.....	87
5.3.1.3.2	Kolokace	92
5.3.1.3.3	Struktury spojení.....	95
5.3.1.3.4	Slovesný způsob	102
5.3.1.3.5	Další pozorování	103
5.3.1.4	Stručné shrnutí.....	103
5.3.2	Druhá případová studie: ověření nalezených struktur na dalších APKS s bezpředložkovým akuzativem	104
5.3.2.1	Získání dat a jejich anotace	106
5.3.2.2	Analýza dat.....	106
5.3.3	Tvorba pravidel pro morfologickou disambiguaci LanGr	109
5.3.3.1	K míře morfologické homonymie predikativních jmen	109
5.3.3.2	Zápis pravidel	111
5.4	Používané skripty	113
5.4.1	Skripty použité v první studii.....	114
5.4.2	Skripty použité v druhé studii	115
5.4.3	Skripty pro analýzu morfologické homonymie	115
5.4.4	Skript pro konverzi dat z Excelu do CSV.....	116
6	Závěr	117
	Zdroje.....	120
	Použité nástroje a korpusy	120
	Literatura.....	121
	Seznam zkratk	129
	Zdroje a autoři příkladů.....	130
	Seznam tabulek	131
	Seznam grafů, obrázků a zdrojových kódů	133
	Seznam grafů.....	133
	Seznam obrázků	133
	Seznam zdrojových kódů	133

Příloha 1	Soupis příkazů, operátorů a komentářů používaných v programovacím jazyce LanGr	134
Příloha 2	Počet slov (vzdálenost) mezi kategoriálním slovesem a predikativním jménem v APKS s <i>chovat/chovávat</i>.....	137
Příloha 3	Predikativní jména tvořící APKS se slovesem <i>chovat</i> v pořadí V-N.....	138
Příloha 4	Predikativní jména tvořící APKS se slovesem <i>chovat</i> v pořadí N-V.....	140
Příloha 5	Číslo predikativních jmen (APKS s kategoriálním slovesem <i>chovat</i>).....	142
Příloha 6	APKS s typem koreference <i>ACT_{subst}-ACT_{verb}</i>	144
Příloha 7	Přehled homonymních tvarů (Asg a Apl) predikativních jmen z Přílohy 6.....	147

1 Úvod

Tato disertační práce se věnuje analytickému predikátu s kategoriálním slovesem (dále APKS), tedy spojením jako *provádět kontrolu* nebo *chovat úctu*. Zaměřuje se na jejich využití pro systém automatické morfologické disambiguace LanGr vyvíjený na Ústavu teoretické a počítační lingvistiky Filozofické fakulty Univerzity Karlovy (dále ÚTKL) a používaný pro morfologickou anotaci jazykových korpusů Českého národního korpusu (dále ČNK).

1.1 Motivace a přínos práce

Většina moderních lingvistických prací využívá jako zdroj jazykových dat nejrůznější korpusy. Důvod je zřejmý: jsou velmi rozsáhlé a oplývají užitečnými údaji (morfologickými značkami, lemmaty, popisem syntaktické struktury aj.), které mohou lingvisté pro svůj výzkum využít. Úspěšnost morfologického značkování korpusů dnes dosahuje 97,32 % (Jelínek et al. 2021, s. 55), přesto existují oblasti, které lze vylepšit.

Jednou z nich je anotace tvarů substantiv s pádovým synkretismem, při níž místy dochází k nesprávnému určení pádu. V případě špatného rozlišení mezi homonymním nominativem a akuzativem se přitom chyba projeví i u syntaktického značkování podmětu a předmětu. Problematiku příhodně ilustrují věty *Soudce vynesl rozsudek* a *Rozsudek vynesl soudce*. Jelikož je čeština jazyk s volným slovosledem a věty samotné neobsahují žádný kontext, lze je jen velmi těžko počítačově analyzovat. Mluvčí jazyka si s touto situací poradí a rozhodne se na základě znalosti sémantiky daného spojení, což zatím počítačové nástroje nedovedou. Místo toho se při rozhodování opírají buď o strojové učení, nebo o pravidla formulovaná lingvisty. Zatímco strojové učení můžeme zlepšit jedině tím, že mu poskytneme kvalitnější a rozsáhlejší trénovací data, pravidla lze upravit, případně vytvořit nová. Právě to je jedním z cílů této práce.

APKS, ve smyslu konstrukcí složených z vyprázdněného slovesa a ze substantiva nesuscího význam celého analytického spojení, jsou velmi produktivním jazykovým fenoménem, který je vlastní různým jazykům a v češtině cítá nižší tisíce spojení. Na první pohled se APKS chovají a vypadají jako běžná spojení slovesa a předmětu nebo podmětu, zároveň je ale charakterizují vlastnosti, které běžná spojení predikativního slovesa a substantiva nemají.

Přestože se při zpracovávání korpusů s víceslovnými jednotkami běžně pracuje, APKS nejsou zatím dostatečně pokryty. Domnívám se, že studiem APKS a přetavením poznatků do pravidel bude možné ambiguitu způsobenou pádovým synkretismem alespoň částečně řešit.

Tato úvaha se zakládá na předpokladu, že známe-li dvojici substantiva a slovesa, která společně do APKS vstupuje, víme, v jakém pádě se substantivum v dané konstrukci nachází. Rozpoznáme-li tedy APKS v příkladových větách, můžeme rozhodnout, že substantivum *rozsudek* z věty *Soudce vynesl rozsudek* je v akuzativu a substantivum *soudce* v nominativu.

1.2 Cíl práce

Předkládaná disertační práce není tradičním teoretickým příspěvkem k tématu APKS, ale klade si za cíl teoretické poznatky aplikovat na zpracování dalších jazykových dat.

Samotné analýze APKS a jejich teoretickému poznání se v posledních patnácti letech věnovala řada lingvistů (zejm. Kettnerová 2017; Kettnerová a Lopatková 2017; Martínek 2008b; 2016; Radimský 2010) a nemalá část příspěvků vyšla během psaní této disertační práce. Co ale stále chybí, je syntéza existujícího poznání; starší práce jsou jen velmi rámcově shrnuty v Radimského (2010) monografii. Právě shromáždění a utřídění existujících poznatků je jedním z cílů této práce. Podrobněji se zaměřuji na poznání související s morfologickou a syntaktickou stránkou těchto spojení, kde spatřuji největší přínos pro korpusovou a počítačovou lingvistiku.

S rozvojem korpusů úzce souvisí proces jejich tvorby, mj. morfologická anotace korpusových dat, jejíž součástí je též systém pravidly řízené redukční automatické morfologické disambiguace LanGr. Ten tvoří nedílnou součást jak procesu značkování korpusů ČNK, tak dalších nástrojů pro počítačové zpracování češtiny.

Se systémem LanGr úzce souvisí praktický a zároveň hlavní cíl práce: nalézt takové pravidelnosti APKS, na jejichž základě bude možné formulovat pravidla pro zmíněnou morfologickou disambiguaci, s jejich pomocí zlepšit úspěšnost morfologické anotace a v návaznosti na ni nepřímo ovlivnit i anotaci syntaktickou. Předpokládám, že nabyté poznatky bude možné zúročit nejen ve vytvořených pravidlech, ale i dále – například v databázi víceslovných lexikálních jednotek LEMUR (Petkevič et al. 2020).

1.3 Struktura práce

Disertační práce sestává ze čtyř kapitol orámovaných úvodem (kapitola 1) a závěrem (kapitola 6), na konci doplněných o používané zdroje a seznam zkratk, tabulek, grafů, obrázků a zdrojových kódů používaných v textu práce. Méně rozsáhlé přílohy jsou přímo součástí práce, rozsáhlejší jazykové zdroje a skripty jsou přiloženy zvlášť.

Do teoretické části práce spadají tři kapitoly, které připravují půdu pro její praktickou část a věnují se APKS, značkování korpusů obecně a disambiguačnímu systému LanGr.

První ze zmíněných tří kapitol pojednává o analytickému predikátu s kategoriálním slovesem (kapitola 2). Ve svém úvodu sjednocuje různorodou terminologii a představuje kritéria a testy pro rozlišování mezi APKS a volným spojením predikativního slovesa se substantivem (2.1). Dále se věnuje stávajícímu výzkumu APKS z pohledu jednotlivých lingvistických disciplín (2.2): nahlíží na stylistickou a komunikační rovinu těchto spojení, představí přístup lexikologický, lexikografický i frazeologický, podrobněji se věnuje syntaktickému poznání a okrajově diachronnímu výzkumu. V dalších podkapitolách se zabývá vlastnostmi predikativního jména (2.3), kategoriálního slovesa (2.4) a APKS jako celku (2.5) z pohledu morfologicko-syntaktického (valence, typologie spojení, desémantizace kategoriálního slovesa aj.). Kapitulu uzavírá rozcestník existujících seznamů APKS (2.6), které je možné využít při tvorbě pravidel.

Následuje kapitola 3 věnovaná značkování jazykových korpusů, která se zaměřuje na celý proces morfologické anotace. Soustředí se zejména na morfologickou disambiguaci (3.3), tj. výběr správné morfologické značky a lemmatu pro každé slovo (token) z nabídky všech tvarově možných.

Do oblasti morfologické disambiguace spadá i poslední kapitola teoretické části (kapitola 4), která se zabývá pravidlovou disambiguací LanGr. Kapitola představuje fungování systému, včetně teoretické (4.1) i praktické části tvorby pravidel (4.2), a zabíhá též do techničtějších podrobností (4.3), jako je stejnojmenný programovací jazyk, ve kterém jsou pravidla zapísána a který bude potřeba v praktické části.

Praktická část (kapitola 5) poznatky z teoretických kapitol využívá k vytvoření pravidel pro morfologickou disambiguaci LanGr. V úvodu přibližuje, jaké možnosti se při řešení problematiky morfologické homonymie pomocí znalosti APKS nabízejí (5.1), dále popisuje způsob práce s daty z korpusu SYNv10 (5.2). Následují dvě případové studie zaměřené na APKS s bezpředložkovým akuzativem jakožto nejrozšířenějším typem těchto spojení.

První ze studií (5.3.1) podrobně zkoumá spojení kategoriálního slovesa *chovat/chovávat* a rozšířeného seznamu 47 predikativních jmen. Samotná deskriptivní analýza se soustředí zejména na povrchovou strukturu nalezených APKS, konkrétně na sekvence slovních druhů, které se mezi oběma členy APKS nacházejí, přičemž hledá formální rysy, o které by bylo možné se opřít při tvorbě pravidel.

V rámci druhé případové (5.3.2) studie jsou vysledované pravidelnosti shrnuty do kritérií, podle nichž je možné vyhledávat APKS v korpusu: za předpokladu, že známe dvojice

kategoriálního jména a predikativního slovesa, by výsledky dotazu díky formulovaným kritériím měly obsahovat téměř výhradně APKS. Kritéria jsou následně úspěšně ověřena na rozsáhlejší seznamu APKS a využita pro tvorbu nových pravidel pro morfologickou disambiguaci LanGr (5.3.3). Praktickou část uzavírá přehled a stručná charakteristika používaných skriptů (5.4).

2 Analytický predikát s kategoriálním slovesem

Zatímco u syntetických predikátů je význam vyjádřený jen tvarem slovesa (*kontrolovat, plavat*), u analytických predikátů s kategoriálním slovesem (dále APKS) je rozdělený mezi dvě části – sloveso a nejčastěji substantivum¹ (*provádět kontrolu, chovat sympatie, tropit hlouposti*). Sloveso tvořící formálně-syntaktickou kostru věty má funkci kategoriální (označuje typ události, ale jinak je významově vyprázdněné); v češtině vyjadřuje pouze způsob, čas a vid (syntakticky nezávislé kategorie) a k nim osobu a číslo (syntakticky závislé shodové kategorie), zatímco samotné sémantické jádro je vyjádřeno obvykle abstraktním substantivem (které může také vyjadřovat vid). Kromě toho má APKS další specifické vlastnosti, kterými se odlišuje od jiných typů spojení (srov. např. Daneš, Grepl a Hlavsa 1987, s. 23; Radimský 2017a; 2017b).

Tyto konstrukce jsou velmi produktivní a zcela běžné nejen v moderních indoevropských jazycích, ale i v jazycích jim zcela vzdálených, jako je např. korejština nebo čínština (Butt 2003; 2010). Na evropských jazycích např. můžeme sledovat, jak se původně latinský výraz *promissum tenere* přenesl do národních jazyků – *držet slib, keep the promise, sein Versprechen halten, tenir sa promesse, držat' obeščanije* (Jelínek 2007, s. 45–46).

Po stopách APKS se můžeme vydat až k historickým počátkům indoárijské jazykové skupiny – z dnešní hindštiny nebo urdštiny až tři tisíce let zpět v čase do sanskrtu, v jehož gerundiu má kořeny i dnešní slovesná složka APKS. Jedná se tak o historicky velmi stabilní fenomén (Butt 2003; 2010). Přestože se ve starší bohemistické literatuře hojně vyskytoval názor, že v češtině jsou APKS typické až pro moderní a funkční jazyk, Martínek (2008a; 2016) spolehlivě zdokumentoval užívání APKS i v češtině sklonku 15. století (k utváření a původu APKS viz 2.2.2.2 a 2.2.4).

Tato kapitola si klade za cíl shrnout předchozí poznatky nikoliv primárně po jednotlivých autorech,² ale spíše uceleně po oblastech výzkumu; více současné bohemistické poznatky provázat a vytvořit rozcestník, který mně samotné zpočátku chyběl. Mým záměrem není zaplňovat prázdná místa, ale naopak poznatky přetavit ve zlepšení anotace korpusů, a zároveň ověřit

¹ Kralčák (2005, s. 34, 60–61) zabývající se slovenštinou řadí mezi APKS i spojení s adverbii (*robiť napriek, vyjst' najavo*). Jako součást periferních APKS je analyzuje ve staré češtině i Martínek (2008b, s. 91–93), zejména kvůli „větší variantnosti a rozkolísanosti normy“ (Martínek 2008b, s. 18). Oba autoři přitom poněkud paradoxně pracují s pojmem *verbonominální*, který taková spojení nezahrnuje (za podnětnou poznámku děkuji Vladimíru Petkevičovi). Také další autoři zmiňují spojení s adjektivy (*být podobný komu*) nebo adverbii (*dát dolů*) (Kettnerová a Lopatková 2017, s. 5; Macháčková 1979, s. 51; Radimský 2017a), nicméně v analýzách se jim nevěnují. I tato disertační práce se zaměřuje pouze na spojení se substantivou, která tvoří centrum APKS.

² Skvělá shrnutí starší bohemistické literatury viz např. Radimský (2010, s. 25–35) nebo Martínek (2008b, s. 18–24), ve slovenské literatuře příp. Kralčák (2005, s. 17–23).

některá předchozí pozorování. V roce 2017 vyšel *CzechEncy – Nový encyklopedický slovník češtiny* (viz Radimský 2017a; 2017b). V něm uvedená hesla věnovaná APKS jsou bezpochyby vynikajícím úvodem do problematiky, který si nicméně zaslouží dále rozvinout. Právě formát disertační práce poskytuje pro ucelenější urovnání dosud publikovaných poznatků ideální prostor – přestože místy za cenu větší či menší generalizace a/nebo pouhého povrchního porovnání jednotlivých přístupů. Podrobněji se zaměřuji na oblasti související s dalšími částmi disertační práce a na ty, které nebyly zpracovány nebo shrnuty jinde (tj. literaturu po roce 2010), případně na ty, u kterých je komentář nezbytně nutný.

V této kapitole nejprve vymezím terminologii APKS (2.1.1) a představím kritéria a testy pro určování APKS (2.1.2). Shrnu základní literaturu pohledem na problematiku APKS pomocí rozdělení na lingvistické disciplíny (2.2). V dalších kapitolách věnovaných predikativnímu jménu (2.3), kategoriálnímu slovesu (2.4) a APKS jako celku (2.4.4) se zaměřím především na syntaktické hledisko. Kapitulu uzavřu přehledem existujících seznamů českých APKS (2.6).

2.1 K definici APKS

Tato podkapitola má čtenáři usouvztažnit označení APKS s dalšími používanými označeními téhož jevu a také hned zkraje představit testy a kritéria pro určování těchto spojení.

2.1.1 Rozdílná pojmenování APKS

Odborných termínů označujících celou strukturu a její jednotlivé složky existuje celá řada. V této práci jsem se rozhodla nevytvářet novou terminologii a užívat termín *analytický predikát s kategoriálním slovesem* (APKS), pro slovesnou složku termín *kategoriální predikát* a pro jmennou složku termín *predikativní jméno*. Termín APKS pochází ze stejnojmenného hesla v *CzechEncy – Novém encyklopedickém slovníku češtiny* (Radimský 2017a). Radimský zde ustoupil od termínu *verbonominální predikát s kategoriálním slovesem*, který používal mj. ve své monografii (Radimský 2010) v návaznosti na termíny zavedené v *Mluvnici češtiny 3* (Daneš, Grepl a Hlavsa 1987, s. 23–24). Tento „nový“ termín klade důraz na obecnou dvou-složkovost spojení, zatímco termín předchozí odkazoval na uskupení sestávající ze slovesa a substantiva. Taková spojení jistě tvoří většinu APKS, nicméně kategoriální sloveso může figurovat i ve spojení s adverbii a adjektivy. Při práci s dalšími zdroji se snažím pojmenování sjednotit, aby byla pro čtenáře práce co nejkonzistentnější a nejpřehlednější.

V českých studiích se kromě zmíněných termínů setkáme i s dalšími označeními, jako *analytický predikát* nebo *slovesně jmenná spojení* (Macháčková 1979; 1983), *analytická*

verbonominální spojení (Martínek 2008a; 2009; 2010; 2011), *verbonominální spojení* (Gregor 2008; Jelínek 2002; 2003; 2007), *multiverbizační/multiverbální spojení* (Vlková 1978; 1990), *složený predikát* (Kolářová 2010; Mikulová et al. 2005), *komplexní predikát* (Kettnerová 2017; Kettnerová a Lopatková 2017) aj. Další existující označení, např. Čermákovo (1974) *víceslovné pojmenování typu verbum-substantivum*, na tomto místě neuvádím, jelikož nejsou dále v literatuře tematizovaná a označují oblast výrazně obecnější. Pro slovesnou složku se v českém prostředí ujal již zmiňovaný termín *kategoriální sloveso* (většina příspěvků k tématu), na základě německého *Funktionsverb* též *funkční sloveso* (Kettnerová 2017; Macháčková 1979; 1983), okrajově také *operační sloveso* (Jelínek 2002; 2003). Jmenná složka se obvykle označuje jako *predikativní jméno*, v některých případech jako *verbonominální* nebo *analytický predikát*.

I v angličtině je terminologie pestrá: *light verb constructions* (Jespersen 1942), *composite predicate* (Cattell 1984), *complex predicate/verb* (Brinton a Akimoto 1999; Butt 2003; 2010), *complex predicate with light verb*, *support verb constructions*, *expanded predicates*, *verbo-nominal phrases* aj. Pro kategoriální sloveso se nejčastěji užívají termíny *light verb* nebo *support verb* a pro predikativní jméno prostý termín *noun*, *predicate noun* nebo *predicative noun*. Ve francouzštině se nejčastěji používá termín *prédicats verbonominaux à verbe support* (verbonominální predikáty s podpurným slovesem), příp. *prédicats nominaux à verbe support* (jmenné predikáty s podpurným slovesem), v německých studiích termín *Funktionsverbgefüge* (struktury s funkčním slovesem). Mnohé termíny (*light verbs*, *verbes support* i *Funktionsverben*) se pojí s konkrétními metodologickými přístupy (Radimský 2010, s. 11).

Seznam uvedený v této podkapitole zdaleka není vyčerpávající. Pro podrobnější výčet viz shrnutí J. Gregora (2008, s. 238–247), který srovnává terminologii napříč mnoha jazyky (s důrazem na ruštinu a češtinu).

2.1.2 Kritéria a testy pro určení APKS

Přestože jsou podle M. Butt³ (2003; 2010) APKS univerzálním lingvistickým fenoménem, mezi jednotlivými jazyky (příp. jazykovými rodinami obecně) se mohou struktury APKS i výrazně lišit. Proto jsou veškeré syntaktické testy vymezující APKS a kategoriální predikáty specifické pro daný jazyk (Butt 2003, s. 24; Cinková 2009, s. 27) a testy fungující v cizím jazyce je nutno nejprve ověřit a případně adaptovat. Podívejme se nyní na to, jak se definují APKS v češtině.⁴

³ Příjmení zahraničních autorek ponechávám záměrně v původní nepřechýlené formě, kromě osobní preference spatřuji hlavní výhodu v jednodušší práci s citovanou literaturou.

⁴ Postupný vývoj starších kritérií detailně popisuje např. Kralčák (2005, s. 27–29), z něhož vychází i Martínek (2008b).

Starší bohemistické práce (zejm. práce stylistické a lexikologické) APKS často vymezovaly poměrně intuitivně na základě neformální charakteristiky. Častým kritériem byla existence synonymního syntetického slovesa (*dát podnět – podnítit*), ze kterého bylo utvořeno multiverbizační spojení. Multiverbizační teorie vzniku APKS se dále rozvíjela a akceptovala spojení vytvářená analogicky (*dát přednost* na základě *dát podnět*), případně se přestala omezovat na stejný základ obou sloves (*chovat úctu – respektovat, dělat rozdíl – rozlišovat*). Objevují se i další kritéria: sloveso je sémanticky vyprázdněné, APKS tvoří jeden významový celek (jako verbonominální spojení se sponou), predikativní jméno je vyjádřené abstraktem atp. (srov. např. Čermák 1974; Jelínek 2002; 2003; 2007; Macháčková 1979; 1983; Vlková 1978; 1990).

Novější studie akcentují zejména skutečnost, že významově oslabené sloveso a substantivum společně vytvářejí jednu jednotku se společnou/transponovanou syntakticko-sémantickou strukturou. Kategoriální sloveso poskytne valenční rámec, do kterého predikativní jméno přinese sémantické participanty – ty jsou vyjádřeny právě jako valenční doplnění kategoriálního slovesa (srov. Cinková 2009, s. 50; Kettnerová 2017; Kettnerová a Lopatková 2017; Kolářová 2010, s. 112; Kralčák 2005, s. 27–29; Macháčková 1983, s. 135; Radimský 2010; 2017a). V angličtině se tento princip nazývá různě: *argument fusion*, *argument transfer*, *argument composition* (srov. např. Kettnerová a Lopatková 2015, s. 192; Kettnerová, Lopatková a Barančíková 2016, s. 882), v češtině bývá tematizován jako koreference/sdílení/slučování argumentů.

V následujících oddílech představím transformační testy, které vycházejí z tradice francouzské Lexikogramatické školy a jsou adaptované pro češtinu. Konkrétně tu jde o redukční test a test koreference argumentů, které dovolí výše zmíněné poznatky formálně ověřit. Dále shrnu kritéria používaná pro anotaci dat Pražského závislostního korpusu (*Prague Dependency Treebank*, dále PDT) a navrhu ke zvážení několik doplňujících kritérií dalších autorů.

2.1.2.1 Redukční test

Redukční test (*cancellation/reduction test*, fr. *test de l'effacement du verbe support*) formalizuje intuitivní rozhodování, zda je sloveso sémanticky vyprázdněné. Zakládá se na poznatku, že sémantické jádro APKS je reprezentováno predikativním jménem (nikoliv kategoriálním slovesem), tudíž predikativní jméno na rozdíl od kategoriálního slovesa nemůže být z věty vypuštěno bez ztráty významu. Redukční test ověřuje, že kategoriální sloveso do APKS nevnaší vlastní lexikální význam a že je přítomno intencí pole predikativního jména. Při redukčním testu jsou z testované věty postupně odstraňována sysémantika,⁵ a to pomocí syntaktické

⁵ Kettnerová a Lopatková (2017, s. 4) poznamenávají, že „[z]e syntaktického hlediska je toto hodnocení do určité míry problematické: slovní druhy považované tradičně za sysémantika (předložky, spojky a částice) neplní úlohu větných členů, zatímco kategoriální slovesa ano“.

operace relativizace a vynechání kategoriálního slovesa (srov. Kettnerová et al. 2013, s. V–VI; Radimský 2010, s. 14–16, 54–65; 2012, s. 68–72; 2017a).

Princip testu je nejvhodnější rovnou dokumentovat na příkladu – vezměme dvě rozdílné věty z řádku (a) v Tabulce 1 níže a aplikujme na ně redukční test. Prvním krokem je relativizace původní fráze (b), kdy predikativní jméno přesuneme do nominativu, ve druhém kroku vynecháme kategoriální sloveso (c) a v posledním kroku (d) ještě provedeme transformaci subjektového argumentu na přídavné jméno přivlastňovací. Poslední krok do francouzského lexikogramatického testu přidal Radimský přímo pro češtinu, aby bylo výsledné syntagma přijatelnější. Pokud během všech kroků transformace zůstává zachován původní význam věty (tzv. *sémantický invariant*),⁶ sloveso ve větě lze charakterizovat jako kategoriální (viz levý sloupec tabulky níže). V opačném případě, tzn. je-li ve větě použito plnovýznamové sloveso, se v průběhu transformace význam od původního značně vzdálí – (c) a (d) nejsou ekvivalenty (a) a (b). Příklady v tabulce jsou převzaty ze slovníkového hesla definujícího redukční test (Radimský 2017a):

	kategoriální sloveso – nést	predikativní sloveso – odmítat
(a)	Petr nese odpovědnost za tuto tragédii.	Petr odmítá odpovědnost za tuto tragédii.
(b)	Odpovědnost, kterou Petr nese za tuto tragédii.	Odpovědnost, kterou Petr odmítá za tuto tragédii.
(c)	Odpovědnost Petra za tuto tragédii.	Odpovědnost Petra za tuto tragédii.
(d)	Petrova odpovědnost za tuto tragédii.	Petrova odpovědnost za tuto tragédii.

Tabulka 1: Ukázka redukčního testu u spojení s kategoriálním a predikativním slovesem
Zdroj: Radimský (2017a)

Zachování sémantického invariantu si můžeme ověřit i přidáním slovesa *být* mezi výsledek transformace (d) a relativizovanou větu (b) (srov. příklady (1) a (2) níže). Pokud výsledné tvrzení platí, je invariant zachován.

- (1) Petrova odpovědnost za tuto tragédii **je** odpovědnost, kterou Petr nese za tuto tragédii. (JR)
- (2) Petrova odpovědnost za tuto tragédii ***je** odpovědnost, kterou Petr odmítá za tuto tragédii. (JR)

⁶ Více k tomuto strukturalistickému konceptu „neměnné referenční sémantiky, kterou sdílejí jednotlivé transformace určité sekvence“ (Radimský 2010, s. 16) viz Ibrahim (2007; cit. dle Radimského).

Radimský redukční test pro češtinu dále modifikuje a vytváří vlastní *test vložení kategoriálního slovesa*, který považuje za přesvědčivější, neboť se hledá sémantická ekvivalence mezi dvěma větami (nikoliv syntagmatem a větou jako u redukčního testu). V tomto testu do věty s kontextualizovaným predikativním jménem (3) naopak vložíme vztažnou větu s testovaným slovesem (4), tj. de facto provedeme redukční test od konce. Shoda významu vynikne zejména v kontrastu s případem, kdy vložené sloveso kategoriální není (5).

(3) Petrovu odpovědnost za tuto tragédii si všichni uvědomují. (JR)

(4) Odpovědnost, **kte**rou Petr za tuto tragédii **nese/má**, si všichni uvědomují. (JR)

(5) Odpovědnost, **kte**rou Petr **odmítá**, si všichni uvědomují. (JR)

U tohoto testu je třeba dbát na to, aby kontextualizace (*si všichni uvědomují*) nebyla v rozporu s modalitou, případně temporálností, kterou nese vkládané kategoriální sloveso. Pokud je *odpovědnost* z příkladu níže předmětem *šetření*, tak je slovesem *má* zpochybněno probíhající *šetření*, a proto je souvětí obtížně přijatelné.

(6) [?]Odpovědnost, kterou Petr má za tuto tragédii, je předmětem šetření. (JR)

Ke zpřesnění testu vložení kategoriálního slovesa a odstranění případných pochybností ještě Radimský přidává doplňkový *test s vložení dvou kategoriálních sloves* (neboli *diferenční test*), při kterém se do výchozí věty (7) s kategoriálním slovesem (*mít odpovědnost*) vloží testovaná slovesa (*nést, vnímat, cítit*) jako v příkladech (8) a (9). Je-li vyznění obou výsledných klauzí redundantní jako v případě (8), mají slovesa výlučně kategoriální význam. Kombinace nevyprázdněného slovesa a kategoriálního slovesa v druhé větě (9) tautologická není.

(7) Petr **má** odpovědnost za tuto tragédii. (JR)

(8) *Petr **nese** odpovědnost, kterou **má** za tuto tragédii. (JR)

(9) Petr **vnímá/cítí** odpovědnost, kterou **má** za tuto tragédii. (JR)

Přestože se původně jedná o binární test ověřující APKS zasazený do jednoduché typové věty, Radimský ve své práci ukazuje, že jednotlivé konstrukce zasazené do skutečného kontextu *parole* mohou mít různou míru přijatelnosti a bylo by lepší k testu přistoupit jako k testu skalárnímu. Více ke skalární desémantizaci kategoriálního slovesa v oddíle 2.4.2.

2.1.2.2 Test koreference argumentů

Na APKS je dále možné v některých případech aplikovat test koreference argumentů⁷ (*coreference test, test de coréférence de sujet*) (srov. Kettnerová et al. 2013, s. VI; Radimský 2010, s. 54–55, 72–82, 166–167; 2012, s. 73–76; Radimský 2017a), který ověří, zda je valenční pole kategoriálního slovesa shodné s intenčním polem predikativního jména. Test se zakládá na skutečnosti, že jeden ze sémantických argumentů (např. levovalenční doplnění *Petr* v predikátu *odpovědnost*) může mít tři rozdílné syntaktické realizace, a to na pozici jmenného komplementu (10) a (11) a v pozici subjektu kategoriálního slovesa (12), a zároveň je stále zachována koreference (u predikativních sloves být zachována nemusí).

(10) odpovědnost **Petra** za tuto tragédii (JR)

(11) **Petrova/jeho** odpovědnost za tuto tragédii (JR)

(12) **Petr** nese odpovědnost za tuto tragédii. (JR)

Radimský si dále všímá, že při vyjádření stejného argumentu vložím posesivního výrazu (13) je význam stále interpretovatelný, přestože se jedná o stylisticky neobratné vyjádření. Pokud je koreference porušena (14), věta s kategoriálním slovesem již interpretovatelná není. U věty s plnovýznamovým slovesem (15) žádné podobné omezení neplatí.

(13) **Petr** nese [?]**svou** odpovědnost za tuto tragédii. (JR)

(14) **Petr** nese ^{*}**mou/Evinu** odpovědnost za tuto tragédii. (JR)

(15) **Petr** zpochybňuje **mou/svou/Evinu** odpovědnost za tuto tragédii. (JR)

Tento test lze aplikovat pouze pro konstrukce s APKS základního typu (k typologii více v oddílu 2.4.2), neboť se v nich první syntaktický argument predikativního jména realizuje na pozici subjektu kategoriálního slovesa. U nenominativního typu podmět není realizován vůbec (*došlo k zemětřesení*), případně je obsazen predikativním jménem (*panuje nejistota*). Testování koreference není možné ani u kauzativních APKS, protože argument v podmětu kategoriálního slovesa není v intenční struktuře predikativního jména (*opatření vyvolalo diskusi – diskuse *opatřením / *z opatření*). U konverzního typu APKS je možnost vyjádřit jeden argument na dvou pozicích současně syntakticky blokována a nelze ji realizovat (*Lukáš dostal facku od Evy – Lukáš dostal facku *Lukášovi*), v případě pravovalenčních doplnění nedává testování koreference smysl, i když je teoreticky možné (*Lukáš dostal tu Evinu facku = Tu facku od Evy dostal Lukáš* × neinterpretovatelné **Lukáš dostal tu Evinu facku od Moniky*), protože jsou syntakticky

⁷ V lexikogramatické terminologii test koreference subjektů (myšleno sémantických subjektů).

navázaná na predikativní jméno a kategoriální/predikativní sloveso je neovlivňuje (*Evina facka bolela = facka od Evy bolela* × neinterpretovatelné **Evina facka od Moniky bolela*) (Radimský 2017a).

Radimský (2010, s. 122) upozorňuje, že u konstrukcí s rezultativy lze koreferenci porušit přidáním přívlasktu neshodného – v příkladu (16) se jedná o požadavek starosty a konstrukci klasifikujeme jako APKS, v příkladu (17) již jde o požadavek Technických služeb, čímž se konstrukce stane interpretovatelnou predikativně. Chybějící kategoriální sloveso můžeme přidat ve vedlejší větě (18). Více viz skalární desémantizace (2.4.2).

(16) Starosta **přednesl požadavek** o zvýšení cen za prodej hrobového místa.⁸ (JR)

(17) Starosta **přednesl požadavek Technických služeb** o zvýšení cen za prodej hrobového místa. (JR, ČNK)

(18) Starosta **přednesl požadavek** o zvýšení cen za prodej hrobového místa, **který vyslovily Technické služby**. (JR)

Test koreference argumentů používá i kolektiv autorů zabývajících se přidáváním APKS do slovníku VALLEX (viz 2.2.3.3.2). Ti koreferenční test přesouvají z povrchové roviny na rovinu tektogramatickou (významovou, hloubkovou) a vymezují APKS jen ze syntaktického hlediska.

V této úpravě test ověřuje koreferenci valenčních doplnění bez ohledu na to, jestli jsou vyjádřená na povrchové rovině, nebo nikoliv. Jediným požadavkem je, aby konatel predikativního jména koreferoval s libovolným doplněním kategoriálního slovesa, což umožňuje spolehlivě testovat APKS, v nichž je predikativní jméno v pozici přímého objektu. Zásadní výhodou je, že oproti testu koreference na povrchové rovině tato varianta umožňuje testovat i predikáty typu *dostat radu* či kauzativní APKS typu *vyvolat pobouření* (**Petr dostal radu Lukáše od Lukáše. *Zpráva vyvolala zděšení obyvatel u obyvatel.*). Generuje tak stejný výsledek pro aktivní i pasivní struktury typu *dát/dostat* (např. Kettnerová 2017, s. 7–8; Kettnerová a Lopatková 2017, s. 5–6). Pokud konatel predikativního jména nekoreferuje s žádným doplněním slovesa, nejedná se o APKS – stejně jako v příkladu (17) (Kettnerová 2017, s. 21–22).

„[Test koreference je také] omezeně aplikovatelný na komplexní predikáty, u nichž je predikativní jméno vyjádřené v subjektu (srov. např. *proběhl experiment*, kde ACT jména *experiment* nekoreferuje s žádným doplněním slovesa *proběhnout*[,] vs. *přepadl ho strach*, kde ACT jména *strach* koreferuje s PAT slovesa *přepadnout*), či na predikáty vytvářející

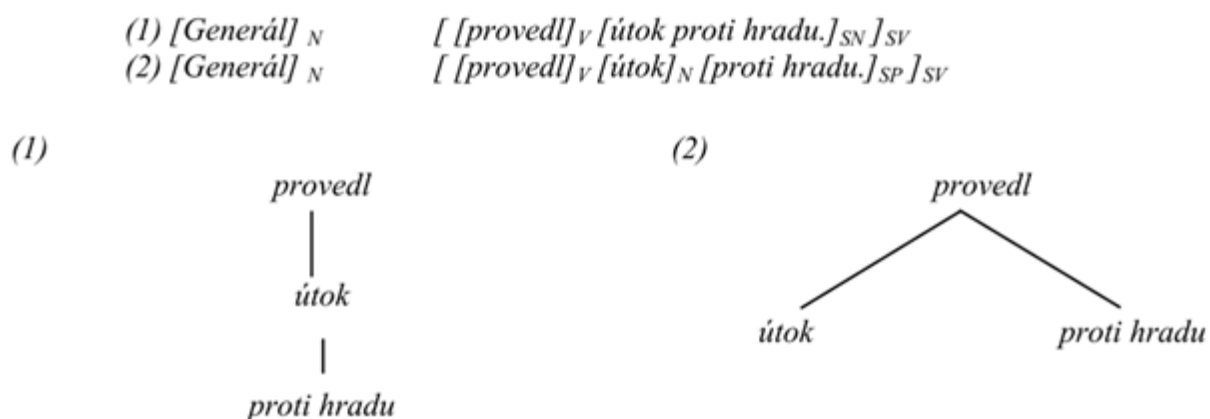
⁸ Vazba *požadavek o zvýšení cen* místo *požadavek na zvýšení cen* je již v příkladech, které Radimský (2010, s. 120) čerpá z ČNK.

bezsubjektové struktury (srov. např. *došlo k neštěstí*, kde ACT jména *neštěstí* nekoreferuje s žádným doplněním slovesa *dojít*)“ (Kettnerová a Lopatková 2017, s. 5).

2.1.2.3 Další pomocná kritéria

Kromě dvou výše zmíněných testů se lze při rozhodování opřít o další specifické vlastnosti APKS, což také mnozí autoři činí.

Lexikogramatikové využívali pro určení některých spojení tzv. dvojí vazbu (fr. *double analyse*) (Radimský 2010, s. 17–18, 166–168). Tato syntaktická anomálie umožňuje objektový argument syntakticky analyzovat dvojím způsobem (a spojení je tak mnohem volnější, než je běžné), viz schéma níže.



Obrázek 1: Schéma dvojí analýzy

Zdroj: Radimský (2010, s. 17)

Představit si to můžeme tak, že při relativizaci původní věty (19) můžeme doplnění *proti hradu* zavést jak na sloveso *provést* (20), tak na dějové jméno *útok* (21) (Radimský 2010, s. 17–18; 2012, s. 59–61).

(19) Generál provedl útok **proti hradu**. (JR)

(20) Útok, který generál **provedl proti hradu**. (JR)

(21) **Útok proti hradu**, který generál provedl. (JR)

Martínek (2008b, s. 27) mezi další kritéria řadí skutečnost, zda je spojení „součástí sítě několika významově souvisejících spojení“. Má tím na mysli to, že existuje řada kategoriálních sloves (včetně existujících sémantických opozicí) pojících se s jedním predikativním jménem a/nebo řada predikativních jmen pojících se s jedním kategoriálním slovesem.

Kolářová (2010, s. 114) v části věnované APKS z pohledu valence deverbativních substantiv navrhuje, že pokud sloveso použité jako kategoriální získá valenční vazbu (*udělat*

něčemu konec/přítrž zřejmě jako analogie k *zabránit čemu*), nebo ji naopak spojením s predikativním jménem ztratí (*podat výkon* – chybí adresát, *dostat chuť/nápad* – chybí origo), je možné celou konstrukci považovat za APKS, ačkoliv ostatní kritéria nejsou zcela splněna. Detailněji viz podkapitola 2.2.4 věnovaná kategoriálnímu slovesu.

Kettnerová et al. (2013, s. VI–VII)⁹ řadí ve své studii mezi kritéria též skutečnost, že u konstrukcí s APKS (22) může být omezená nebo méně přijatelná možnost nahrazovat predikativní jméno zájmenem (23) nebo se na něj zeptat pomocí tázacích zájmen (24). Tuto vlastnost využívá jako další z kritérií a ilustruje ji na následujících příkladech:

(22) Petr upadl do rozpaků. (VK et al.)

(23) *Petr upadl do toho. (VK et al.)

(24) *Do čeho Petr upadl? (VK et al.)

V této studii také autoři ověřovali úspěšnost určování APKS pomocí zmíněného redukčního testu, testu koreference argumentů¹⁰ a výše uvedeného doplňkového kritéria: shoda tří anotátorů na 5900 větách byla tzv. dobrá (85,3 %, Cohenovo vážené kappa 0,686) (Kettnerová et al. 2013, s. VII–IX).

2.1.2.4 Kritéria pro anotaci APKS v PDT¹¹

Při anotování PDT (více k APKS v PDT viz 2.2.3.3.1) byla pro tato spojení stanovena dvě obecná kritéria níže a k nim série vlastností určujících jednotlivé části (Mikulová et al. 2005, s. 329–331). Obecná kritéria jsou následující:

- (a) jedná se o kombinaci významově vyprázdněného slovesa a substantiva (často označujícího děj nebo stav) nesoucího hlavní lexikální význam celého spojení
- (b) APKS tvoří víceslovnou lexikální jednotku, která má ve většině případů odpovídající synonymní vyjádření pomocí jednoslovného predikátu (*hovořit* – *vést hovor*, *plánovat* – *mít plán*)

K tomu jsou doplněny tyto vlastnosti a bližší určení vlastností slovesné části (Mikulová et al. 2005, s. 331):

⁹ Přestože Kettnerová v tomto případě neuvádí žádný zdroj, totožně jev charakterizuje Macháčková (1979, s. 63–64) a v návaznosti na ni i Radimský (2010, s. 157–171; 2017a), více viz pozn. 55.

¹⁰ Oba testy byly provedené jen na povrchové rovině, k testu koreference argumentů na tektogramatické rovině dochází až v pozdějších studiích (viz 2.2.3.3.2).

¹¹ Pravidla a kritéria jsou ve většině případů kondenzovanou citací z anotátorské příručky a uvedených pravidel u každé sekce (domnívám se, že klasická forma citace by zde byla příliš matoucí i s ohledem na rozdílné formátování této práce a příručky), příklady pocházejí z příručky vždy.

- (c) významově vyprázdněné užití slovesa je odlišné od významu, který má dané sloveso v bezpříznakovém užití (srov. *dostat knihu – dostat rozkaz*)
- (d) význam celého složeného predikátu je dán významem jmenné části, nikoliv primárně významem slovesa, které vyznačuje zejména aktualizační slovesné významy
- (e) jednotlivá slovesa jsou často provázána videm (*dostat, dostávat*), synonymickou řadou (*dostat, získat*) a příp. fází průběhu děje (*dostat, mít, ztratit*)

Pro jmennou část složených predikátů uvádí anotační příručka jako charakteristické tyto vlastnosti (Mikulová et al. 2005, s. 331–332):

- (f) ve většině případů se jedná o deverbativní substantiva, ale může jít i o jména nede-
verbativní (srov. *učinit rozhodnutí, přiznání... a mít možnost, povinnost...*)
- (g) ve většině případů se jedná o abstraktum (*mít auto – mít možnost*)
- (h) je třeba rozlišit, jestli primárně abstraktní substantivum neoznačuje v dané klauzi
konkrétním (*Nájemce může podle svých požadavků dostat **nabídku** i po telefonu,
popřípadě navštívit kancelář společnosti a **nabídky** dostane vytištěné.*)
- (i) jména tvoří často sémantické řady (např. citové prožitky: *důvěra, dojem, rozčarování, soustrast, nadšení*), synonymické řady (*kontakt, spojení, styk, vztah*) a méně často antonymické řady (*milost – trest, souhlas – zákaz*)
- (j) jmenná část může být stejná pro několik slovesných částí (*dostat/mít/ztratit chuť*)

2.2 Problematika APKS z pohledu lingvistických disciplín a rozcestník odborné literatury

Než se pustím do shrnutí literatury pojící se s APKS, považuji za vhodné upozornit na rozsáhlou totožnost tohoto jazykového fenoménu a dalších jevů, které s ním úzce souvisí, jako např. problematika víceslovných jednotek, celá oblast deverbativních substantiv, která jsou nejčastěji predikativním jménem atd. V této práci se snažím co nejuceleněji pokrýt klíčovou bohemistickou literaturu věnovanou APKS s důrazem na výzkum založený na korpusových datech.

Počátky četného zahraničního zkoumání APKS jsou spojeny s osobnostmi, které položily základní kameny výzkumu, ať už pro „své“ jazyky, nebo v obecném pojetí: obecný lingvista a anglista Jespersen (1942), Polenz (1963) se svou studií o *funkčních slovesech*, Gross (1981), jeho Lexikogramatická jazykovědná škola a na ni navazující metodologie objektových tříd, z generativní teorie řízení a vázání vycházející Cattell (1984), německý gramatik Helbig

(1984) a další. Za poslední více než půlstoletí byly existující poznatky mnohokrát přetaveny do nových (i bohemistických) studií, které na zmíněná díla navazují. V bohemistických studiích můžeme např. pozorovat vliv německé tradice, ze které čerpá Macháčková (1979; 1983; 1994) a později i Martínek (2008b; 2016), francouzské lexikogramatické tradice u Radimského (2010; 2012) i anglofonní literatury, na níž staví Kettnerová a další lingvisté seskupení okolo PDT a VALLEXu.

Pořadí následujících oddílů velmi zjednodušeně sleduje vývoj výzkumu APKS.¹² V počátcích se zkoumala stylistická příslušnost APKS a možnosti, které tyto konstrukce do vyjadřování přinášejí (díky tomu jsou charakterizovány i další rysy APKS, tyto rysy jsou později systematictěji analyzovány), zároveň probíhalo lexikologicko-frazeologické mapování. Následuje fáze detailnějšího popisu z hlediska morfologie, syntaxe a sémantiky, často s „přesahem“ do jiných oblastí lingvistického bádání. Do samostatné podkapitoly pak začleňují výzkum starších vrstev češtiny, které byly dlouho neprávem opomíjeny a jimž se konečně dostalo kýžené pozornosti.

Výzkum APKS a dalších víceslovných výrazů (*multiword expressions*),¹³ mezi které APKS patří, je v posledních zhruba patnácti letech na vzestupu. Je motivován zejména těmito faktory:

- (a) rozvoj korpusů – k dispozici je velké množství jazykových dat, a to i bez nutnosti zdlouhavé ruční excerptce
- (b) potřeba APKS a další víceslovné jednotky formálně uchopit, aby se s nimi dalo lépe pracovat při počítačovém zpracování jazyka – tady mám na mysli nejen anotaci korpusů zmíněných v bodu (a), ale i automatický překlad, odpovídání na otázky, extrakci informací apod. (Barančíková a Kettnerová 2017; Kettnerová, Lopatková a Bejček 2012; Kettnerová et al. 2013)
- (c) skutečnost, že používání APKS sice nevyžaduje zvláštní pozornost od rodilých mluvčích, pro studenty cizích jazyků (anebo cizince snažící se naučit česky) ale tento jev představuje poměrně velký problém, který dlouho pozornosti lingvistů

¹² Jsem si vědoma, že dělení podle různých jazykovědných disciplín je poněkud zjednodušené, ale pro základní přehled by mělo být, doufám, postačující. Detailněji k predikativnímu jménu a kategoriálnímu slovesu v podkapitolách 2.3 a 2.4.

¹³ Těm se v poslední době věnuje mnoho článků, např. Sag et al. (2002), Hnátková et al. (2018) a Petkevič et al. (2020). Pro češtinu jsou dnes např. zpracovávány v databázi víceslovných lexikálních jednotek LEMUR (Petkevič et al. 2020).

unikal; v učebnicích a slovnících se APKS zpracovávají nesystematicky nebo vůbec (Cinková 2009; Radimský 2010)¹⁴

Uplatnění a role APKS při akvizici cizího jazyka a při překladu/tlumočení inspirovaly k výzkumu např. Cinkovou (2009; švédština), Gregora (2008; ruština), Radimského (zejm. 2010; francouzština, italština), jeho diplomantku Povolnou (2012; francouzština) a patrně i další. V této práci se překlady/tlumočením nebudu podrobněji zabývat, zmíněné práce nicméně přispívají k celkovému lingvistickému poznání jevu, tudíž se k nim ještě vrátím. Zmiňovaná kniha Radimského je jednou z nejrozsáhlejších prací věnovaných fenoménu APKS v české bohemistice vůbec a vycházejí z ní téměř všichni pozdější autoři (včetně mě samotné).

Kromě shrnutí základních poznatků o APKS založeného zejména na literatuře, která patří ke kánonu poznání českých¹⁵ APKS, se budu věnovat neprávem opomíjenému Martínkovi (zejm. 2008b; 2011). Poté se zaměřím na větší množství krátkých článků, které byly od vydání Radimského monografie k tématu publikovány a které, pokud vím, nikde shrnuty nejsou.

Mám tím na mysli články, které publikovala skupina lingvistů okolo Kettnerové z Ústavu formální a aplikované lingvistiky MFF UK (dále ÚFAL). Tento kolektiv se v průběhu posledních let¹⁶ APKS věnuje poměrně intenzivně a mezi lety 2015–2017 získal podporu grantového projektu *Kombinování slov: Syntaktické vlastnosti českých víceslovných výrazů s tzv. lehkými slovesy*.¹⁷ Projekt si kladl za cíl na základě teoretického studia vybraných syntaktických, sémantických a morfologických aspektů APKS vytvořit jejich formální model pro lexikografickou reprezentaci a ten pak dále využít ve *Valenčním slovníku českých sloves – VALLEX* (Lopatková et al. 2016). Díky zmíněnému grantu a úsilí jeho řešitelů je anotace APKS od verze 3.5 publikované v roce 2018 součástí online verze slovníku (ke zpracování APKS ve slovníku více oddíl 2.6.2).

¹⁴ APKS se v dostatečné míře neuvádějí v překladových ani výkladových slovnících, což přináší problémy dvojího typu: (a) v některých případech není nerodilý mluvčí schopen takové výrazy náležitě interpretovat a (b) při produkci cizího jazyka je pro nerodilého mluvčího problém nalézt k substantivu správné kategoriální sloveso, protože spojení je těžko předvídatelné (Cinková 2009, s. 11; Radimský 2010, s. 184).

¹⁵ Kralčáková (2005) monografie, kterou považuje za primární Martínek (2008b, s. 23–24), se věnuje obecně APKS ve slovenštině (kritéria, struktury a povaha komponentů, sémantika, valence, lexikální sémantika), zejména se zabývá slovenštinou štúrovskou. V této práci z Kralčákovy publikace čerpám jen zřídka, a to ze tří důvodů: (1) již existuje dostatek literatury věnované specificky češtině, (2) bylo by nutné provést další rešerši případných novějších slovenských zdrojů, na což už v této práci nezbyvá prostor, (3) předmětem této práce jsou české, nikoliv slovenské APKS.

¹⁶ Z podstatné části se zmíněné období překrývá se samotným psaním této disertační práce, které započalo již v roce 2014. Analýza APKS v kontextu jazykových korpusů jako zdroje dat v posledních letech značně pokročila a je třeba poznatky reflektovat, aby se výzkum zbytečně nepřekrýval tam, kde to není nutné.

¹⁷ Projekt GA15-09979S byl v letech 2015–2017 financován Grantovou agenturou ČR, detaily k projektu na webu <https://starfos.tacr.cz/cs/project/GA15-09979S> a <https://ufal.mff.cuni.cz/node/1124>.

2.2.1 Stylistická a komunikační rovina

APKS nalezneme na všech komunikačních úrovních, uplatní se napříč všemi styly a jako takové jsou dnes chápány jako stylově neutrální.¹⁸ Typické jsou zejména pro publicistické, administrativní a odborné texty, ale pronikají i do dalších oblastí, např. jako neotřelé metafory (*proci-tala v něm zvědavost*). Existují obecná spojení stylově neutrální (*mít strach, mít vliv*), nespisovná (*vlepit facku, vyfasovat trest*), APKS se vyskytují dokonce i ve slangu (*vstřelit branku, inkasovat gól*) (Macháčková 1979, s. 75–77; 1983, s. 140–141; Radimský 2017a; Vlková 1990, s. 3–4). Přestože ve starších pracích byly APKS často přijímány negativně (Jelínek (2003, s. 46) jmenuje např. Bečku (1933; 1946), který říká, že s použitím APKS dochází k zesložitování textu a posouvání významu ke zbytečné knižnosti), jejich postavení je dávno narovnáno (Bečka 1992, s. 80–81).

APKS se ze stylisticko-komunikační perspektivy věnovalo množství bohemistů (zejm. Jelínek 2002; Jelínek 2003; 2007; Macháčková 1979, s. 169–178; 1983; Mathesius 1966; Povolná 2012; Radimský 2010; Vlková 1978; 1990) a pro podrobnější popis odkáží čtenáře rovnou na zmíněné studie.

Morfologických, syntaktických a stylistických vlastností nebo zjednodušeně předností APKS, které vedou k popularitě těchto obrátů ve srovnání s případnými konkurenčními prostředky, je celá řada. Zároveň tyto charakteristiky poměrně výstižně vykreslují celé spojení:

- (a) APKS umožňují vyjádřit jádro věty odborným termínem, aniž by bylo potřeba odvozovat sloveso (*provádět kontrolu – kontrolovat*) (Cinková 2009, s. 48–49; Jelínek 2002, s. 37–38; 2003, s. 47–48; 2007, s. 46; Martínek 2008b; Radimský 2017a; Vlková 1990)
- (b) APKS vyplňují mezery v lexikálním systému, pokud neexistuje jednoslovný ekvivalent (*udělat kotoul – *kotoulovat, dělat kotrmelce*) (Macháčková 1983, s. 134; Martínek 2008a, s. 188; 2008b)
- (c) predikativní jméno lze rozvíjet přívlastky (*poskytnout cennou radu – *cenně poradit*) a tím vytvářet komplexnější termíny (*provést šikmou incizi pravé síně, vydat rozhodnutí o předběžné otázce*) (Jelínek 2007, s. 46–47; Radimský 2017a; Vlková 1978, s. 108; 1990)

¹⁸ Nicméně např. ještě Jelínek (2003, s. 43, 48; 2007, s. 40) považoval APKS za knižní nebo alespoň nesoucí „mírné knižní zabarvení“, a to i u spojení jako *podniknout průzkum* místo *prozkoumat*. Možná svoje tvrzení založil na interpretaci konkrétních vět, ty ale bohužel součástí jeho článků nejsou, nemám proto, oč se opřít. Jeho tvrzení ale můžeme srovnat např. s francouzštinou, kde jsou podle Macháčkové (1983, s. 123) APKS řazeny k elegantnímu stylu.

- (d) APKS mohou být stylovou alternativou syntetických predikátů (*podniknout útok – zaútočit*) nebo pasiva (*zlepšilo se – došlo (ke) zlepšení*) a obecně zvyšovat stylistickou pestrost textu (Jelínek 2002; 2003, s. 48–49; Macháčková 1983, s. 134; Martínek 2008b, s. 33; Radimský 2017a; Vlková 1978; 1990)
- (e) oproti plnovýznamovým slovesům má spojení kategoriálního predikátu a několika predikativních jmen shodnou valenci i rekcí (*Obviňujeme ho, vyčítáme mu a stěžujeme si na něj – Vznášíme proti němu stížnosti, obvinění a žaloby.*) (Cinková 2009, s. 49; Helbig 1984, s. 177, cit. dle Martínek (2008b, s. 33))
- (f) užití APKS může podtrhnout závažnost/rozsah akce/děje nebo jeho intenzitu (*vykonat návštěvu ve smyslu oficiálnosti – navštívit přítele; provést úklid celé školy – uklidit doma; prohloubit vliv; propuknout v pláč*) (Jelínek 2002; 2003, s. 49; Macháčková 1983, s. 130, 157–158; Martínek 2008b, s. 131–132; Vlková 1990)
- (g) APKS mohou dějovost naopak i utlumit (*Díky moderním technologiím dochází k výraznému přiosťrování konkurenčního boje. – Výsledky měření parametrů iontového svazku slouží k ověření korektnosti počítačového modelování.*) (Jelínek 2007, s. 43)
- (h) APKS umožňují rozlišit vid i v případech, kdy to jednoslovné protějšky neumožňují (*projevit/projevovat zájem* proti nedokonavému *zajímat se*) (Macháčková 1979, s. 60; 1983, s. 132; Vlková 1990, s. 9)
- (i) rozdílnou kombinací vidu kategoriálního slovesa a predikativního jména mají APKS schopnost vyjadřovat semelfaktivnost (*provést kontrolu*), iterativnost (*provádět kontroly*), distributivnost (*provést kontroly*) nebo intenzitu (*mít úzkosti – tonout v úzkostech*) aj. (Jelínek 2002; Kolářová 2010, s. 111; Macháčková 1979, s. 58–60; Radimský 2010, s. 115–118, 170–176; 2017a)
- (j) výběrem kategoriálního slovesa lze vyjádřit různý způsob slovesného děje (aktions-art) (*získat vládu nad kým – mít vládu nad kým*) (Čermák 1974, s. 290; Jelínek 2003, s. 44; Macháčková 1983, s. 129–132; Radimský 2017a)¹⁹
- (k) kategoriální sloveso může i přes svou vyprázdňenost nést pozitivní (*těšit se pevnému zdraví*) nebo negativní konotace (*provést lotrovinu, spáchat sebevraždu*) (Radimský 2010, s. 111–112; 2017b)
- (l) na rozdíl od syntetických sloves APKS přinášejí velkou flexibilitu z hlediska aktuálního členění – je možné umístit děj na pozici tématu/rématu, příp. odkázat na

¹⁹ Další sémantické třídění a odkazy na literaturu viz oddíl 2.2.2.

děj nebo stav bez nutnosti odkazovat na účastníky nebo původce děje (*Média jsou tu od toho, aby ve veřejném zájmu vládu kontrolovala. Tuto kontrolu provádějí tím, že otevřeně informují o všem. Bez právní ochrany však takovou kontrolu provádět důsledně nelze*).²⁰ (Cinková 2009, s. 49; Jelínek 2002; 2003, s. 46–47; 2007, s. 46–47; Kolářová 2010; Macháčková 1983, s. 132, 135–136, 142–143; Martínek 2008b, s. 33; Radimský 2010, s. 76, 176–177; 2017a; Vlková 1990, s. 9)

(m) u APKS je většina argumentů intenčního pole fakultativní, zatímco u syntetických sloves je obvykle vyžadován subjekt a často i další valenční doplnění (*dělník opravil vodovod – dělník provedl opravu (vodovodu)*) (Kolářová 2010, s. 116–118; Macháčková 1979, s. 70; 1983, s. 135; Martínek 2011, s. 125–126; Mathesius 1966; Radimský 2010, s. 73–74, 176)

Některým ze zmíněných bodů se budu věnovat i v dalších částech této kapitoly, ostatní nejsou pro tuto práci zásadní, a zůstanu tak u stručného přehledu.

2.2.2 Lexikologie, lexikografie a frazeologie

K APKS lze přistupovat také jako k frazémům (a idiomům)²¹ či kolokačním jednotkám, které mají nepředvídatelné syntaktické vlastnosti (k APKS jako analytickému predikátu viz 2.5). I pokud APKS chápeme spíše jako syntaktickou jednotku se speciálními vlastnostmi, můžeme souhlasit s tím, že některé APKS jsou již více či méně frazeologizované, protože v některých případech nesplňují běžné rysy APKS (např. nelze oddělit slovesnou složku: *vejít ve známost – *známost, ve kterou vešlo*) (Radimský 2010, s. 155–156). Čermák (1974, s. 296) poukazuje též na přechod k idiomům u ustálených spojení *jít svou cestou* nebo *mít poslední slovo*. Kettnerová, Kolářová a Vernerová (2017, s. 213–216) považují za idiomatické APKS ty, ve kterých je deverbativní substantivum vyjádřeno jako adverbiale kategoriálního slovesa (tj. pomocí předložkových skupin – *mít v plánu*), neboť taková spojení vykazují menší míru pravidelnosti v morfematickém vyjádření valenčních komplementů než APKS s přímým předmětem bez předložky (*mít plán*).

²⁰ Příklady jsou mou vlastní variací na jeden z příkladů uváděných Radimským (2017a). Na APKS *provádět kontrolu* mají ilustrovat postavení tématu (T) a rématu (R).

²¹ Frazémy (spojení analyzované z hlediska formálních rysů) a idiomy (analyzované z hlediska sémantiky) chápu jako kombinaci slov se zčásti nebo zcela neodvoditelným významem, která je zároveň jedinečná a ustálená (Čermák 2017a).

V tradičních slovnících nejsou APKS příliš dobře pokryté – bývají zahrnuty jen v omezené míře, a to pro vysokou frekvenční kolokabilitu substantiva a slovesa; mezi frazeologismy jsou řazeny jen zřídka (srov. Cinková 2009, s. 15–16; Radimský 2010, s. 159). Velkou výjimkou je valenční slovník VALLEX (viz 2.2.3.3.2 a 2.6.2), ve kterém APKS zahrnuty jsou. Některé z nich jsou též součástí lexikální databáze víceslovných jednotek LEMUR, která verbo-nominální spojení s abstraktem zpracovává jako kvazifrazémy (Petkevič et al. 2020, s. 43).

V tomto oddílu se soustředím na přístup k APKS jako k ustálené jednotce, na původ a tvorbu tohoto spojení a na jeho sémantiku. Dalších rysů se okrajově dotýkám v rámci celé této kapitoly věnované APKS.

2.2.2.1 APKS jako ustálená jednotka

Čermák (1974)²² se ve svém příspěvku zeširoka věnuje ustáleným dvouslovným spojeníům verbo-substantivum a mimo APKS analyzuje i obecná spojení predikativního slovesa s objektem, jako *splnit podmínku, vyslechnout projev* (srov. Macháčková 1979, s. 11–12, 62; Radimský 2010, s. 25–26). Zmíněná spojení chápe jako „ustálené jednotky zásadně neidiomatické a většinou též nemetaforické“ (Čermák 1974, s. 287) a přistupuje k nim na základě frekvence a kolokací. Díky tomu si mj. všímá, že spojitelnost sloves je přímo úměrná jejich frekvenci (podle frekvence sestupně slovesa *mít, být, dát, ztratit, dostat...*) (Čermák 1974, s. 298–299), což přejímá např. Radimský.

V pozdějších studiích (shrnutých v *CzechEncy* (Čermák 2017d), ze které vycházím) Čermák tato spojení nazývá *verbonominálními frazémy*, které uvádí jako nejběžnější případ tzv. *kvazifrazému*²³ a charakterizuje je kombinací slovesa a abstrakta v roli objektu. Za specifický nepovažuje význam jednotlivých komponentů, ale malý počet sloves, se kterými se podstatné jméno tvořící frazém pojí. Jako příklad uvádí mj. substantivum *pozor*, které se pojí pouze se slovesem *dát/dávat pozor*.

Z příkladů uvedených ve zmíněném slovníkovém hesle *verbonominální frazém*, na nichž Čermák ilustruje fázovost (viz dále), najdeme některé ve slovníku VALLEX (*zájem – projevit, jevit, mít; spor – ? vyvolat, ? mít, vést*),²⁴ u jiných VALLEX kategoriální čtení neregistruje (*spor – začít, řešit, vyřešit, ukončit; zájem – ztratit*). Čermák dodává, že „[z] obecného hlediska jsou tyto specifické kvazifrazémy, a zvláště jejich konfigurace, ukazující na složitější

²² Detailní a kritické shrnutí tohoto článku viz Martinek (2008b, s. 19–22) nebo Radimský (2010, s. 25–26).

²³ „Frazém, jehož idiomatičnost, měřitelná počtem anomálií a nerealizovatelných transformací, je relativně malá a taková kombinace se pak výrazně blíží výrazu pravidelnému; není s ním však totožná“ (Čermák 2017c).

²⁴ Znakem ? jsou značena slovesa, u kterých VALLEX registruje i kategoriální funkci, ale podstatné jméno *spor* není uvedeno mezi jejich predikativními jmény.

kombinatorické struktury v nominaci, jedinečné a většinou nenahraditelné, tj. jsou (kromě pracovních autorských metafor) jediným způsobem, jak takováto abstrakta používat“. Slovníkové heslo poté končí odkazem na heslo *verbonominální predikát*. Z toho důvodu usuzuji, že Čermák vidí možný průnik množiny verbonominálních frazémů a APKS, ač k nim tak sám nepřistupuje. Ohledně preferenci jednoho nebo druhého označení souhlasím s Radimským (2010, s. 155–156), že zcela zásadní je nejen záměr případné studie, ale také hlavní hledisko, které aplikujeme (omezená kolokabilita vs. syntaktické vlastnosti APKS).

2.2.2.2 Původ a tvorba APKS

Z hlediska původu APKS se v bohemistické literatuře kromě kalků z cizích jazyků a diachronní motivace (srov. Jelínek 2002; 2003; Martínek 2008b; 2009; 2010) akcentuje multiverbizační přístup (Jedlička 1969). V něm je sloveso nahrazeno slovesem a závislým substantivem (*dát podnět – podnítit*); multiverbizace je typická pro odborný styl. Jak plyne z oddílu 2.2.1, věnovaného stylistické a komunikační rovině, není dobré APKS chápat jen jako stylistickou alternativu bez dalších charakteristik, protože koexistence dokonale synonymního páru sloves by neměla v jazykovém systému opodstatnění (Macháčková 1979, s. 51–52; Radimský 2010, s. 158–159).²⁵ Komentář k jednoslovným synonymním ekvivalentům dále v podkapitole 2.4.4.

Multiverbizaci v souvislosti s APKS se věnuje Vlková (1978; 1990), která mezi APKS zařazuje i jednotky utvořené na základě analogie (*dát přednost* podle *dát podnět*) (Vlková 1978, s. 106; 1990, s. 2–3). Vztah jednoslovných a multiverbálních spojení považuje za symetrický a poukazuje na tendenci univerbizace, kdy na základě dvouslovných pojmenování „zpětně vznikají i pojmenování jednoslovná, která se však ve většině případů zatím pohybují mimo rámec spisovného jazyka“ (Vlková 1990, s. 6).²⁶

Kromě Vlkové se motivaci tvoření APKS (ovšem ve slovenštině) věnoval i Kralčák (2005, s. 93–98), který kromě multiverbizace a vytváření APKS na základě analogie zmiňuje i nemultiverbizované jednotky, kdy vzniká naopak univerbizovaný výraz (*dozerat' → dozor → vykonávat' dozor → dozorovat'*). Přidává také skupinu APKS bez syntetického slovesného korelátu (*mat' prehľad*). Zejména poslední skupinu Vlková ve svých analýzách úplně opomíjí, čehož si všímá již Macháčková (1979, s. 9).

²⁵ Delší komentář k multiverbizačnímu přístupu viz Radimský (2010, s. 26–27, 32–33).

²⁶ Po více než třiceti letech od publikování autorčina článku jsou již mnohá zmiňovaná spojení (*dát přednost – upřednostnit, mít/vzít ohled – zohlednit*) součástí úzu. Některá se patrně neuchytila vůbec (*vzít/brát zřetel – zřetelovat*) nebo možná zůstala součástí mluveného jazyka, případně žargonu (*provést generální opravu – zgenerálkovat, provést/vysadit kurzívou – kurzívovat*). Rychlá kontrola v korpusu SYNv10 neukázala žádné výskyty, ale bylo by zajímavé tuto problematiku prozkoumat v korpusech specializovaných.

2.2.2.3 Sémantika spojení

S APKS se pojí více specifík souvisejících s lexikálním významem, mj. fakt, že díky desémantizaci (viz 2.4.3) může do spojení s jedním predikativním jménem vstupovat několik různých sloves, která ve svém predikativním významu vůbec nejsou synonymní (*mít/utrousit/vyslovit/vznést/podat/předestřít/formulovat/adresovat námitku*). Naopak u podstatných jmen fungují synonymní řady tak, jak jsme zvyklí mimo rámec APKS (*podat informaci/zprávu/sdělení/oznámení*) (např. Radimský 2017a; Vlková 1990, s. 5–6).

K sémantickému třídění APKS (ať už celých spojení, nebo jen kategoriálních sloves) bylo publikováno mnoho příspěvků a komentářů, které si dovoluji shrnout jen stručně.²⁷ Kategoriální slovesa mohou vyjadřovat různé fáze (nastání, počátek, průběh i ukončení) (srov. Čermák 1974, s. 290; Kralčák 2005, s. 45–54; Macháčková 1983, s. 125–127, 146–155, 166–171; Martínek 2008b, s. 26; Radimský 2017a). Kolářová (2010, s. 111) dodává, že většina sloves získá fázový význam až ve spojitosti s predikativním jménem (*dostat chuť, ztratit chuť*). Kromě toho mohou být slovesa rozdělena v různých variacích na Danešovu sémantickou klasifikaci predikátů (stavy, procesy, mutace) (Kralčák 2005, s. 45–54; Macháčková 1983, s. 143–145; Martínek 2008b, s. 116–117, 119–124).

Mnozí autoři se na vlastních datech věnovali též zařazování kategoriálních sloves do větných vzorců. Čermák (1974, s. 300) si všímá, že mezi sledovanými konstrukcemi nejsou verba dicendi (až na *mluvit pravdu*), slovesa duševních stavů (*zapomenout*), citových stavů (*hodit se, líbit se*) a slovesa jako *potřebovat, pomáhat, platit, opakovat a vypadat*. Oproti tomu Macháčková (1983, s. 153–154) i Radimský (2010, s. 118) dokazují, že verba dicendi APKS běžně tvoří. Kettnerová et al. (2013) pozorují, že nejfrekventovanější kategoriální slovesa lze rozřadit do několika málo sémantických skupin: vyjadřují výměnu (např. *vzít, dát*), umístění (např. *pokládat, položit*), pohyb (např. *přicházet, přijít*), přesun (např. *vést*) nebo se vztahují k obecné akci (např. *dělat*).²⁸

Pozoruhodným jevem je korelace čísla predikativního jména a vidu kategoriálního slovesa a jejich vliv na sémantiku spojení. Čermák (1974, s. 290, 291) předpokládal nepříznačkovou kvantitativní korelaci (*provést pokus – provádět pokusy*), ostatní kombinace považuje za příznačkové. S tím nesouhlasí Macháčková (1979, s. 56; 1983), která takové pojetí považuje za zjednodušující, nicméně k další analýze se dostává až Martínek (2008b, s. 117–119) s Radimským (2010, s. 82–90, 172–178). Patrně nezávisle na sobě docházejí ke stejným výsledkům a

²⁷ Detailněji např. Martínek (2008b, s. 103–113) a Radimský (2010, s. 90–104).

²⁸ Sémantické skupiny jsou přeloženy do češtiny, v anglickém originálu se jedná o *exchange, location, motion, transport, an action in a generic way* (Kettnerová et al. 2013, s. IV).

poukazují na významové rozdíly, kterých je možné pomocí těchto prostředků docílit – sg. může vyjadřovat semelfaktivnost (*udělat kotrmelec*) a pl. iterativnost (*udělat kotrmelece*), distributivnost (*dal důtky všem zaměstnancům*) či intenzitu (*mít obavy, dělat si naděje*). Radimský se noří do větších detailů a srovnává APKS i se syntetickým slovesem.

2.2.3 Syntax

V této kapitole rozdělím přístupy k APKS podle převažující syntaktické metodologie, stručně shrnu starší literaturu a zaměřím se zejména na články publikované po roce 2010 – tedy od doby vydání Radimského (2010) monografie. Ucelenější charakteristiku nadto podávají tři následující podkapitoly věnované predikativnímu jménu (2.3), kategoriálnímu slovesu (2.4) a APKS jako celku (2.4.4).

2.2.3.1 Dvourovinná valenční syntax

V českém prostředí APKS ze syntaktického hlediska jako vůbec první zkoumala Macháčková (1979; 1983; 1994).²⁹ Její výzkum je metodologicky zasazen do bohemistům dobře známé dvourovinné valenční syntaxe (dále DVS) F. Daneše, z jehož pojetí vychází i několikrát zmiňovaná *Mluvnice češtiny 3* (Daneš, Grepl a Hlavsa 1987). U APKS jakožto analytického spojení (v terminologii DVS analytického predikátoru) je slovesný význam rozdělen mezi dvě složky. Sloveso nese obecné významové rysy a je jádrem syntaktické valence. Substantivum je považováno za predikát a nese rysy lexikálněsémantické (Macháčková 1979, s. 68–74; 1983, s. 127; Radimský 2010, s. 28–29).

Macháčková ve své kandidátské práci a jejím pozdějším shrnutí (1979; 1983) zkoumá, jak se APKS odlišují od volných a frazeologických spojení, snaží se popsat, jak jsou spojení utvářena (viz 2.2.2), navrhuje sémantické třídy sloves i abstrakt (viz 2.2.2), popisuje vyjadřování slovesného děje, možnosti vyjádření diateze (Macháčková 1983, s. 155–157) a další charakteristiku těchto spojení z hlediska jejich funkce a stylu (viz 2.2.1). Ze syntaktického hlediska se zaměřuje zejména na valenci substantivní z hlediska zachování intenčního pole (Macháčková 1979, s. 68–74), valenci kategoriálních predikátů (Macháčková 1979, s. 67–74) a navrhuje třídění z hlediska počtu doplnění. Pozornost věnuje i dalším dílčím vlastnostem APKS.

²⁹ Zmíněná kandidátská práce Macháčkové (1979) nebyla nikdy publikována a jediná kopie strojopisného výtisku je dostupná v konzervačním fondu Národní knihovny v Hostivaři pod systémovým číslem 000876538, pořizovat kopie vlastním fotoaparátem je možné. Většina pozdějších autorů s výjimkou Radimského odkazuje až na shrnutí práce (Macháčková 1983), v němž chybí zejména různé seznamy, které jsou součástí kandidátské práce (více viz 2.6.1).

Na rozdíl od pozdějších příspěvků k syntaktickému výzkumu APKS Macháčková neměla možnost pracovat s korpusovými daty – veškeré poznatky pocházejí z ručně excerpovaných dat.

Je na místě podotknout, že z DVS vychází i Martínek (2008b; 2009; 2011), např. u sémantického třídění predikátů (Martínek 2008b). Tento autor se ale zaměřuje spíše na starou češtinu, k jeho práci se ještě vrátím v části věnované diachronnímu výzkumu (2.2.4), cituji ho též u dílčích poznatků.

2.2.3.2 Lexikogramatická škola

Na Macháčkovou navazuje Radimský (2010), jehož rozsáhlá práce je dalším zcela zásadním příspěvkem k popisu APKS v češtině. Radimský si klade za cíl na češtinu adaptovat a kriticky zhodnotit metody francouzské Lexikogramatické školy (*Lexique-grammaire*, dále LG), které se do té doby používaly pouze pro jazyky románské. Kromě adaptace je konfrontuje s přístupy k APKS v bohemistice i italianistice a soustředí se zejména na kvalitativní analýzu a popis dat.

LG založil strukturalista Maurice Gross, který vycházel z přesvědčení, že syntaktická pravidla (gramatika) jsou svázaná s lexikem (v opozici ke generativní gramatice N. Chomského). Lexikogramatikové pracují s jednoduchou modelovou větou s plným počtem valenčních doplnění, na níž analyzují syntaktické vlastnosti každé lexikální jednotky. Opěrným bodem celé teorie jsou transformační pravidla, která zachycují pravidelné vztahy mezi konstrukcemi. Lexikogramatikové se kromě detailního popisu sloves a dalších slovních druhů zaměřují zejm. na pokrývání takových oblastí lingvistiky, jako jsou frazeologismy a přísloví, objektové třídy, APKS, jmenná kompozita a později elektronické slovníky i automatické zpracování přirozeného jazyka. APKS se na přelomu 70. a 80. let věnovali M. Gross a J. G. Schniderová a v jejich pojetí má spojení taktéž dvě složky. Stejně jako u DVS je predikátem jmenná složka vyjadřující děj a aktualizátorem je podpůrné sloveso bez vlastního významu, které autoři považují za synsémantikum, jež neváže vlastní argumenty. Z LG se dále odštěpila metoda objektových tříd (*classes d'objets*) Gastona Grosse, jejíž hlavní hypotézou je, že výběr kategoriálního slovesa je závislý pouze na sémantických vlastnostech predikativního jména, s čímž Radimský v teoretickém popisu APKS nesouhlasí (Radimský 2010, s. 11–24).

Díky tomu, že Radimský metodologii adaptoval v prostředí, kde LG není původní, měl volnější ruku v její aplikaci na česká data. Zachoval dva transformační testy pro identifikaci APKS, které pro češtinu mírně upravil (viz oddíl 2.1.2), a v několika bodech se od původní metodologie zásadně odchýlil:

- 1) Při analýze postupuje od predikativního jména ke kategoriálnímu slovesu (stejně jako pozdější metoda objektových tříd),³⁰ čímž je schopen mj. dokázat, že kategoriální funkci může mít téměř libovolné sloveso, a také lépe popsat predikativní jména jako taková.
- 2) Analyzuje korpusová data³¹ – může se blíže zaměřit na kvalitativní rozdíly mezi jednotlivými realizacemi a v některých případech i na kvantitativní charakteristiku.
- 3) Za (analytický) predikát považuje obě složky APKS, zatímco LG chápe jako „predikát“ pouze jmennou složku spojení.

Radimského závěry a pozorování výrazně ovlivnily podobu následujících podkapitol o APKS jako analytickém predikátu, predikativním jméně a kategoriálním slovese, v nichž Radimského doplňuji o zjištění z novějších výzkumů. Soustředím se v nich na poznatky o češtině, odlišnosti Radimského pojetí oproti LG ani výzkumy italštiny tematizovány nejsou.

2.2.3.3 Funkční generativní popis

Nejnovější studie věnované APKS jsou spojeny s funkčním generativním popisem (dále FGP),³² který vznikl v 60. letech 20. století (tedy ve stejné době jako DVS) na UK a byl zamýšlený především pro češtinu (Sgall 1967). Jedná se o formalizovaný popis jazyka, který je postavený na závislostní syntaxi a zachycuje vztahy mezi jednotkami různých rovin jazykového systému. FGP je stále živý a neustále rozvíjený zejména na základě zpracovávání dat při anotacích PDT,³³ kde je jeho prostřednictvím popsána tektogramatická (významová, hloubková) rovina a zjednodušeně i analytická (povrchová) rovina jednotlivých vět.

2.2.3.3.1 Značení APKS v PDT

Na začátek si dovoluji krátkou, ale důležitou odbočku ke značení APKS v PDT. Značení v PDT detailně popisuje anotátorská a referenční příručka (Mikulová et al. 2005, s. 301–357; Mikulová

³⁰ Podle Radimského je původní lexikogramatický přístup od slovesa ke jménu v rozporu s východiskem LG, že predikátem je jmenná část selektující si sloveso a že jsou zároveň kategoriální slovesa díky desémantizaci ze synchronního pohledu arbitrární. Radimský připouští, že u francouzštiny, kde se APKS tvoří zejména s několika základními slovesy (*faire* – dělat, *donner* – dát, *tenir* – držet, *prendre* – vzít, *avoir* – mít), bylo v tomto směru jednoduché vyjít od sloves, ale pro češtinu je takový přístup stěží realizovatelný. Radimský proto vychází z frekvenčních seznamů a vybírá potenciální predikativní jména, ke kterým hledá kategoriální slovesa.

³¹ Tj. neanalyzuje umělé, introspektivně vytvořené příklady posouzené rodilým mluvčím jako (ne)přijatelné. Takový zdroj příkladů byl u LG běžný, neboť v době zpracování tématu nebyly korpusy tak rozvinuté jako dnes, a patrně proto je LG odmítá (Radimský 2010, s. 12–13).

³² Zdrojem ke stručnému shrnutí mi bylo heslo z *CzechEncy – Nového encyklopedického slovníku češtiny* (Panevová 2017), které může sloužit jako vhodný rozcestník pro případnou další literaturu.

³³ Projekt PDT probíhá od roku 1996 na ÚFAL a v Centru počítačnické lingvistiky na MFF UK; stránky projektu: <https://ufal.mff.cuni.cz/prague-dependency-treebank>.

et al. 2006, s. 88–97).³⁴ APKS, ve smyslu pojednáváném v této práci, se v příručkách ukrývají pod dvěma níže definovanými termíny; kritéria pro anotaci jsou v obou případech stejná (kromě typu sloves, se kterými se spojení pojí) a jsou rozepsána v pododdíle 2.1.2.4. V datech nejsou skupiny rozlišeny a charakterizuje je stejná značka CPHR (viz dále).

- (a) složené predikáty (*provést kontrolu, učinit rozhodnutí*) – centrální skupina APKS, bezpředložkové i předložkové pády s výjimkou dativu (viz níže)
- (b) kvazimodální a kvazifrázové predikáty³⁵ (*mít schopnost, dostat chuť, sbírat odvalu, být schopen, být možno, být nutno*) – část odpovídá pojetí APKS s kategoriálními slovesy omezenými na slovesa modální, zároveň tato skupina označuje i případy, kdy jsou na pozici jmenné části adverbia (která ostatně někteří lingvisté mezi APKS řadí)

Mikulová et al. (2005, s. 330–331) zároveň podotýká, že kvůli striktním pravidlům značení možná nebyly označeny všechny APKS – a naopak tuto možnost považuje za pravděpodobnou. Kolářová (2010, s. 112) dodává, že záměrem anotace PDT bylo „vytipovat všechny konstrukce s gramatickou koreferencí (tj. tzv. složené predikáty kontroly) a dále složené predikáty, jejichž jmenná část má nějaké vlastní valenční doplnění, jehož slovosledné postavení může způsobovat neprojektivní konstrukce“.

V PDT není zachycena nominalizace slovesné části složených predikátů (*věnování pozornosti, pozornost věnovaná dětem*) (Mikulová et al. 2005, s. 330). Samotné označení *složené predikáty* bylo zvoleno kvůli tomu, že pojem *verbonominální predikáty* je v příručkách k PDT rezervován pouze pro spojení slovesa *být* se substantivou a adjektivou (*být veselý, být vítězem*) a pojem *analytický* zase pro analytickou rovinu popisu (Kolářová 2010, s. 109).

Dělení víceslovných predikátů v PDT, potažmo anotátorské příručky (Mikulová et al. 2005), kritizuje Martínek (2008b, s. 31–32). Nesouhlasí s úzkým pojetím verbonominálních predikátů, jehož důvod je ale dohledatelný a z mého pohledu ospravedlnitelný. Druhá Martínková výtku směřuje k tomu, že APKS s předložkou jsou hodnoceny jako *slovesná frazeologická spojení* (uvádí příklad *mít v užívání*). Mám za to, že se bude jednat o jednotky příkladů dané vysokou kolokabilitou spojení, nicméně s touto výtkou lze souhlasit. Problém nicméně spatřuji v tom, že Martínek tvrdí, že příručka složené predikáty charakterizuje nesprávně jako

³⁴ Nenašla jsem nikde zmínku o tom, že by se příručky ve smyslu dále zmíněného měnily, ale doporučuji případnému budoucímu čtenáři informace zkontrolovat.

³⁵ Na kvazimodální a kvazifrázové predikáty se odkazuje jako na podtyp složených predikátů (Mikulová et al. 2005, s. 329).

„bezpředložková spojení funkčního slovesa a substantiva v akuzativu“ (Martínek 2008b, s. 32), taková definice se ale v anotátorské příručce nevyskytuje. Naopak Mikulová et al. (2005, s. 332) uvádí, že „jmenná část složeného predikátu může mít následující formy: bezpředložkový akuzativ (ve většině případů): *mít potíže*, nominativ: *padlo rozhodnutí*, bezpředložkový instrumentál: *hořet nenávisť*, bezpředložkový genitiv: *pozbyt vtipnosti* [a] předložkový pád: *dát se do práce, přijít o možnost*“. Jedinou výjimkou je dativ, kdy „[s]pojení slovesa se jménem v dativu (*propadnout apatii, podlehnout zmatku*) prozatím jako složený predikát neoznačujeme“ (Mikulová et al. 2005, s. 333). Autoři bohužel (minimálně v této části příručky) nevysvětlují, proč tomu tak je.

Pro samotnou práci s PDT nicméně stačí vědět, že APKS jsou zachyceny dvěma uzly – jedním pro kategoriální sloveso (obvykle predikát) a druhým pro predikativní jméno, které je přímým potomkem kategoriálního slovesa a je označeno funktorem³⁶ CPHR (*compound phrase*). CPHR zde vystupuje namísto Patientu³⁷ a v tomto specifickém případě nedochází k posouvání aktantů,³⁸ které je jinak pro FGP charakteristické (Kettnerová a Bejček 2016, s. 516; Kolářová 2010, s. 110).

Kolářová (2010, s. 109–111) zdůrazňuje, že přestože FGP považuje APKS ze sémantického hlediska za jednu jednotku, jsou v PDT víceslovné jednotky zachyceny každá svým vlastním uzlem (až na výjimky v některých souslovích), a to i na tektogramatické rovině, kde je substantivum vyznačeno jako závislé na řídicím slovesu. Na tektogramatické rovině se APKS jako jedna lexikální jednotka projevuje referenční totožností části valenčních doplnění – v povrchové struktuře jsou zpravidla tato valenční doplnění vypuštěna, ale na tektogramatické rovině jsou zobrazena a je jim přiřazen speciální uzel #QCor (Quasi-Control), do kterého vede koreferenční šipka k valenčnímu doplnění slovesa, s nímž je uzel totožný.

Pro ukázkou přikládám tektogramatický strom věty *Televizní reportér klade otázky náhodně vybraným lidem* (Obrázek 2 níže).³⁹ Kategoriální predikát *klást* je řídicím členem

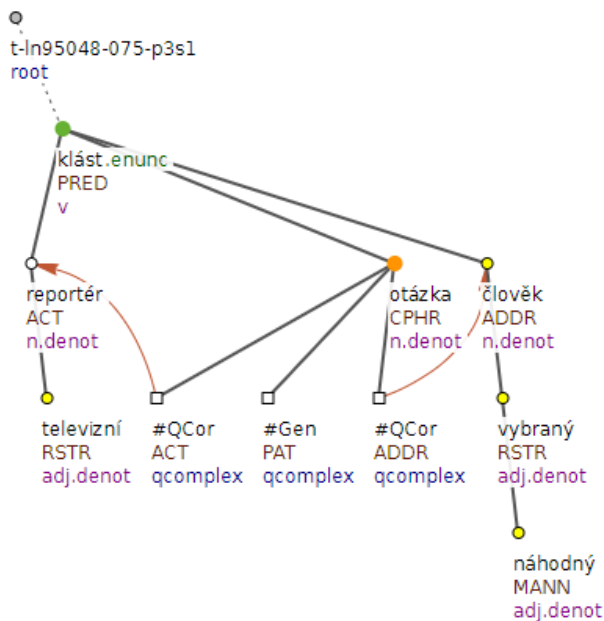
³⁶ Syntakticko-sémantické označení vztahu uzlu řídicího k uzlu závislému, zde valenčního doplnění k řídicímu slovesu (Kettnerová a Bejček 2016, s. 516).

³⁷ FGP píše názvy aktantů (Konatel, Patiens, Adresát, Původ (Origo) a Výsledek (Efekt)) s velkým písmenem na začátku slova. V této části věnované FGP ponechávám názvy v souladu s terminologií, v jiných částech práce sjednocuji na malé písmeno.

³⁸ V rámci FGP jsou první dva aktanty určeny na základě syntaktických kritérií: syntaktický subjekt je označen jako Konatel, předmět pak jako Patiens. Až pro určení dalších aktantů se používají kritéria sémantická. Pokud sloveso se dvěma aktanty nemá Patiens, posune se druhý aktant do jeho pozice (*oslovit někoho*, Adresát → Patiens) (Kolářová 2010, s. 35–36).

³⁹ Ilustrační tektogramatický strom pochází z webového rozhraní PDT dostupného na adrese <https://lindat.mff.cuni.cz/services/pmltq/#!/home>. K vyhledání jsem použila následující dotaz v dotazovacím jazyce PML-Tree Query (PML-TQ) na věty obsahující sloveso *klást* následované funktorem CPHR:

predikativního jména *otázky* (s funktorem CPHR), ze dvou uzlů s tektogramatickým lemmatem #QCor vedou koreferenční šipky k uzlům *reportér* a *člověk*, tyto uzly jsou zároveň valenčním doplněním slovesa. Ve stromu si můžeme povšimnout ještě jednoho uzlu, který není povrchově reprezentován a má tektogramatické lemma #Gen. Ten označuje systémovou gramatickou elipsu, která může být určena na základě širšího kontextu; tento aktant není součástí žádného typu koreference.



Obrázek 2: Zobrazení APKS a koreferenčních vztahů mezi jeho členy v PDT – bílé čtverce značí uzly, které nejsou povrchově vyjádřené, oranžové šipky znázorňují koreferenci
Zdroj: PDT 3.5

V následujících studiích se navíc uplatní ještě další podmínka: oba členy se musí vyskytovat v rámci jedné klauze. Jsou tedy vyloučeny případy, kdy je predikativní jméno relativizováno (*Příkaz, který Petr dostal.*) a funktor CPHR je přiřazen vztažnému zájmenu (Kettnerová a Bejček 2016, s. 517).

2.2.3.3.2 Příspěvky k APKS (zejm. po roce 2010)

Ve své kontrastivní práci o švédských základních slovesech⁴⁰ se Cinková (2009) věnuje užití základních sloves v přeneseném významu s omezenou kolokabilitou – nejčastěji právě APKS. Tato spojení podle Cinkové (2009, s. 11) nepůsobí pokročilým studentům problémy při

t-node [t_lemma="klášť", t-node [functor = "CPHR"]]. Více informací k dotazovacímu jazyku zde: <https://ufal.mff.cuni.cz/pmltq>.

⁴⁰ „Základní slovesa (*basic verbs*), tj. frekventovaná významová slovesa, jež zpravidla popisují fyzický pohyb, umístění, stav, nebo děj, procházejí řadou sémantických posunů, díky kterým se používají k vyjádření druhotných, přenesených významů“ (Cinková 2009, s. 11).

porozumění textu (jsou sémanticky transparentní), ale při jeho produkci ano (jsou jazykově specifická a neprediktabilní). Rodilí mluvčí je považují za frazeologismy, a proto jim při výuce švédštiny jako cizího jazyka není věnována dostatečná pozornost. Cinková se snaží tuto mezeru vyplnit – poznatky zobecňuje a začleňuje do valenčního slovníku švédských sloves (podobně jako český VALLEX) a do slovníku nominálních částí analytických predikátů (oba reflektují valenční teorii FGP).

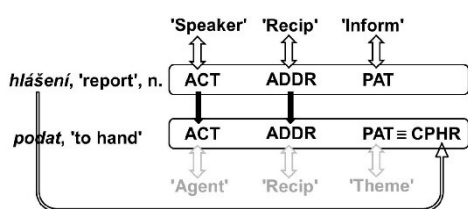
Na substantiva jako součást složených predikátů v PDT se v rámci monografie věnované valenci deverbativních substantiv zaměřuje Kolářová (2010, s. 109–126). Autorka pracuje zejména se záznamy z PDT-VALLEXu⁴¹ a teoreticky navazuje především na Čermáka (1974), Macháčkovou (1983) a zpracování anglických APKS. Soustředí se na to, jaké důsledky na valenční vlastnosti substantiva má jeho zapojení do APKS – charakterizuje třetí valenční doplnění kategoriálního slovesa, změny ve valenčním rámci kategoriálního slovesa oproti bezpříznakovému užití sloves, analyzuje konkurenční doplnění a sdílení některých doplnění oběma složkami, nakonec se věnuje aktuálnímu členění a neprojektivním konstrukcím. To vše prokládá značným množstvím příkladů a objasňuje zachycování APKS v PDT. K jejím poznatkům o valenci se vracím u predikativního jména (oddíl 2.3.2) a kategoriálního slovesa (pododdíl 2.4.4.4).

Na výsledky obou uvedených disertačních prací navazuje kolektiv lingvistů kolem Václavy Kettnerové z ÚFAL v sérii článků vznikajících v rámci zmíněného grantového projektu *Kombinování slov: Syntaktické vlastnosti českých víceslovných výrazů s tzv. lehkými slovesy* (viz pozn. 17). Hlavní motivací autorů je kromě rozšíření slovníku VALLEX také snaha popsat syntaktický mechanismus chování APKS, lexikální materiály existující v době řešení grantu, které APKS popisují, totiž neposkytují dostatečné informace pro správné zachycení povrchových a hloubkových struktur (Kettnerová a Lopatková 2015, s. 192). Teoretické výstupy shrnuje příspěvek *Syntaktická struktura komplexních predikátů v češtině* (Kettnerová 2017), ke kterému se ještě vrátím. Níže se snažím podat přehled jednotlivých článků zmíněného kolektivu.

⁴¹ PDT-VALLEX je valenční slovník, který vznikal průběžně při anotaci tektogramatické roviny PDT, aby zajistil konzistenci při anotování valence sloves i podstatných a přídavných jmen. Slovník obsahuje záznamy o významech vyskytujících se výhradně v PDT, tj. v anotovaných datech. Kromě jiného jsou ve slovníku zachyceny i valenční rámce sémantických substantiv tvořících jmennou část složeného predikátu, tj. takových, která mají funktor CPHR. PDT-VALLEX se výrazně odlišuje od VALLEXu, do kterého jsou slovesa vybírána podle frekvence, dělena do sémantických tříd a zpracovávána detailněji a s ohledem na všechny významy (Kolářová 2010, s. 48–49). Poslední verze PDT-VALLEXu je dostupná na <http://hdl.handle.net/11234/1-3499>.

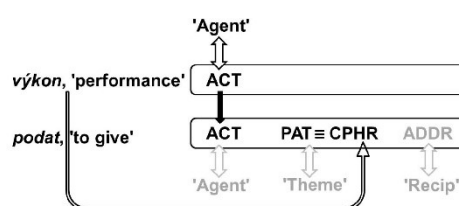
Článek Kettnerové et al. (2013) byl věnován identifikaci APKS, resp. sestavení seznamu APKS pro další výzkum a otestování základních kritérií (jak bylo představeno v oddíle 2.1.2). Výzkum pracuje s 59 nejčastějšími slovesy pojícími se s funktorem CPHR v PDT-VALLEXu a 100 náhodnými vzorky z ČNK ke každému z nich. Tři anotátoři tyto vzorky prošli a rozhodli, jde-li o kategoriální nebo predikativní použití slovesa. Získaný seznam čítá 893 kombinací kategoriálních sloves a predikativních jmen. Během výzkumu také autoři poukazují na to, že analyzovaná slovesa lze rozřadit jen do pěti sémantických tříd (viz 2.2.2.3).

Budoucí podobu hesel zachycujících APKS ve VALLEXu pokrývá několik studií (Kettnerová a Lopatková 2013; 2015; Kettnerová, Lopatková a Barančíková 2016). Z hlediska hloubkové struktury věty autorky charakterizují sémantické participanty obou členů APKS zvlášť (např. Kettnerová a Lopatková 2013, s. 148–150; Kettnerová, Lopatková a Bejček 2012), a následně navrhují způsob, jak propojit slovesné i jmenné doplnění. Pozici Patientu kategoriálního slovesa obsadí predikativní jméno (tedy funktoř CPHR), zbývající pozice nasytí doplnění predikativního jména (viz Obrázek 3). Pokud není valenční pozice kategoriálního slovesa obsazena, je z rámce odstraněna (Obrázek 4). Z hlediska vyjádření těchto doplnění v povrchové struktuře věty (a následně i ve slovníku) autorky vycházejí z toho, že ze slovesného valenčního rámce jsou zobrazené jen ty argumenty, které jsou nasycené jmenným doplněním. Naopak na povrchové struktuře zůstávají ta doplnění, která nejsou propojená s žádným slovesným doplněním (např. Kettnerová a Lopatková 2013, s. 150–153; 2015, s. 196–199; Kettnerová, Lopatková a Barančíková 2016, s. 887–888; Kettnerová et al. 2018, s. 33–35).



Obrázek 3: Propojení slovesných valenčních doplnění se jmennými doplněními (černé šipky) a jejich saturace nominálními sémantickými participanty v APKS *podat hlášení*

Zdroj: Kettnerová a Lopatková (2013, s. 150)

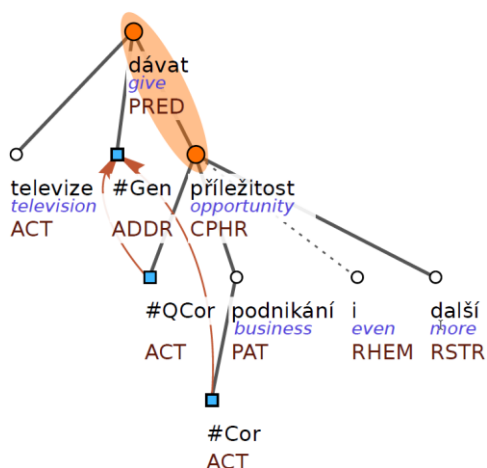


Obrázek 4: Propojení slovesných valenčních doplnění se jmennými doplněními (černé šipky) a jejich saturace nominálními sémantickými participanty v APKS *podat výkon*, viz zejm. nepropojený a odstraněný Adresát

Zdroj: Kettnerová a Lopatková (2013, s. 150)

Toto pozorování později zpřesňují Kettnerová a Bejček (2016, s. 519) s tím, že Konatel je vyjádřen vždy. Tento poznatek přetavují v hlavní kritérium určování APKS, společně s koreferencí (Kettnerová 2017, s. 7–8), viz příklad (25) a jeho struktura (Obrázek 5) níže.

(25) Televize_{ACT} dává i další příležitosti k podnikání. (PDT)



Obrázek 5: Zjednodušená tektogramatická struktura věty z příkladu (25)

Zdroj: Kettnerová a Bejček (2016, s. 519)

Ke konkrétním návrhům na samotnou reprezentaci lexikografického hesla zachycujícího kategoriální sloveso a predikativního jméno ve slovníku a řešení vzájemného mapování obou členů viz Kettnerová a Lopatková (2013, s. 154–155; 2015, s. 193–196); Kettnerová, Lopatková a Barančíková (2016, s. 883–887); Kettnerová et al. (2018, s. 45–52).

Kettnerová a Bejček (2016) se snaží postihnout principy distribuce valenčních doplnění APKS v datech PDT a stanovit kritéria pro rozlišení valenčních doplnění predikativního jména a kategoriálního slovesa. Největším problémem, se kterým se anotátoři musejí vypořádat, je rozlišení mezi slovesným a jmenným doplněním. Ve většině případů sice situace bývá jednoznačná, někdy je ale povrchová struktura valenčního doplnění stejná pro kategoriální sloveso i predikativní jméno. V těchto případech jsou doplnění plošně přisouzena kategoriálnímu slovesu⁴² (Kettnerová a Bejček 2016, s. 516–517). Autoři pracují s 1695 větami, které rozdělili na dvě skupiny. Menší část posloužila pro stanovení principů stojících za valenční distribucí, větší část autoři společně s dalšími daty z české verze PCEDT⁴³ využili pro automatizovaný

⁴² Oproti přístupu Kolářové (2010, s. 115), která považuje za silnější valenci predikativního jména.

⁴³ Prague Czech-English Dependency Treebank 2.0, dostupné z <http://ufal.mff.cuni.cz/pcedt2.0/en/index.html>.

experiment.⁴⁴ Výsledkem je poznatek, že doplnění z valenčního rámce kategoriálního slovesa jsou přítomná na povrchové struktuře, pokud existuje koreference (tzv. kvazikontrola) se jmennými doplněními, jedinou výjimkou je Konatel, u kterého není kvazi kontrola nutná. Naopak v případech doplnění predikativního jména jsou povrchově zachycená pouze ta doplnění, která nejsou ve vztahu koreference ke kategoriálnímu predikátu. Toto pravidlo se nevztahuje na další doplnění nad rámec valenčních doplnění obou členů APKS, která vyjadřují okolnosti děje, jako místo, čas, způsob apod. Podle autorů tato pravidla až na drobné výjimky⁴⁵ v analyzovaných datech platí.

Pro samotné rozšíření VALLEXu o popis kategoriálních predikátů v něm autoři (zjednodušeně řečeno) nejprve na základě průniku se značením v PDT-VALLEXu dohledali slovesa, která mohou utvářet APKS, nicméně ve VALLEXu zatím měla pouze kategoriální význam. K nim pomocí kolokací v nástroji Sketch Engine doplnili substantiva v přímém předmětu v akuzativu a nechali anotátory rozhodnout, jestli spojení splňují kritéria APKS. Anotátoři pak vybrali správný valenční rámec predikativního jména podle podkladů z PDT-VALLEXu a přiřadili roli Kauzátoru, pokud to bylo pro spojení relevantní (Kauzátor se týkal jen 13 % spojení) (Kettnerová, Lopatková a Barančíková 2016, s. 888–890; Kettnerová et al. 2018, s. 47–52). Do slovníku VALLEX 3.5 bylo nakonec zahrnuto téměř 1500 APKS (Kettnerová et al. 2018, s. 54).

Kettnerová, Kolářová a Vernerová (2017) popisují deverbativní substantiva v nominálních konstrukcích i v APKS a dokazují, že APKS jsou výsledkem běžných syntaktických operací. V článku se vracejí k hloubkové i povrchové struktuře APKS, a dále představují slovníky NomVallex a VALLEX a lexikografickou reprezentaci nominálních a verbálních struktur. K výsledné verzi VALLEXu a charakteristice a práci s APKS viz 2.6.2.

Teoretické poznatky z grantového projektu jsou shrnuty ve dvou člancích. Prvním je příspěvek *Ke koreferenci u komplexních predikátů s kategoriálním slovesem* (Kettnerová a Lopatková 2017), který podle druhů koreference rozděluje APKS do skupin (odpovídajících atributu *map* ve VALLEXu); tomuto článku se stručně věnuji v pododdíle 2.4.4.1. Druhým příspěvkem je *Syntaktická struktura komplexních predikátů v češtině* (Kettnerová 2017). Zde autorka charakterizuje čtyři typy APKS (viz 2.4.2), na nichž ukazuje, že „utváření syntaktické struktury vět s komplexními predikáty, u nichž predikativní jména obsazují pozici přímého

⁴⁴ Pro ruční analýzu to bylo 649 APKS s 2034 valenčními doplněními. Pro verifikaci bylo využito zbývajících 1036 APKS z PDT (3264 valenčních doplnění) společně s 2116 APKS z PCEDT (počet doplnění není ve studii specifikován) (Kettnerová a Bejček 2016, s. 517–518).

⁴⁵ Po zpřesnění pravidel s Konatelem byla úspěšnost v manuálně zpracovaných datech 95 %, na datech z automatizovaného experimentu dokonce 97 % (Kettnerová a Bejček 2016, s. 518–519).

objektu, je pravidelný proces a že zásadní úlohu v tomto procesu sehrává požadavek funkčního slovesa na koreferenci některého (některých) z jeho valenčních doplnění s doplněním predikativního jména a též případná přítomnost [K]auzátoru u kauzativních funkčních sloves“ (Kettnerová 2017, s. 22).

Díky analýze shromážděného materiálu vzniká také slovník parafrází českých komplexních predikátů s kategoriálním slovesem ParaDi (Barančíková a Kettnerová 2017). Pro tvorbu slovníku autorky využily dříve pořizených seznamů APKS⁴⁶ (viz Kettnerová, Lopatková a Barančíková 2016; Kettnerová et al. 2013). Poté tyto APKS identifikovaly v několika spojených korpusech (přes 20 % APKS se v datech nevyskytlo ani jednou) a pomocí algoritmu mělkých neuronových sítí (*shallow neural network*) k nim hledaly jejich jednoslovný ekvivalent. Získané parafráze⁴⁷ byly ručně rozříděny dvěma anotátory, kteří upřesnili, o jaký typ parafráze se jedná. Z 1776 analyzovaných APKS identifikovaných v korpusech jich 999 mělo alespoň jednu jednoslovnou parafrázi. Praktické využití slovníku bylo dále demonstrováno v experimentu s automatickým překladem (Barančíková a Kettnerová 2017, s. 6–7). Anotátoři měli posoudit kvalitu překladů vět obsahujících APKS – v jedné variantě byly věty přeloženy rovnou, v druhé byly před samotným překladem APKS nejprve nahrazeny parafrází. Ve většině případů si oba překladače vedly lépe při použití parafrázovaného APKS.

Autoři na nasbíraném materiálu dále zkoumali zachycení reciprocit (Kettnerová a Lopatková 2020), která zatím ve slovníku zachycená není. Nedořešenou výzvou zůstává mj. to, jak zachytit číslo predikativního jména a jak zobrazit dvojí (stejnou) koreferenci (Kettnerová a Lopatková 2013, s. 154–155).

2.2.4 Diachronní⁴⁸ výzkum

Jak již bylo zmíněno, APKS je možné vystopovat v mnoha jazycích a velmi hluboko do minulosti; nejinak tomu bylo v češtině a jazycích češtinu ovlivňujících. Tyto konstrukce ale dlouho zájem badatelů neprávem nepřitahovaly.

⁴⁶ Zahrnoval celkem 2257 APKS skládajících se z 154 sloves a 1061 substantiv (Barančíková a Kettnerová 2017, s. 3 elektronické verze).

⁴⁷ Autorky definují tyto podoby parafrází: (a) synonymní parafráze včetně stejných valenčních doplnění (*mít zájem – chtít*), (b) antonymum – nutnost změnit kontext (*zaznamenat propad – stoupnout*), (c) synonymní parafráze s nutností změnit valenci (*dostat nabídku – nabídnout*), (d) synonymní parafráze, ale je třeba přidat reflexivní morfém *se/si* (*nést název – nazývat se*), (e) synonymní parafráze v případě rozvíjení predikativního jména určitým adjektivem (*podat trestní oznámení – žalovat*).

⁴⁸ Termínem diachronní nemíním nutně jen práce, které srovnávají explicitně dvě časová období. Chápu ho i v širším smyslu zkoumání jazykového vývoje, jako opak synchronnímu pohledu na lingvistiku.

Prvním, kdo se za APKS alespoň částečně obrací do historie, je Jelínek (2002; 2003; 2007). Ke studiu APKS přistupuje ze stylistického hlediska, ale vrací se zejména k češtině 19. století a starší. Domnívá se, že popularita APKS v odborném vyjadřování do češtiny přišla z latiny, francouzštiny i němčiny. Stejně tak předpokládá, že větší množství „nových“ APKS bylo během 17. a 18. přejato z němčiny (většina obrozeneckých autorů byla bilingvní) společně s tím, jak se čeština začala používat i v odborné komunikaci. Naopak v poslední třetině 19. století byl tento trend výrazně zasažen jazykovým purismem. Ten považoval APKS za nesystémový import německého nominálního vyjádření a snažil se proti těmto „germanismům“ v podobě podstatných jmen slovesných zakročit. Jelínek (2003, s. 45–46; 2007, s. 44–45) tuto éru odmítání APKS dokládá od prvního vydání matičního *Brusu jazyka českého* v roce 1877 až do roku 1942, kdy *Naše řeč* varuje před napodobováním některých výrazů.

Starší a humanistické češtině (a také latině a němčině) se zevrubně věnuje Martínek (zejm. Martínek 2008b; 2010; 2011; 2016),⁴⁹ který se ostře vymezuje proti starším tvrzením, že APKS jsou projevem výhradně moderního jazyka (Martínek 2008b, s. 43–47). Martínek APKS chápe spíše jako jednotku lexikální (ale odlišnou od frazémů), zaměřuje se na jejich jazykové a stylové charakteristiky, a konečně poznatky o APKS v současném jazyce hodnotí a upravuje pro práci s češtinou historickou. Zachycuje nejen spojení se jmennou částí v akuzativu, ale i v genitivu, dativu, instrumentálu a lokálu. Navrhuje také svou vlastní sémantickou klasifikaci APKS (2008b, s. 113–130). Z důvodu omezenosti a dostupnosti jazykového materiálu analyzuje vybrané texty a zaměřuje se zejména na kvalitativní výzkum a třídění podle jednotlivých rysů (stav vs. děj, fáze, kauzativita atp.). V neposlední řadě se věnuje jazykovým kontaktům v období humanistické češtiny. Poukazuje na působení německých a latinských vzorů na česká APKS, ale také na schopnost jazyka vytvářet případné chybějící APKS při překládání (Martínek 2010).

Pro další studium APKS v kontextu staré češtiny a diachronního pohledu odkáží čtenáře nejprve na výše zmíněné práce; přehled starších článků a publikací k tématu viz Martínek (2008b, s. 47–52; 2016, s. 26–45), který je zároveň podrobuje analýze a kritickému zhodnocení.

2.3 Predikativní jméno

Mezi predikativní jména se nejčastěji řadí verbální substantiva (*rozhodnutí*), dějová jména odvozená od sloves (*analýza*) nebo od přídavných jmen (*možnost, povinnost, zodpovědnost*), dále

⁴⁹ Disertační práce (Martínek 2016) je napsaná v němčině, kterou bohužel neovládám, takže v tomto případě čerpám informace ze zkrácené verze v přiloženém autoreferátu a abstraktu.

podstatná jména neodvozená, jako jsou stavy a nálady (*povodeň, chladné počasí*), fyzikální jevy (*ticho, tma, teplo*), obecné významy s průběhem (*podzim, vhodný okamžik*) a další nede-
verbativní substantiva (*právo, šance*) (srov. Kolářová 2010, s. 111–112; Macháčková 1979, s. 55; 1983, s. 128, 158–165; Radimský 2010, s. 57–58; 2017a).

Predikativní jméno má standardní morfologické kategorie substantiva a můžeme u něj sledovat jisté tendence k vidu (viz deverbativa dále). Kombinaci počitatelnosti a teličnosti a jejich souvislosti se věnoval Radimský (2010, s. 82–90), který dokonce navrhl poměrně komplexní typologii pro počitatelná jména. Z hlediska dalších vlastností jsou predikativní jména vstupující do APKS často sémanticky provázána (např. citový prožitek: *důvěra, něha, soucit, soustrast, touha*), vytvářejí synonymické řady (*dohoda, kontrakt, smlouva*) a někdy i antonymické páry (*souhlas – zákaz, milost – trest*) (Kolářová 2010, s. 111–112; Mikulová et al. 2005, s. 331–332).

Radimský (2010, s. 90) upozorňuje, že predikativní jména sice obvykle vyjadřují děje nebo stavy, to ale neplatí vždy.⁵⁰ Dokumentuje to příklady *dát dárek* a *předat vyznamenání*, která označují jak rezultativum, tak hmatatelný objekt, a proto je možné tato spojení interpretovat predikativně i kategoriálně (k sémantickému vymezení predikativních jmen viz oddíl 2.2.2). Na základě analýzy Radimský dochází k tomu, že výše zmíněné charakteristiky jsou sice výrazným, ale ne vždy realizovaným znakem predikativních jmen. Stanovuje proto dvě klíčové podmínky, z nichž alespoň jednu musí predikativní jméno splňovat (Radimský 2010, s. 104–105):

- (a) přítomnost dějového nebo rezultativního významu (splňují téměř všechna predikativní jména, výjimkou jsou konkrétní jména relační (*přítel, matka, vlast*) splňující podmínku (b))
- (b) existence valenčního a intenčního pole, tj. schopnost jména vázat na sebe argumenty (nesplňují ji pouze dějová jména, která žádné participanty neinterpretují – *došlo k bouřce, panuje mráz*)

Radimský ale upozorňuje, že „[p]otenciální schopnost jména tvořit VNP ještě automaticky neznamená, že se v reálném úzu dané jméno ve VNP skutečně uplatňuje“, a uvádí jako příklad frekventované substantivum *oddělení*, které se jako VNP uplatní jen 21 případech z více než 11 000 celkových výskytů slova v korpusu SYN2000 (*dojít k oddělení, oddělení probíhá*) (Radimský 2010, s. 105).

⁵⁰ Srovnáním starších přístupů k rozlišování mezi abstrakty a konkréty se ve své práci věnuje např. také Martínek (2008b, s. 58–61).

V následujících oddílech představím charakteristiku deverbativ, která tvoří základnu predikativních jmen, a zaměřím se i na valenční vlastnosti predikativního jména. V posledním oddíle se zaměřím na to, čím může být predikativní jméno rozvíjeno, jelikož to souvisí s empirickou částí práce.

2.3.1 Deverbativa

Pro deverbativní podstatná jména⁵¹ je typická pluralita forem a funkcí – od jedné slovesné báze je možné odvodit celou škálu dějových jmen (*omlouvání – omluvení – omluva – omluvenka*), přestože ne všechny typy jsou vždy obsazené, případně jeden tvar plní více funkcí (*psaní*). Kromě dějového významu se s deverbativy pojí i významy nedějové, jako obory lidské činnosti (*doprava, obchod*), výsledek děje (*výtisk zisk, náklad*), prostředek (*jídlo, postřik, výtah*), písemnosti (*smlouva, žádost*), místo (*průjezd, výčep*) a původce děje (*doprovod, návštěva*) (Radimský 2010, s. 59–60).

Hodnocení dějovosti deverbativ probíhá na škále (dokonavá a nedokonavá verbální substantiva, dějové jméno prosté a konkrétní odvozeniny). Nejblíže slovesu jsou verbální substantiva zakončená na *-ní, -tí*, která uchovávají vid, v některých kontextech si mohou ponechat zvrtné zájmeno *se* (*vzdělávání se – vzdělávání mládeže*), jsou nejprocesuálnější, i když v některých případech mohou metonymicky získávat i nedějový význam (*psaní*), ale jen v některých případech je dějové čtení ztraceno (*brnění, jmění, sklepení*) (Jelínek 1967, s. 570–591; Radimský 2010, s. 60–61).

Kettnerová, Kolářová a Vernerová (2017, s. 208–209) dodávají, že tato verbální substantiva dědí valenční strukturu ze slovesa, a to včetně funktořů a obligatornosti; jen u doplnění v nominativu (→ gen, inst, přivlastňovací přídavné jméno nebo zájmeno) a akuzativu v přímém objektu (→ gen, přivlastňovací přídavné jméno nebo zájmeno) proběhne v závorce uvedený systémový posun. U deverbativ označujících abstraktní výsledek děje kromě výše uvedeného dojde i k nesystémovému posunu, a to u vyjádření patientu (*na + acc, o + loc, pro + acc*). Valenční rámce deverbativních substantiv s konkrétním výsledkem děje se mohou výrazně lišit od svých základových sloves.

Radimský (2010, s. 59–71; 2017a) na základě sondy v ČNK popisuje typy deverbativ vstupujících do APKS. Do jeho původní tabulky shrnující jednotlivé typy deverbativ doplňuji

⁵¹ K deverbativům detailněji viz např. Jelínek (1967, s. 570–591), Kolářová (2010), Stehlíková (2010), Kettnerová, Kolářová a Vernerová (2017) a z diachronního pohledu také Martínek (2013).

pro přehlednost zakončení, podle nichž Radimský vyhledává v korpusu, a také terminologii pro stejný jev zavedený v Grimshaw (1990).⁵²

proces (CEN)	↔		designace (SEN)
A	B	C	D
verbální substantivum		dějové jméno prosté	nomen rei (výsledek děje)
nedokonavé nepočitatelné	dokonavé nepočitatelné	(dokonavé) (počitatelné)	--- počitatelné
<i>omlouvání</i>	<i>omluvení</i>	<i>omluva</i>	<i>omluvenka</i>
<i>-o[u]vání</i>	<i>-ení</i>		<i>-ka</i> ⁵³

Tabulka 2: Derivační škála deverbativ na ose proces–designace podle Radimského
Zdroj: Radimský (2010, s. 59–60, 62, 64)

Závěrem je pozorování, že nejčastějším typem predikativního jména jsou dokonavá verbální substantiva (B) a dějová jména (C+D) doplněná o běžné kategoriální sloveso. Nedokonavá dějová substantiva (A) tvoří APKS méně často a obvykle jsou doprovázena fázovými slovesy (*zahájit, začít, probíhat* atp.), případně slovesy obecnějšími (*provést, provádět, věnovat se*). Důvodem může být, že tyto APKS s predikativním jménem s dějovým významem (*provádět fackování*) jsou navíc často jen ekvivalentem syntetického slovesa (*fackovat*). Místo toho je pro APKS typičtější substantivum s rezultativním významem (C+D, *dát facku*), kdy se kombinací čísla predikativního jména a vidu kategoriálního slovesa vyjadřují sémantické nuance (srov. *dát facku – dát facky – dávat facku – dávat facky*), což jde proti původnímu lexikogramatickému předpokladu. Kategoriální sloveso těmto predikativním jménům propůjčuje dějové čtení, sloveso tedy označuje proces a jméno jeho výsledek (podrobněji viz Radimský 2010, s. 99–104).

Ze sémantické stránky se Radimský (2010, s. 90–99) zaměřuje také na polysémní dějová jména (konkrétní vs. predikativní), kdy spojení *podepsat smlouvu* může znamenat jak uzavření smlouvy, tak fyzický podpis papírové smlouvy (tzv. souběžná aktualizace významů). U některých jmen je sice možná obojí interpretace, ale sloveso aktualizuje jen jedno čtení, jako v případě *dát branku* (*branka* jako gól anebo jako předmět vymezení průchod či cíl), tzv. oddělená aktualizace významů.

⁵² Opozice dějových jmen – *complex event nominals* (CEN) – a nedějových jmen – *simple event nominals* (SEN).

⁵³ Zde Radimský (2010, s. 65) ještě musel ručně odebrat zhruba polovinu výskytů, protože koncovka zahrnuje i konkrétní typu *ruka* nebo jména neodvozená typu *láska*.

2.3.2 Valence

Základní vlastností predikativního jména je existence vlastního valenčního a intenčního pole (např. Kolářová 2010, s. 116–119; Radimský 2010, s. 73–82). Na pozici syntaktického podmětu se realizuje jedno z valenčních doplnění predikativního jména (*sémantický subjekt* v LG) a ostatní argumenty zůstávají v pozici přívlastků, což Radimský (2010, s. 73) ilustruje na následujícím příkladu.

(26) válka (Německo, Polsko) (JR)

(27) válka **Německa s Polskem** / proti Polsku (JR)

(28) **Německo** vede válku **s Polskem** / proti Polsku (JR)

Výjimkou jsou spojení, která žádné participanty nemají – neosobní konstrukce a takové konstrukce, ve kterých je podmětem predikativní jméno (u Macháčkové *subjektový typ*).

Pro tvoření APKS jsou klíčové ty argumenty, které mohou být syntaktickým subjektem a zároveň se v LG kryjí se *sémantickým subjektem*, který odpovídá subjektovému genitivu (původci děje) predikativního jména jako v příkladu (29). Pokud tomu tak není (30), jedná se o predikativní spojení (Radimský 2010, s. 73–74).

(29) Peníze mají vliv na Maxe. (= vliv **peněz** na Maxe) (JR)

(30) Peníze vysvětlují tento vliv na Maxe. (JR)

V pozici syntaktického subjektu se může u tzv. konverzních konstrukcí realizovat i jiný argument (32) (Radimský ho nazývá *sémantický objekt*), který se v češtině nápadně kryje se změnami v aktuálním členění (u věty (31)). V češtině je v některých případech možné vytvořit i vazbu s druhým objektovým argumentem (33), který v tomto případě odpovídá objektovému genitivu (jako alternativa k dativu) (Radimský 2010, s. 75–78).

(31) **Lukáš** učinil **Marii** nabídku k sňatku. (JR)

(32) **Marie** dostala od **Lukáše** nabídku k sňatku. (JR)

(33) **Sňatek** byl předmětem **Lukášovy** nabídky **Marii**. (JR)

Kolářová (2010, s. 116) na základě analýzy deverbativních substantiv tvořících APKS v PDT usuzuje, že mají téměř vždy svou vlastní valenci (u trojvalenčních struktur), a to ve formě prostého pádu (*provést opravu něčeho, budit pocit něčeho*), předložkové skupiny (*mít rozhovor s někým, vzbudit zájem o něco*) nebo infinitivu či vedlejší věty (*vytvořit názor, že..., vydat pokyn + infinitiv*). Pokud jde o nedeverbativní substantiva, Kolářová uvádí, že mají často valenci

získanou od slov, od kterých byla odvozena (to platí zejména pro deverbativní adjektiva), a že u primárních substantiv jde o získání valence analogií (*věrnost, oddanost někomu, zodpovědnost za něco*).

Substantivum, které vstoupí do APKS, si může podle Macháčkové (1983, s. 68–69, 136) zachovat svou valenci (*chovat úctu ke komu*), nebo převzít valenci slovesa (u trojvalenčních sloves jako *dát, poskytovat, vzdát, projevit*; srov. *důvěra ke komu*, ale *projevit důvěru komu*). Pokud je jediným valenčním doplněním kategoriálního slovesa predikativní jméno (a podmět), je realizace dalších participantů v režii predikativního jména (*mít zalíbení v kom/čem*). V některých případech si mohou valence konkurovat – *budit obdiv všech, budit Janův obdiv* (přivlastek ovlivněný substantivem), ale *budit v kom (u koho) obdiv* (přísloušné určení místa ovlivněné slovesem).

Macháčková (1983, s. 135), Kolářová (2010, s. 116–118) i Radimský (2010, s. 73–74) poukazují na to, že někdy nemůže být vlastní substantivní doplnění vůbec vyjádřeno (34) a (35), případně bývají doplnění neshodná s valenčními členy kategoriálního slovesa vypuštěna (*údržbář opravil vodovod* ale *údržbář provedl opravu*).

(34) *Projevil Petrovi důvěru **k němu/Petrovi**. (VeK)

(35) *Poskytl Janovi péči **o něj/Jana**. (VeK)

Kromě adresáta nebo pacientu vstupuje do vazby s podstatným jménem i konatel, který bývá totožný s konatelem slovesné části a obvykle není povrchově vyjádřen, nebo dokonce ani vyjádřen být nemůže (36). Jen v některých případech je možné využít přivlastňovacího zájmena *svůj* (37) (Kolářová 2010, s. 116–117; Macháčková 1983, s. 135). Naopak běžně je konatel vyjádřen tam, kde konatel substantiva není totožný s konatelem slovesa (38). V PDT jsou tyto informace reflektovány na tektogramatické rovině z důvodu konzistence slovníku i dat (viz 2.2.3.3.1).

(36) *Jan dostal **Janův** strach. (VeK)

(37) Jan Petrovi projevil **svoji** důvěru. (VeK)

(38) Na konci svého vystoupení chci využít příležitosti a obrátit **vaši** pozornost na osudy oněch jinak smýšlejících lidí v našich zemích, kteří [...]. (VeK, ČNK)

Macháčková (1983, s. 135) také upozorňuje, že se toto valenční doplnění umí stát podmětem věty ve spojení s kategoriálním slovesem, a to stejně jako u základového slovesa (*Zemědělci osévají půdu. – Zemědělci provádějí osev.*).

Kolářová (2010, s. 118–119) poukazuje na skutečnost, že pokud se substantivum osamostatní od kategoriálního slovesa, může přebrat jeho třetí valenční doplnění (substantivum by tuto formu samo nemělo, příp. ji u deverbativ nemá ani základové sloveso). Pokud příklady a pozorování stručně shrnu, ve většině případů se jedná o dativní valenci (*podpořit někoho, vyjádřit podporu někomu* → *podpora někomu; otázat se někoho, dát/položit někomu otázku* → *otázka někomu*), doplnění s předložkou *od* a genitivem (*pacient byl ošetřen od lékaře* → *ošetření pacienta od lékaře*) nebo u spojení předložky *na* a akuzativu většinou patrně o původní vazbu substantiva (*důraz na něco, konkurz na něco*).

Z pozorování Kolářové (2010, s. 120–121) doplněného četnými příklady plyne, že některá valenční doplnění členů APKS mohou být referenčně totožná, resp. že určité valenční doplnění sdílejí (*poskytnout Petrovi pomoc – pomoc Petrovi*), nejčastěji se jedná o konatele, ale může jít i o valenční doplnění s jiným významem (v PDT jsou na tektogramatické rovině zachycena všechna valenční doplnění). Pro sdílená valenční doplnění Kolářová (2010, s. 121) zavádí termín *konkurenční valenční doplnění* a podrobně nastiňuje problematiku neprojektivních konstrukcí v rámci APKS (Kolářová 2010, s. 121–126).

2.3.3 Rozvíjení predikativního jména

Predikativní jméno může být v rámci APKS dále rozvíjeno i jinak než dalšími substantivy v rámci vlastní valence. Většina autorů (např. Čermák 1974, s. 296; Kralčák 2005, s. 34; Macháčková 1983, s. 136; Radimský 2017a) zmiňuje rozvíjení predikativního jména pomocí adjektiv (*chovat nábožnou úctu, poskytnout cennou radu*) v roli přívlastku shodného. Radimský (2010, s. 165) se na základě svých pozorování domnívá, že predikativní jméno lze rozvíjet adjektivy odpovídajícími adverbiím u syntetického ekvivalentu, a to za podmínky, že mezi adjektivem a adverbiem existuje derivační vztah a zároveň se nejedná o frazeologismus (*tupý stríh – *tupě stríhat*). Také si všimá (Radimský 2010, s. 164; 2017b) toho, že adjektivní přívlastek lze u verba dicendi vyjádřit pomocí adverbií:⁵⁴

(39) Ruský parlament **jasně/přesně/jednoznačně** formuloval/vyslovil/*vznesl/*dal/
*předložil požadavek na jednostranné zrušení sankcí vůči Srbsku. (JR)

⁵⁴ Macháčková (1979, s. 64) o adverbiích tvrdí, že rozvíjejí celé APKS. Radimský (2010, s. 164–165) ve svých datech nenachází žádné empirické důkazy pro potvrzení této hypotézy, nicméně potvrzuje neexistenci vazby k predikativnímu jménu (test oddělení jmenové a slovesné části – *rychle dával rozkazy – rozkazy, které rychle dával – *rychle rozkazy, které dával*). Skutečnost, že lze rozvíjet predikativní jména adjektivy, která odpovídají adverbiím u kategoriálního slovesa, ale podle něj poskytuje nepřímý důkaz pro validitu tvrzení jeho předchůdkyně.

(40) Ruský parlament vznesl/dal/formuloval/předložil/přednesl/vyslovil **jasný/přesný/jednoznačný** požadavek na jednostranné zrušení sankcí vůči Srbsku. (JR)

Adjektiva rozvíjející predikativní jméno lze podle mnohých charakterizovat jako hodnotící, kvalifikující, intenzifikační a kvantifikující (srov. Čermák 1974, s. 296; Macháčková 1979, s. 64; Martínek 2008b, s. 139). Martínek (2008b, s. 139) ještě doplňuje, že mohou nést „lexikální význam místo rozvíjeného velmi obecného substantiva (*věc*) nebo konečně tvoří se substantivem víceslovný lexém a přecházejí do terminologické oblasti (*služba boží*)“.

Čermák (1974, s. 296) dodává, že je „[a]djektivní modifikace naproti tomu silně omezená, ne-li nemožná, u substantivních spojení typu: *výroba, výchova, smích, služba, počátek, projev, cit, učení, správa* apod“. Toto pozorování nicméně moje vlastní data nepotvrzují. Např. se spojením *chovat cit(y)* může být použita celá řada adjektiv: *něžné, hluboké, vřelé, silné, extrémní, jednosměrná, podobné, mateřské, stejné, svaté, neopětované, jisté, ryzí, vlastenecké, opravdové, přátelské* atp.

Macháčková (1983, s. 136) i Kralčák (2005, s. 34) dále uvádějí, že v některých případech vůbec není možné adjektivum vynechat (*páchat trestnou činnost – *páchat činnost*) a APKS musí mít tři konstituenty.

Radimský (2010, s. 166) selekci určitého typu adjektiv a adverbii a případně obligatornost některých doplnění chápe jako běžný syntaktický fenomén. Za specifickou vlastnost APKS nicméně považuje transformační souvislost mezi adjektivy a adverbii, která poukazuje na vzájemnou provázanost obou částí APKS.

Mezi řádky autoři zmiňují i další doplnění, např. přivlastňovacím zájmenem (viz předchozí oddíl věnovaný valenci) či zájmenem ukazovacím, nevěnují jim ale žádnou systematickou pozornost.

2.4 Kategoriální sloveso

Jako kategoriální se v češtině může uplatnit jakékoliv sloveso, které v kombinaci s predikativním jménem ztratilo lexikální význam a je sémanticky vyprázdňené (k desémantizaci oddíl 2.4.2). Nejedná se tedy o speciální typ sloves, jako jsou slovesa modální. Jinými slovy predikativní jméno aktivuje kategoriální funkci slovesa (v terminologii LG je kategoriální sloveso jeho *aktualizátorem*), jinak by bylo sloveso interpretováno predikativně (srov. *chovat zášť* a *chovat slepice*). K případné aktivaci kategoriální funkce může dojít dokonce jen v rámci

určitého kontextu.⁵⁵ Každé sloveso může být tedy potenciálně užito i jako sloveso kategoriální (výjimkou by byla slovesa s výhradně kategoriální funkcí jako, např. *(z)tropit, podniknout*) (srov. např. Kolářová 2010, s. 111; Radimský 2017a; 2017b).

Na rozdíl od francouzštiny, angličtiny a dalších jazyků, které mají jen omezené množství kategoriálních sloves (s obecným významem *dělat, mít, dát, brát* atp.), je u češtiny a dalších slovanských jazyků paleta kategoriálních sloves mnohem pestřejší a omezený je naopak okruh predikativních jmen (srov. např. Cinková 2009; Radimský 2010, s. 15).

V následujících oddílech shrnu základní poznatky o variabilitě českých kategoriálních sloves, zaměřím se na jejich typologii a problematiku desémantizace a zastavím se u jejich valence.

2.4.1 Variabilita kategoriálních sloves

Radimský (2010, s. 107, 122–124) dochází přizpůsobením lexikogrammatických kritérií češtině k následujícím vlastnostem ovlivňujícím variabilitu kategoriálních sloves v APKS:

- (a) **syntaktická realizace argumentů predikativního jména** (viz 2.4.2 Typologie) – tato realizace vychází z toho, že přinejmenším některé argumenty predikativního jména jsou realizovány jako valenční
 - (a1) varianty „diateze“ (realizace syntaktického subjektu) – základní, konverzní, nenominativní, kauzativní
 - (a2) realizace ostatních argumentů – fakultativní, obligatorní
- (b) **sémantický přínos kategoriálního slovesa ve smyslu stylové expresivity a povahodějového významu** (Radimský 2010, s. 107–110) (viz 2.2.1 Stylistická a komunikační rovina) – sémantický přínos by neměl být v ideálním případě žádný, ale specifickější slovesa mohou měnit povahu děje a ve spojení s rezultativy nemusí být desémantizace úplná
 - (b1) povahodějové varianty
 - i. fáze děje – durativnost, inchoativnost, terminativnost ...
 - ii. kvantita – intenzita, iterativnost, úsilnost...

⁵⁵ Jak zmiňuje Radimský (2010, s. 157–171; 2017a) v návaznosti na E. Macháčkovou (1979, s. 63–64), otázky směřující na predikativní jméno jako podmět nebo předmět slovesa (*Co vůči vám chová? Nedůvěru.*) jsou samy o sobě obtížně přijatelné, protože existuje silná vazba mezi oběma členy spojení. Pokud ale použijeme sloveso s primárně kategoriálním čtením, k problému s interpretovatelností nedochází (*Co to tam tropí? Hlouposti.*). Kategoriální interpretaci můžeme také aktivovat upravením věty a přidáním predikativního jména do původní otázky (*Chová vůči vám nedůvěru? Ne, spíše nenávisť.*). Stejným způsobem funguje i pronominalizace predikativního jména (*Jan vůči nim choval to. – Často zmiňoval nedůvěru. Tu vůči nim Jan choval;* vlastní příklad).

- (b2) skalární desémantizace – kompletní, částečná (kategoriální a predikativní)
- (c) **míra kolokability kategoriálního slovesa s predikativními jmény** (viz 2.4.3 Desémantizace) – desémantizace nemusí být úplná, a proto ovlivňuje i kolokabilitu a substituovatelnost
- (c1) = (b2)
- (c2) sémantická restrikce – bezpříznakové, příznakové
- (c3) stylistické varianty

2.4.2 Typologie

Radimský (2010, s. 112–115; 2017a) v návaznosti na Lexikogramatickou školu (zejm. Gross 1989) a DVS (Macháčková 1979; 1983) navrhuje pro češtinu typologii APKS i kategoriálních sloves z pohledu od predikativního jména ke kategoriálnímu slovesu. Na tomto rozdělení závisí i jeden ze zmíněných testů pro určení APKS – test koreference argumentů (viz oddíl 2.1.2.2) – který je (na povrchové rovině) uplatnitelný pouze pro APKS základního typu.

Základem rozdělení APKS je předpoklad, že pro každé predikativní jméno může existovat více kategoriálních sloves, která se od sebe liší podle toho, jakou syntaktickou roli má predikativní jméno, resp. které doplnění je v rámci APKS syntaktickým subjektem. Získáme tak čtyři typy konstrukcí níže (příklady převzaty z citované práce Radimského, tučně je vyznačen podmět).⁵⁶

- (a) základní konstrukce – v pozici syntaktického subjektu je subjektový argument; ve jmenném syntagmatu mu odpovídá genitiv (*požadavek vlády*) nebo u silně tranzitivních jmen instrumentál (*poškození přístroje Petrem*) – ***Vláda předložila sněmovně požadavek na navýšení rozpočtu.***
- (b) konverzní konstrukce – syntaktickým subjektem je jeden z objektových argumentů – ***Sněmovna obdržela od vlády požadavek na navýšení rozpočtu. Navýšení rozpočtu bylo požadavkem vlády (ve sněmovně).***
- (c) kauzativní/faktitivní konstrukce – v pozici syntaktického subjektu je kauzátor, který ale není doplněním predikativního jména – ***Tato situace vyvolala požadavek vlády (ve sněmovně) na navýšení rozpočtu.***

⁵⁶ Distribuci sloves podle jednotlivých konstrukcí zkoumá Radimský na vzorku kategoriálních sloves pojících se 70 vybranými frekventovanými jmény (viz Radimský 2010, s. 124–153).

- (d) nenominativní konstrukce – pozice syntaktického subjektu není realizována vůbec – *Vládou byl sněmovně předložen požadavek na navýšení rozpočtu. Objevil se požadavek vlády na navýšení rozpočtu (ve sněmovně).* Tato konstrukce může být tvořena neosobním slovesem (*Došlo k navýšení rozpočtu.*), použitím kategoriálního slovesa s predikativním jménem na místě podmětu (*Navýšení rozpočtu proběhlo hladce.*) nebo pasivem základního kategoriálního slovesa v případě akuzativního predikativního jména (*Navýšení rozpočtu bylo provedeno okamžitě.*). Patří sem také zvrtná slovesa, která v nezvratné formě kategoriální funkci nemají (*vynořila se otázka – *vynořit otázku*).

Podobné pozorování ohledně syntaktického subjektu má i Kettnerová (2017, s. 12–21), tedy kromě nenominativních struktur, protože se podobným strukturám nevěnuje. Díky změně kategoriálního slovesa se na pozici syntaktického subjektu může vyjádřit jiný sémantický participant, APKS jsou proto jedním z prostředků hierarchizace větné struktury. Z povrchověsyntaktického hlediska Kettnerová vysvětluje hierarchizaci skutečností, že v koreferenci konatele je pokaždé jiné ze jmenných doplnění, a také možnou kauzativností. Navrhuje proto vlastní typologii APKS založenou právě na těchto dvou rysech: na vyšší rovině kauzaci (KAUZ) a na nižší konverzi (KONV), viz Tabulka 3. Rys kauzace vyjadřuje (ne)přítomnost kauzátora jakožto sémantického participantu a konverze zohledňuje koreferenci konatele slovesa.

Tabulka níže srovnává typologii navrženou Radimským (viz výše) a rozdělení zmíněných rysů kauzace-konverze podle Kettnerové. Příklady jsou zkrácenou variací těch, která uvádí ve své studii Kettnerová.

Radimský	Kettnerová			příklad
typ konstrukce	KAUZ	KONV	koreference	
základní	-	-	konatel kategoriálního slovesa ↔ konatel predikativního jména (v aktivní konstrukci)	(dát slib) <i>Petr dal slib mamince, že nepřijde pozdě.</i>
konverzní	-	+	konatel kategoriálního slovesa ↔ doplnění predikativního jména	(dostat slib) <i>Maminka dostala slib od Petra, že nepřijde pozdě.</i>
kauzativní	+	-	kauzátor ↔ konatel kategoriálního slovesa	(dát možnost) <i>Úřad dal občanům možnost vyjádřit se ke stavbě.</i>
	+	+	kauzátor ↔ doplnění kategoriálního slovesa	(dostat možnost) <i>Občané dostali od úřadu možnost vyjádřit se ke stavbě.</i>

Tabulka 3: Srovnání typologií kategoriálního slovesa navržených Radimským a Kettnerovou

2.4.3 Desémantizace

Zásadním rysem kategoriálních sloves v APKS je desémantizace, tj. ztráta významu (kategoriální slovesa jsou označována také jako oslabená, významově vyprázdněná, depletivní atp.), a tím částečný posun směrem k sysémantikům (odtud anglický termín *light verb*). Díky desémantizaci vytvářejí kategoriální slovesa ve spojení s jedním predikativním jménem i poměrně nečekané synonymní řady (*dát/vytnout/vrazit/vlepit/uštedřit facku*), přičemž některá synonymická paradigmata mohou být blokována (*mít námitku* – **vlastnit/držet námitku*). V případě APKS se ale nejedná o úplnou desémantizaci⁵⁷ – sloveso kromě aktualizačních kategorií vyjadřuje i obecné významové rysy: význam a změnu stavu nebo způsobení změny stavu (viz sémantické třídění APKS v oddíle 2.2.2). V některých případech zůstávají zachovány i povahodějové rysy jako fázovost a kvantita nebo expresivnost (viz 2.2.1) (srov. např. Čermák 1974, s. 300; Macháčková 1983, s. 134; Radimský 2010, s. 181–124, 154; 2017b). Martínek (2008b, s. 27–28) dokonce navrhuje namísto o desémantizaci implikující „neplnohodnotnost“ hovořit spíše o „přenesení významu a zachování obecnějších sémantických rysů“ nebo o metaforičnosti spojení (ilustruje ji na spojeních *zahrnout výčítkami*, *obklopit péčí*, *chovat nedůvěru*).

⁵⁷ Srov. sloveso *být* jako pomocné sloveso, nebo např. význam předložky ve valenčních vazbách (*myslí na maminku* – *sedá na maminku*) (Hladká 2017).

Pro testování desémantizace kategoriálních sloves můžeme použít redukční test (viz 2.1.2.1), při kterém po odstranění kategoriálního slovesa syntagma ztratí jen své povahové rysy, nikoliv význam.

Vyprázdňenost kategoriálního slovesa se projevuje i při překladu z jednoho jazyka do druhého, jak uvádí ve svých studiích Radimský (2010, s. 55; 2017a),⁵⁸ a to hned dvěma způsoby. Při překladu je někdy možné kategoriální sloveso bez náhrady vypustit (srov. *Not all the decisions he made were wise.* – *Ne všechna jeho rozhodnutí byla moudrá*). Pokud se naopak kategoriální sloveso zachová i v překladu, je třeba ho vybrat s ohledem na použité predikativní jméno (srov. *If you want to ask any further questions...* – *Chcete-li položit nějaké další otázky...*).

Radimský (2010, s. 118–122; 2017b) se vymezuje proti lexikogramatickému pojetí desémantizace jako binární opozice mezi kategoriálním a predikativním užitím slovesa a dokazuje, že tato opozice neplatí ve všech kontextech. Ilustruje to na příkladu níže, kdy pro větu z novinového článku z ČNK (původně se slovesem *vznesl*) doplňuje další kategoriální slovesa, která jsou v rámci daného kontextu navzájem zaměnitelná (viz (41) a (42), snad jen *vyrukoval* má rys náhlého děje s „militantním“ naladěním činitele):

(41) Ruský parlament **vznesl/dal/formuloval/předložil/vyslovil/vznesl požadavek** na jednostranné zrušení sankcí proti Srbsku [...]. (JR)

(42) Ruský parlament **přišel/vyrukoval s požadavkem** na jednostranné zrušení sankcí proti Srbsku [...]. (JR)

Zmíněná slovesa lze zařadit do různých větných vzorců: slovesa transmise (*dát, předložit*), verba dicendi (*formulovat, vyslovit, přednést*) a pohybová slovesa (*přijít, vyrukovat*). Metaforicky je možné nalézt opodstatnění pro volbu jednotlivého slovesa, ale v rámci kontextu uvedené věty jsou slovesa synonymní. Radimský poukazuje na dvě protichůdné tendence při výběru kategoriálního slovesa – ten je motivován výchozím větným vzorcem (příp. vnější motivací jako kalk z cizího jazyka), a zároveň je arbitrární (pokud můžeme sloveso při transformaci odstranit beze změny významu, nezáleží až tolik na tom, které konkrétně uijeme – což je velký přínos pro stylistiku) (Radimský 2010, s. 118–119).

Podstata skalární desémantizace spočívá v tom, že může existovat kontext, v němž nelze použít všechna slovesa z daného rámce – srov. příklady (43) a (44). V prvním případě je možné

⁵⁸ Problematice překládání APKS do francouzštiny a češtiny v právních textech Evropské unie se věnuje M. Povolná (2012).

použít libovolné sloveso, zatímco v druhém pouze verba dicendi – kontext vyžaduje predikativní jméno s aspektem komunikace a sloveso má tedy omezenou substituovatelnost.

(43) Před pár týdnů Státní ústav pro kontrolu léčiv **přednesl** svůj finální požadavek: ke konečnému schválení léku dodat studii o jeho trvanlivosti [...]. (JR, ČNK)

(44) Živě si představil, jako Bob vytáčí číslo a bez slůvka vysvětlení **přednáší** svůj požadavek. (JR, ČNK)

Radimský z analyzovaných příkladů nakonec vytváří čtyřstupňovou škálu charakterizující skalární desémantizaci slovesa. Jednotlivá spojení tvoří kontinuální přechod mezi predikativním a kategoriálním významem, neboť v jistých kontextech může být desémantizace dokonce blokována např. porušenou koreferencí (ač k ní dochází zřídka). Tuto škálu Radimský ilustruje na konstrukci *přednést/zaslat požadavek* v tabulce níže:

funkce slovesa		příklad
kategoriální	kategoriální	Před pár týdnů Státní ústav pro kontrolu léčiv přednesl svůj finální požadavek [...].
	kategoriální i predikativní, omezená substituovatelnost	Živě si představil, jak Bob vytáčí číslo a bez slůvka vysvětlení přednáší svůj požadavek .
↓	kategoriální i predikativní, velmi omezená substituovatelnost	Máte-li všechny vynaložené náklady spočítané, zašlete svůj požadavek majiteli obchodu co nejdříve.
↑	predikativní	Starosta přednesl požadavek Technických služeb o zvýšení cen za prodej hrobového místa.

Tabulka 4: Škála skalární desémantizace kategoriálního slovesa podle Radimského
Zdroj: Radimský (2010, s. 121)

V příkladu z prvního řádku lze substituovat sloveso *přednést* libovolným dalším z příkladu (41), protože je desémantizované zcela. Ve dvou následujících případech jsou všechny formální požadavky na APKS splněné, ale *přednáší* má význam ústního projevu a *zašlete* naopak písemné komunikace, jsou tedy hůře nahraditelné.⁵⁹ V posledním příkladu přidáním neshodného přísudku *Technických služeb* dojde k porušení koreference (*požadavek* již není starostův) a sloveso má proto predikativní funkci.

V souladu s podstatou skalární desémantizace je i rozdělení sloves podle Macháčkové (1979, s. 57; 1983, s. 129, 134) na slovesa s obecným významem (*mít, být, dát*), slovesa

⁵⁹ Omezená substituovatelnost, resp. sémantická restrikce, může být zjevnější u spojení se slovesem *spáchat*, které nese negativní konotace: *spáchat zločin – *spáchat dobrý skutek* (Radimský 2010, s. 111–112).

částečně desémantizovaná (*přijít do rozpaků*) a slovesa užitá metaforicky (*zahrnout/zasypat výčitkami*).

Generická slovesa *provést/provádět* jsou sice vysoce kolokabilní, ale z hlediska frekvence v úzu méně preferovaná; jediným frekventovaným a zároveň „univerzálním“ slovesem je sloveso *mít* (Radimský 2010, s. 154). Za nejméně příznaková kategoriální slovesa pro dané predikativní jméno považuje Radimský (2010, s. 65, 106) ta, se kterými toto jméno tvoří APKS (*pokuta – uložit, porážka – ušetřit*). Kolokability sloves se samozřejmě dotýkají i starší práce (např. Čermák 1974, s. 297–299; Macháčková 1983, s. 134; Vlková 1990, s. 3–4), ale bez přístupu ke korpusovým datům.

2.4.4 Valence

Jak již bylo zmíněno, kategoriální slovesa jsou nositeli gramatických kategorií (stejně jako běžná plnovýznamová slovesa), nicméně mají jen strukturní význam (jejich valenční rámec tvoří syntaktickou strukturu věty, ale neodpovídá situačním doplněním) a nenesou žádné lexikálněsémantické charakteristiky. Situační význam a participanty do spojení přináší až predikativní jméno (Kettnerová a Lopatková 2017, s. 4–5, 8). Z formální stránky se kategoriální slovesa chovají stejně jako běžné konstrukce s plnovýznamovými slovesy a jejich sémantické participanty jsou na povrchu vyjádřené pouze jedenkrát – až na řídké výjimky s přivlastňovacími zájmeny (Kettnerová et al. 2018).

V tomto oddíle se věnuji různým úhlům pohledu na valenci kategoriálních sloves. Jedná se zejména o problematiku koreference, sémantické a syntaktické řízenosti, pozici umístění predikativního jména a fungování třetího valenčního doplnění v rámci APKS. Valenčnímu doplnění, které do APKS vnáší predikativní jméno, byl věnován oddíl 2.4.2.

2.4.4.1 Koreference a mapování valenčních struktur

Radimský s koreferencí operuje zejména v kontextu testů pro určování APKS (2.1.2). Mnohem podrobnější rozbor provádí studie *Ke koreferenci u komplexních predikátů s kategoriálním slovesem* (Kettnerová a Lopatková 2017), kterou v tomto pododdíle stručně shrnuji. Podrobněji k syntaktickému chování na hloubkové a povrchovésyntaktické rovině viz Kettnerová (2017); tématu jsem se také částečně dotkla v poddíle 2.2.3.3.2.

Studie (srov. Kettnerová 2017, s. 8–11; Kettnerová a Lopatková 2017, s. 7–9) operují se dvěma pojmy – tzv. situačním a strukturním významem. První zmíněný význam autorky popisují jako abstraktní situaci s určitým počtem situačních participantů charakterizovaných typem a vzájemnými vztahy; tj. participantům jsou přisouzeny sémantické role. Oproti tomu

strukturní význam je gramatická část aktantu, kterou lze zachytit valenčními rámci (v FGP odpovídající tektogramatické rovině). Oba druhy významu společně utvářejí celý význam APKS. Kategoriální sloveso má na rozdíl od predikativního jména jen strukturní význam (výjimkou je kauzátor), ale díky koreferenčním vztahům (odkazům ke stejným situačním participantům) získává i význam lexikálněsémantický.

Ve studii autorky popisují spojení valenčních rámců (již zmiňované v 2.2.3.3.2) zjednodušeně tak, že ve valenčním rámci kategoriálního slovesa zaujme jednu pozici predikativní jméno (v FGP reprezentované funktorem CPHR, obvykle *patiens*), zbylé pozice mohou díky koreferenci lexikálněsémanticky specifikovat jeho doplnění. Pokud se pozice nevyužije, je dané doplnění zredukováno. Např. spojení *dát radu* se skládá z rámců substantiva *rada* (ACT ↔ mluvčí, ADDR ↔ recipient, PAT ↔ informace) a kategoriálního slovesa *dát* (CPHR = predikativní jméno, ACT, ADDR). Ke stejnému situačnímu participantu koreferuje ACT i ADDR predikativního jména i kategoriálního slovesa. Pokud ACT predikativního jména není koreferenční s žádným doplněním slovesa, nejedná se o APKS, ale o kolokaci predikativního slovesa s predikativním jménem. Totéž ukazuje Radimský u testu koreference (viz 2.1.2.2).

Na povrchověsyntaktické rovině jsou (při zanedbání elipsy, nevyjádření podmětu na povrchu a příp. nepřítomnosti fakultativních doplnění) vyjádřena všechna doplnění kategoriálního slovesa (a)–(c) a z rámce predikativního jména se může uplatnit (d) (Kettnerová 2017, s. 11–12; Kettnerová a Lopatková 2017, s. 10–11).⁶⁰

- (a) valenční doplnění CPHR lexikálně vyjádřené predikativním jménem
- (b) valenční doplnění korespondující s kauzátorem (jde-li o kauzativní kategoriální sloveso)
- (c) ostatní valenční doplnění kategoriálního slovesa, neboť všechna získávají (skrze koreferenci s valenčními doplněními predikativního jména) lexikálněsémantické naplnění
- (d) ta doplnění, která nejsou v koreferenčním vztahu se žádným doplněním slovesa⁶¹

Autorky tuto skutečnost ilustrují na následujícím příkladu:

⁶⁰ Definice (a)–(d) přebírám v původním znění (shodně Kettnerová 2017, s. 10–11; Kettnerová a Lopatková 2017, s. 10).

⁶¹ Jedinou řídkou se vyskytující výjimkou je souběžné vyjádření konatele slovesného (*nikdo*) i jmenného (*svůj*) ve větách jako *Nikdo nevěnuje lásce celý svůj čas* (Kettnerová a Lopatková 2017, s. 12); dále viz 2.5.

(45) Farář_{ACT-verb} pak dal rodině_{ADDR-verb} radu_{CPHR-verb}, (aby svěcenou křídou udělali okolo chalupy čáru a modlili se růženec)_{PAT-subst} (SYN2013pub)

Specificky se chová skupina kauzativních sloves, kdy jedno z doplnění slovesa slouží jako iniciátor děje – kauzátor. Ten není součástí predikativního jména, ale je sémantickým jádrem celého APKS (Kettnerová a Lopatková 2017, s. 11–12; ve shodě s Radimský 2010, s. 79–80). Viz také typologie podle Kettnerové (2017) v oddíle 2.4.2.

U některých nekauzativních APKS může dojít k porušení principu (d) a na povrchu je vyjádřen konatel predikativního jména i kategoriálního slovesa, přestože referují ke stejnému sémantickému participantu. Doplnění predikativního jména je v tomto případě vyjádřeno přivlastňovacím zájmenem (srov. (46) a (47)) (Kettnerová 2017, s. 21).

(46) Musíme počkat, až **skončí** své_{N-ACT} **vyšetřování** orgány_{V-ACT} činné v trestním řízení. (VaK, SYN2013pub)

(47) Skoro tři měsíce uplynuly od doby, kdy hráči předložili svůj poslední návrh, a ještě jsme **nedostali** od vedení ligy_{V-ORIG} její_{N-ACT} **protinávrrhy**. (VaK, SYN2009pub)

Kettnerová a Lopatková (2017, s. 12–18) na základě korpusového jazykového materiálu rozdělují typy koreference do 16 skupin podle toho, které aktanty jsou propojeny. Ze spojení s nekauzativními slovesy jsou ve zkoumaných datech nejzastoupenější slovesa s dvouvalečným doplněním ACT + CPHR (*činit, konat, podniknout, pojmout*) a dále se třemi doplněními ACT + ADDR + CPHR (*dát, poskytnout, projevit, uložit*) nebo s ORIG na místě ADDR (*brát (si), dostat, přijmout*). Podrobnější popis a přehledné tabulky viz studie citovaná výše.

2.4.4.2 Sémantická vs. syntaktická řízenost APKS

Na valenci kategoriálního slovesa se dá pohlížet z různých stran, nejvíce ji ale ovlivňuje problematika syntaktické a sémantické řízenosti. Predikativní jméno je sémanticky řídicím prvkem celého spojení (připisuje argumentům sémantické role, vybírá si lexikálně vyprázdňené sloveso a aktivuje jeho kategoriální funkci), ale stále je syntakticky podřízeným doplněním kategoriálního slovesa (aktantem). Rozpor mezi syntaktickou a sémantickou řízeností může mít v různých přístupech také vliv na to, která složka APKS se považuje za predikát; u formálních syntaxí je za něj považováno kategoriální sloveso (zápisy v PDT (2.2.3.3.1) i např. Čermák (1974)), u sémantického přístupu naopak kategoriální jméno (např. DVS i LG, více oddíly 2.2.3.1 a 2.2.3.2). Formálně se navíc nejedná o jeden větný člen; je možné ho rozdělit do dvou klauzí (*otázka, kterou mi včera položil*), a dokonce může být i podmětem věty (*byla provedena kontrola*). Radimský (2010, s. 18–20, 55, 178, 180; 2017a) proto navrhuje pohlížet na APKS jako

na duální jednotku v *langue*, která „umožňuje větné jádro, prototypicky vyjadřované slovesem, v případě potřeby rozštěpit na složku sémantickou a syntaktickou“ (Radimský 2017a).

Zmíněný rozpor ovlivňuje i analýzu těchto spojení. Je možné postupovat od predikativního jména ke kategoriálnímu slovesu (a např. zkoumat, jaká kategoriální slovesa se pojí s jedním predikativním jménem; *cit – chovat, probouzet/probudit, projevit/projevovat, vzbouzet/vzbudit/vzbuzovat, získat/získávat, ztrácet/ztratit*), anebo opačně od kategoriálních sloves k predikativním jménům (a zkoumat např. vlastnosti, chování i kolokabilitu jednotlivých kategoriálních sloves; *chovat – cit, naději, náklonnost, nedůvěru, nenávisť, obdiv, podezření, přátelství, respekt, sympatie, úctu, zášť* atd.) (srov. Radimský 2010; 2017a).

Stejně jako v PDT (viz 2.2.3.3.1) budu v empirické části této disertační práci považovat z hlediska formální syntaxe za predikát kategoriální sloveso, predikativní jméno pokládám za jedno z jeho valenčních doplnění. Protože je mým cílem řešit mj. pádovou homonymii, a tudíž je vhodné pracovat s APKS se stejným valenčním rámcem, postupuji od kategoriálního slovesa směrem k predikativnímu jménu. Ze sémantického hlediska ale považuji za nutné počítat s celým spojením.

2.4.4.3 Pozice predikativního jména a jeho pád

Někteří autoři (srov. Jelínek 2003, s. 40–41; Kolářová 2010, s. 111–112; Kralčák 2005, s. 34–35; Macháčková 1979, s. 51–55; Martínek 2008b, s. 10) APKS rozdělují podle toho, je-li predikativní jméno na pozici syntaktického objektu (*chovat úctu, klást otázky*), nebo subjektu (*popadl ho vztek, zmocnil se ho strach*), podle této charakteristiky APKS dělí na subjektový a objektový typ,⁶² případně o něm mluví jako o pravo- a levovalenčním doplnění kategoriálního slovesa. Subjektovému typu se většina z autorů (s výjimkou Macháčkové) nevěnuje a místo toho se soustředí na nejrozsáhlejší skupinu APKS s predikativním jménem v bezpředložkovém akuzativu.

Jak již bylo řečeno, nejčastější je u objektového typu APKS predikativní jméno v bezpředložkovém **akuzativu** (*chovat úctu, klást otázky*), příp. s předložkou (*brát v úvahu, vzít na vědomí*). Vyskytují se ale i APKS s ostatními pády: **genitivem** (*doznat nápravy, upadnout do rozpaků, dát se do pláče*) – kromě knižních vyjádření také velice často ve spojitosti se zvrátnými slovesy, **dativem** (*podrobit analýze, přistoupit k vyhlášení výsledků*), velmi zřídka **lokálem** (*získat na síle, udržet v paměti*), **instrumentálem** (*kypět zdravím, hořet touhou, přijít s nápadem udělat něco*) (srov. Čermák 1974, s. 292–293; Jelínek 2007, s. 43; Kolářová 2010, s.

⁶² Pozor ale na záměnu u Kralčáka (2005, s. 73–80), který uvádí čtyři typy APKS (S-V-O(-O), S-V, V(-O) a V) a chápe subjektový/objektový typ jako spojení, které sice má subjekt/objekt, ale na jeho pozici nemusí být nutně predikativní jméno.

111–112; Macháčková 1979, s. 110–193; 1983, s. 138–140; Mikulová et al. 2005, s. 332; Vlková 1990, s. 5).

Macháčková (1979, s. 68) poukazuje na to, že v případě substantiva v instrumentálu jde často původně o příslovečné určení příčiny (*hořet nenávistí, pukat zlostí*) anebo příslovečné určení prostředku (*naplnit koho žalem, zasypat koho otázkami*), které přechází v objekt. Stejně tak u předložkových pádů u sloves pohybu (*přijít do rozpaků, dospět k rozhodnutí*), kdy se jedná vlastně o příslovečné určení.

2.4.4.4 Třetí valenční doplnění

Kategoriální sloveso nemusí mít pouze dvě valenční doplnění (zpravidla bývá v subjektu agens a druhou valenci obsadí predikativní jméno) nebo doplnění jedno, kdy je predikativní jméno zároveň subjektem (*kontrola byla provedena*). V mnoha případech má doplnění tři – poslední z nich zahrnuje adresáta, recipienta nebo zdroj, stimul nebo cíl děje/stavu (*komise poskytla radu tazatelům*) (srov. Čermák 1974, s. 292–293; Kolářová 2010, s. 113; Macháčková 1983, s. 139).

Na třetí pozici APKS s bezpředložkovým akuzativem (na datech PDT) se nejčastěji vyskytují následující formy (Kolářová 2010, s. 113–114; shodně Macháčková 1983, s. 156):

- dativní doplnění (*dát někomu možnost, poskytnout někomu příležitost*) – obvykle aktivní konstrukce
- předložka *od* + gen (*dostat od někoho úkol/pokutu, získat od někoho právo k něčemu*) – obvykle pasivní vyjádření dativní konstrukce, na povrchové rovině věty často vypuštěno
- předložka *z* + gen (*nabýt z něčeho dojem, udělat z něčeho závěr*)
- předložka *na* + acc (*klást na někoho nároky, obrátit pozornost na něco, uvalit vazbu na někoho*) – typická u sloves, která mají v nepřeneseném významu směrové určení
- předložka *v* + loc nebo *u* + gen (*budit v někom nepříjemný pocit, vzbuzovat u někoho pochybnosti*)

Kolářová (2010, s. 114–115, 120–121)⁶³ popisuje valenci kategoriálních sloves i z pohledu změn oproti bezpříznakovému užití téhož slovesa. Rozlišuje následující možnosti valenčního doplnění (na základě analýzy APKS s bezpředložkovým akuzativem):

⁶³ Kolářová se problematice valence deverbativních substantiv specificky věnuje v celé knize, proto při zpracování této části vycházím z jejích poznatků. Ostatní autoři s ní nejsou v rozporu. Kralčákovo (2005, s. 84–85) dělení vykazuje jisté nesrovnalosti, na které upozorňuje Martínek (2011, s. 127), nicméně i tak se domnívám, že u něj panuje hrubá shoda s Kolářovou, přestože Kralčák ještě sporně přidává analogii k valenci slovesa plnovýznamového.

- (a) zachování stejného valenčního rámce, který má sloveso v bezpříznakovém použití
- (b) získání vazby při vytvoření APKS konstrukce
- (c) ztráta valence při vytvoření APKS konstrukce

Změnu ve valenčním rámci oproti bezpříznakovému použití slovesa (varianty (b) a (c)) považuje Kolářová (2010, s. 114) za důvod, proč celou konstrukci případně považovat za APKS, i když nejsou zcela splněna všechna ostatní kritéria.

K **získání nové vazby** (b), přestože pro ni v predikativním použití není žádný důvod, dochází patrně na základě analogie. První možností je přijetí valence podle vzoru jednoslovného synonymního predikátu – Kolářová uvádí, že se to týká zejména sloves *dělat, udělat, činit, učinit, tvořit, vytvořit*, která tvoří APKS jako *(u)dělat/učinit na někoho dojem* (patrně analogie k *působit na někoho dojem* anebo *zapůsobit na někoho*), *udělat něčemu konec/přítrž* (analogie k *zabránit čemu*). Druhou alternativou je přijetí valence jiného slovesa, které do APKS vstupuje, jako *položít/klást někomu otázku* (analogie k *dát někomu otázku* vs. *otázet se koho*) nebo *vyjádřit někomu úctu/podporu* (analogicky k *projevit někomu úctu* vs. *vážít si koho/čeho, uctívat koho/co*). Bez vyjádření predikativního jména pak takové spojení dodatečné doplnění mít nemůže:

(48) Otázku, kterou **mi** včera položil, jsem nechtěla slyšet. (vI)

(49) Co ***ti** včera položil? (vI)

V některých případech může valenci ovlivňovat i predikativní jméno (zejm. v případě, že samo sloveso má pouze valenční doplnění obsazené predikativním jménem), se kterým sloveso APKS vytváří – zejm. doplnění pomocí předložkových pádů, jako *mít zájem o něco, mít obavu o někoho*, nebo genitivně vyjádřené doplnění jako *dělat rekonstrukci bytu*. V obou případech je zřejmé, že se jedná o valenci substantiva, které tento valenční potenciál má. Sporná situace může nastat u bezpředložkového dativu (*klást/tvořit/vytvářet/stavět něčemu překážku*), kde může jít jak o analogii k jinému slovesu (*bránit čemu*), tak o vliv predikativního jména (sloveso *překážet*, ze kterého bylo substantivum *překážka* utvořeno, dativní valenci má). Kolářová (2010, s. 115) i Cinková (2009, s. 44) považují v takových případech valenční doplnění substantiva za silnější, a proto ho považují za valenci jmenné části.⁶⁴ Naopak podle studií syntaxe

⁶⁴ Možná opačný postoj má Macháčková (1983, s. 136), která uvádí, že má-li doplnění formu příslovečného určení místa, je řízeno slovesem (srov. *budit Janův obdiv – budit obdiv u Jana*). Cinková (2009, s. 44) ale podotýká, že uvedený případ je nejednoznačný, protože není jasné, zda jde o konstrukci *budit Janův obdiv pro něco* nebo *budit něčí obdiv pro Jana*.

v PDT se zdá, že ve většině případů se povrchově vyjadřuje doplnění kategoriálního slovesa (např. Kettnerová 2017, s. 12–14; Kettnerová et al. 2018, s. 42–44).

U některých sloves může naopak docházet ke **ztrátě valenčního doplnění (c)** – *podat výkon* (chybí doplnění s významem adresáta), *dostat chuť/nápad* (chybí doplnění s významem origa).

Problematika **konkurence mezi valenčními doplněními** obou složek APKS je poměrně zapeklitá a rozlišení nebývá vždy zcela jednoznačné (Kolářová 2010, s. 115, 126). U desémantizovanějších sloves je možné mluvit o ovlivnění valencí samotného predikativního jména. Jednoznačné jsou případy, kdy je doplnění vyjádřeno předložkovým pádem (*mít zájem o něco* – *zájem o něco*) nebo genitivem (*dělat rekonstrukci bytu*), detailněji (2.3.2). U spojení s instrumentálem Kolářová (2010, s. 115) navrhuje brát v úvahu slovosled (jednoznačné *vyjádřit pohrdání něčím*, ale srov. *udělal pohyb pravou rukou* (= pohyb rukou) – *udělal tou rukou* (= udělat pomocí ruky)). Sporné mohou být podle Kolářové případy s dativem, kdy se může jednat jak o valenci predikativního jména, tak o valenci kategoriálního slovesa získanou na základě analogie (*klást/tvořit/vytvářet/stavět něčemu překážku/bariéru* – může jít o analogii k *bránit čemu*, ale také *překážet komu*); Kolářová nicméně považuje valenční vztah predikativního jména za silnější, takže neurčitelná doplnění připisuje jmenné části.

2.5 APKS jako analytický predikát

Pokud APKS nepovažujeme za volné syntaktické spojení slovesa s predikativním jménem (přestože formálně vypadá jako pravidelné spojení slovesa a předmětu), můžeme k němu přistupovat jako k frazeologické kolokační jednotce (viz 2.2.2), anebo ho považovat za specifické syntaktické spojení, které se odlišuje svými speciálními rysy (srov. Macháčková 1979, s. 60–66; 1983, s. 133–137; Radimský 2010, s. 155). Rysy specifické pro celé spojení shrnu na základě přehledu sestaveného Radimským (2010, s. 155–169), který dle potřeby doplňují poznatky z další literatury.

K APKS často existuje **syntetické slovesné synonymum** (*provést analýzu* – *analyzovat*), čehož se dříve využívalo jako kritéria pro určování spojení (obzvláště v souvislosti s multiverbizací – viz 2.2.2.2). Dnes tento faktor autoři považují spíše za vodítko pro rychlou orientaci (Kolářová 2010, s. 109, 111; Macháčková 1979, s. 51–52,61; Radimský 2010, s. 156–157), které ale, jak podotýká Radimský (2010, s. 157), „o povaze jednotek nic nevyovídá“. Radimský své tvrzení opírá o následující argument: pokud by existence syntetického slovesného synonyma byla kritériem pro určování APKS, znamenalo by to, že *vykonat rozsudek (smrti)* určíme jako APKS, protože v češtině existuje sloveso *popravit*. To ale neplatí, protože *rozsudek*

(*smrti*) zdaleka neoznačuje stejný děj, který získává se slovesem *vykonat*. Macháčková (1983, s. 127) podotýká, že v záporu se mohou slovesa chovat jinak (v tomto případě se rozdíl stírá: *Neohlížej se na mne. – Neměj/neber na mne ohledy.*) Kolářová (2010, s. 111) dodává, že část APKS, ke kterým není jednoduše možné najít syntetický predikát, lze přiřadit k tzv. kvazimodálním slovesům a „slovesům záměru“ (*mít právo – moci, mít povinnost – muset, mít plán – chtít*).

K párům APKS a syntetického slovesa (nebo několika sloves) se později vrací slovník parafrází českých komplexních predikátů s kategoriálním slovesem ParaDi (Barančíková a Kettnerová 2017); detaily viz oddíl 2.6.3.

Dalším z rysů APKS jako analytického predikátu je **kompozicionalita významu**, který je na rozdíl od frazémů předem očekávatelný a odvoditelný, přestože je posunutý oproti samotnému predikativnímu významu slovesa. Pro správnou interpretaci APKS stačí vědět, že pro predikativní jméno má sloveso kategoriální platnost, protože se na významu prakticky nepodílí. Při úplné desémantizaci je význam závislý jen na predikativním jméně, při desémantizaci částečné je význam motivovaný, a tedy stále kompozicionální. Predikativní jméno je řídicím členem spojení (viz 2.4.4.2) a bez něj není interpretace možná, ale v některých případech je pro správnou interpretaci substantiva nezbytné i kategoriální sloveso (srov. *stanovit/vydat předpis* ve smyslu pravidla a *napsat/vystavit předpis* ve smyslu lékařský předpis) (Cinková 2009, s. 42–43; Macháčková 1979, s. 63; Radimský 2010, s. 158–159). Cinková dále podotýká, že díky kompozicionalitě významu jsou APKS stále produktivní při tvoření nových spojení.

Nominalizace APKS je v češtině možná v případě, že pro ni existuje stylistický důvod, jak ukazuje Radimského (2010, s. 161) výzkum na datech ČNK. U některých spojení je relativně běžná (*dosažení, kladení, podávání, podání, poskytování, provádění, provedení, předložení, přijímání, vytváření*), u jiných velmi ojedinělá (*buzení (pozornosti), chování (úcty)* atp.); téměř se nevyskytuje u slovesa *mít* (kromě předložkového dativu *k máni*). V případě nominalizace se ze slovesného předmětu v bezpředložkovém akuzativu stane substantivum vyžadující genitivní doplnění, případně může substantivum tvořit konstrukci s deverbativním adjektivem (*kontrola prováděná technikem*) (např. Kolářová 2010, s. 109; Martínek 2011, s. 126). K tématu se později okrajově vrací Kolářová (2014) v rámci analýzy nominalizovaných struktur se dvěma aktanty v genitivu adnominálním (*Tato banka se ujímá vlády v impériu IPB. → ujímání se vlády této banky v impériu*).

Macháčková (1979, s. 63–64) zmiňuje **blokování běžných synonymických paradigm** (*mít/*vlastnit/*držet pochybnost*), která jsou sice typická pro frazeologismy, ale u APKS je

možné vytvářet jiné paradigmatické řady, nebo dokonce řady antonymní, jde tedy spíše o specifický rys APKS, jak na řadě příkladů dokazuje Radimský (2010, s. 162–164).

Charakteristická je samozřejmě provázanost **valenčního a intenčního pole kategoriálního slovesa a predikativního jména**. Vychází z ní jak test koreference subjektů, tak test dvojí analýzy (viz 2.1.2), viz také část věnovaná koreferenci (2.4.4.1). V bohemistické literatuře se na toto téma zaměřují mnohé další práce (Kettnerová a Bejček 2016; Kettnerová a Lopatková 2017; Kettnerová et al. 2018, s. 36–40; Kolářová 2010; Macháčková 1979; 1983).

Poměrně skepticky se k pojetí valence staví Martínek (2011), který tvrdí, že APKS jako celé spojení nemá vlastní specifickou valenci (tak nemůže být kritérium určení APKS), nýbrž že valence APKS vzniká pouhou kombinací valencí substantiva a kategoriálního slovesa.⁶⁵ Ukazuje na to, že i predikativně užitá slovesa může mít rozdílnou valenci ve spojení s různými podstatnými jmény (srov. *těšit se něčí přízni – těšit se nač / z čeho, propadnout zoufalství – propadnout roštem*). Kralčákovo třídění odmítá pro nesrovnalosti a nedostatečné příklady, a o systematickém přístupu Kolářové (reprezentovaném v předchozích oddílech 2.3.2 a 2.4.4.4) patrně neví. Dále se vymezuje proti (patrně německému pojetí) APKS a jeho označování valence (*dal Petrovi jablko* jako dvě pravá valenční doplnění, *dal Petrovi pokoj* pouze jako jedno), s čímž jsem se sama nikde nesetkala.

Možná je vyznění článku dáno nedostatkem prostoru pro podrobnější argumentaci, protože i autor sám uvádí (Martínek 2011, s. 125–126), že z některých obligatorních doplnění predikativně použitého slovesa se stanou pouze potenciální doplnění APKS, což je jeden z důležitých komunikačních přínosů APKS. Domnívám se, že „specifická valence“ APKS, kterou Martínek marně hledá, vzniká právě kombinací valenčních rámců, protože bez přítomnosti predikativního jména tuto valenci sloveso mít nemůže, jak bylo prokázáno výše.

Otazníky se vznášejí u vyjadřování **stejně koreference dvakrát** – Kolářová (2010, s. 117) a Macháčková (1983, s. 135) uvádějí, že je možné vyjádřit stejný argument dvakrát bez porušení koreference alespoň v některých příkladech (*Petr Karlovi znovu položil svoji otázku. – Jan dostal Janův strach*). Radimský (2010, s. 167) se domnívá, že k tomu neexistuje systematická překážka, protože v některých kontextech to možné je (*Jan opět dostal ten svůj strach*).

Mezi další rysy související s APKS jako analytickým predikátem patří obtížná **pronominalizace predikativního jména** (viz pozn. 55) a **rozvíjení APKS adjektivy a adverbii** (2.3.3).

⁶⁵ Toto tvrzení ostatně po formální stránce potvrzují i výzkumy na PDT (viz 2.2.3.3.2).

2.6 Seznamy českých APKS

Pro praktickou část této disertační práce je klíčový seznam APKS k dalšímu využití/zpracování. Na počátku psaní této práce se zdálo, že žádný ucelenější seznam APKS neexistuje (např. Radimský (2010) neuvádí ucelený soupis APKS, jelikož se soustředí na kvalitativní rozbor těchto konstrukcí, a *Slovník české frazeologie a idiomatiky – výrazy slovesné* (Čermák a Hronek 2009) obsahuje jen vybrané APKS). Později se nicméně ukázalo, že takové zdroje existují a některé jsou i veřejně dostupné. V oddílech níže následuje jejich stručný výčet, který snad bude nápomocný dalším badatelům zabývajícím se stejnou problematikou.

2.6.1 Přílohy ke studiím

V kandidátské práci E. Macháčkové (1979) jsou dostupné dva seznamy, které vznikly pečlivou excerpcí ze sedmi knih publikovaných v 60. a 70. letech doplněnou o další příklady z tisku. První seznam je součástí sémanticko-syntaktického rozboru kategoriálních sloves tvořících APKS objektového typu (Macháčková 1979, s. 109–193). U většiny hesel autorka uvádí jednoduchý příklad, výklad významu, gramatický větný vzorec a přiřazení větných členů a příklady z literatury. Druhý seznam v příloze autorčiny práce obsahuje abstraktní substantiva, které se vyskytovala ve zkoumaných APKS. Nevýhodou obou seznamů je stáří práce, respektive kvalita tisku (strojopisná kopie přes 40 let stará). Z tohoto důvodu se práce špatně čte, natož digitalizuje. Předností je ale to, že seznam obsahuje APKS s doplněními ve všech pádech.

Pokud jde o starší češtinu, Martínek (2008b) ke své práci o Hájkově *Kronice české* připojil přílohu se soupisem excerpovaných APKS a podobných spojení.

V nedávné době byl publikován také užitečný seznam APKS, který je součástí článku *Ke koreferenci u komplexních predikátů s kategoriálním slovesem* (Kettnerová a Lopatková 2017, s. 21–26). Příloha obsahuje seznamy kategoriálních sloves s akuzativní valencí, která jsou roztříděna podle typu koreference a jsou k nim doplněna predikativní jména. Seznam je pořízen na základě anotace dat z korpusu SYN2010 pomocí nástroje Word Sketch Engine (Kettnerová a Bejček 2016).

2.6.2 VALLEX

Informaci o tom, že sloveso je součástí APKS (v terminologii slovníku se APKS označuje jako *komplexní predikát*), uvádí ve svých heslech také valenční slovník VALLEX 4.0⁶⁶ (Lopatková

⁶⁶ Komplexní predikáty jsou značeny již ve slovníku verze 3.5 publikovaném v roce 2018. Více k doplňování APKS do VALLEXu např. Kettnerová a Lopatková (2015), Kettnerová, Lopatková a Barančíková (2016).

et al. 2020; Lopatková et al. 2021). Aktuálně publikovaná verze zatím obsahuje výhradně APKS s predikativním jménem v akuzativu.⁶⁷ Kromě online rozhraní dostupného na adrese <https://ufal.mff.cuni.cz/vallex/4.0/> je slovník dostupný i ke stažení na adrese <http://hdl.handle.net/11234/1-3524>. Jednotlivá slovesa je možné vyhledávat jako slovníková hesla, případně lze zobrazit jejich přehled, a to rozkliknutím položky *MWE (multiword expressions, víceslovné výrazy)* v hlavním menu a zvolením záložky *complex predicates* (komplexní predikáty).

U hesel týkajících se potenciálně kategoriálních sloves je použita lexikální jednotka s významem *complex predicates (light verb)*. Ta je charakterizována valenčním rámcem (*frame*) s funktorem CPHR, který vyznačuje pozici predikativního jména (více k pojetí APKS v FGP viz pododíl 2.2.3.3). Ve slovníkovém hesle jsou dále uvedeny upřesňující atributy (Kettnerová et al. 2018, s. 45–46; Lopatková et al. 2021, s. 36–38, 41, 49):

- *lvc* – seznam predikativních jmen (resp. lexikálních jednotek), která se slovesem tvoří komplexní predikáty (atribut je použit i u některých predikativních jmen, kde obsahuje naopak výčet kategoriálních sloves); tento parametr propojuje oba rámce
- *map* – informace o koreferenci a slovesných a jmenných doplnění (pro činný rod), resp. tento argument obsahuje pravidla pro namapování valenčního rámce kategoriálního slovesa a predikativního jména
- *instig/instigator* – korespondence participantu kauzátor s některým z valenčních doplnění (pouze u kauzativních kategoriálních sloves)

Na adrese <https://quest.ms.mff.cuni.cz/vallex/> je navíc dostupné pokročilé vyhledávání, které umožňuje prohledávat mj. i nepublikovanou pracovní verzi slovníku (pomocí funktoru CPHR), což rozšíří řady uvedených APKS. Nepatrnou nevýhodou je skutečnost, že jednotlivé volby vyhledávání na webu nejsou nedostatečně dokumentovány (nebo se mi nepodařilo podrobnou dokumentaci nalézt). Nepříliš podrobná dokumentace omezuje potenciál tohoto veřejně dostupného nástroje – na více místech by specifikace možností filtrování zamezila formulování dotazů metodou pokus–omyl.

Na základě hesel z VALLEXu, resp. formálního popisu zahrnujícího hloubkověsyntaktické informace o valenční struktuře kategoriálního slovesa i predikativního jména a informace

⁶⁷ Jako kategoriální slovesa s genitivním doplněním uvádí slovesa: *dostat/dostávat se (ocenění, poděkování, podpory, uznání, výchovy, zadostiučinění a odvahy)* a *chápat/chopit se (iniciativ, práce, pravomoci, příležitosti)*; v dativu pouze *věnovat se (diskusi, fotografování, hledání, hraní, pátrání, pozorování, práci, psaní, sledování)*.

o koreferenci jejich valenčních doplnění a příp. kauzátoru lze generovat nebo analyzovat struktury zahrnující APKS.

2.6.3 ParaDi

Nedávno vznikl také slovník parafrází českých komplexních predikátů s kategoriálním slovesem ParaDi (Barančíková a Kettnerová 2017), který sice nemá žádné vyhledávací rozhraní, ale je pohodlně dostupný jako tabulka s daty, a to v repozitáři LINDAT na adrese <http://hdl.handle.net/11234/1-2605>. Aktuální verze 2.0 obsahuje 1621 párů kategoriálního slovesa a predikativního jména a k nim vždy jedno nebo více syntetických sloves. Více k tvorbě slovníku viz část 2.2.3.3.2.

3 Značkování korpusů⁶⁸

V současných psaných korpusech je každé slovo opatřeno mnoha doplňujícími údaji, mezi nejdůležitější patří *lemma* (základní, slovníkový tvar, vztažený k izolovanému slovnímu tvaru) a *morfologická značka* neboli *tag* (slovnědruhové a morfologické interpretace daného tvaru), poměrně nově je také doplňují pro slovesa tzv. *verbtagy* (gramatické kategorie slovesa, jak pro složené, tak prosté tvary). Korpusy jsou dnes tak velké, že není v lidských silách a časových možnostech je opatřit značkami ručně, jsou proto morfologicky značkovány (anotovány) automaticky. Dodatečné informace k surovému textu práci s korpusy výrazně zjednodušují a nabízejí téměř neomezené možnosti jeho využití, ovšem pouze za předpokladu, že tagy jsou přiřazeny správně. Značkování je přitom velmi zapeklitý úkol, zejména kvůli volnému slovosledu češtiny a velkému množství homonymních tvarů (65 % všech slovních tvarů).⁶⁹ Rodilí mluvčí si homonymii často neuvědomují – dané slovo je pro ně zjednoznačňené kontextem, sémantikou, syntaktickými vztahy, komunikační situací nebo znalostí okolního světa –, počítače se při zpracování textu zprvu dopouštěly značného množství chyb, protože nedokázaly se sémantikou pracovat (Petkevič 2014c, s. 533–534).

Cílem této kapitoly je přiblížit proces morfologického značkování korpusů psané češtiny⁷⁰ zejména čtenáři problematiky neznalému. Podává základ pro kapitolu následující, jež bude věnována dílčí části celého procesu, kterou tato disertační práce dále rozvíjí. Je vhodné také poznamenat, že značkování korpusů se neustále vylepšuje.

Morfologické značkování je ale sám o sobě komplexní úkol a sestává z množství dílčích procesů. *Tokenizace* rozdělí vstupní text na jednotlivé *tokens/pozice* (zjednodušeně řečeno na jednotlivá ortografická slova a interpunkční znaménka). *Větná segmentace* rozpozná jednotlivé věty. *Morfologická analýza* přiřadí každému tokenu (včetně interpunkce) množinu všech slovníkově přípustných dvojic morfologických vlastností a lemmat. *Morfologická disambiguace* (odstranění víceznačnosti) vybere na základě kontextu dané věty právě jednu správnou

⁶⁸ Celá kapitola je sestavena z poznatků zejména z těchto zdrojů: Jelínek a Petkevič (2011), Jelínek et al. (2021) a Petkevič (2006; 2014d), Pořízka (2014).

⁶⁹ Další informace a zejména zevrubný popis morfologické homonymie v češtině viz Petkevič (2014c).

⁷⁰ Až na výjimky se nebudu zabývat zahraničními výzkumy nebo přístupy (přestože proces zůstává podobný), není to ani v intencích této disertační práce, jelikož se dále budu věnovat jen ryze českému modulu podílejícímu se na značkování korpusů ČNK.

interpretaci (kombinaci morfologických vlastností včetně slovního druhu a lemmatu) a ta je pak obsažena ve výsledném korpusu.

V označovaném korpusu je možné vyhledávat podle lexémů, nikoli jen podle jednotlivých tvarů, hledat nejrůznější gramatické a slovosledné struktury a vztahy mezi nimi, zkoumat kombinace tvarů, lemmat a morfologických značek, sledovat kontexty lexémů a tvarů, které jsou dané jejich morfologickými vlastnostmi, a zjišťovat statistické údaje (srov. Petkevič 2014d, s. 195). V neposlední řadě je správné morfologické značkování základním předpokladem pro značkování syntaktické.

3.1 Tokenizace, větná segmentace a morfologická analýza

Při značkování korpusů ČNK (korpusu SYN2020, SYNv9 a novějších) je pro tokenizaci, větnou segmentaci i morfologickou analýzu používán nástroj MorphoDiTa (Straka, Straková a Hajič 2019; Straková, Straka a Hajič 2014) s využitím morfologického slovníku MorfFlex (Hajič et al. 2020) a databáze neologismů Neomat (Goláňová 2011). MorphoDiTa a MorfFlex byly významně aktualizovány v souvislosti s přidáním verbtagů (k nim další podkapitola, ke změnám v nástrojích viz Jelínek, Petkevič a Skoumalová (2022, s. 105–107)), stejně tak pravidlový disambiguační systém LanGr (detailně kapitola 4).

Vstupní text je pro budoucí morfologickou analýzu nejprve nutné připravit, k čemuž nám dopomohou programy *tokenizer* a *segmentátor* (nebo *segmenter*). Tokenizer rozdělí text na jednotlivé *tokens*, tj. na slova, čísla, zkratky a interpunkční znaménka. Musí se přitom vyrovnat s jejich víceznačností – např. zkratky končící tečkou nebo čísla s desetinnou čárkou.

Segmentátor text rozdělí do jednotlivých vět, a to na základě interpunkce, velkých písmen, soupisů různých názvů aj. Nejproblémovější je tečka, která je nejen znakem konce věty, ale také součástí zkratk (50) nebo označení elipsy (...), navíc může tečka ukončovat zkratku stojící zároveň na konci věty.

(50) Takovými prostředky získala v letech 650–500 př. </s><s> Kr. kontrolu ve státě řada „tyranů“.⁷¹ (SYNv10)

Výsledek tokenizace a segmentace se následně ještě koriguje – jednoduchými pravidly se spojí nebo rozdělí věty podobné příkladu výše a drobně se přizpůsobí výsledek tokenizace, aby odpovídal úzu používanému v daném korpusu (např. *e-mail* je v případě ČNK z původních tří

⁷¹ Značky </s><s> značí hranici věty; </s> znak konce věty, <s> znak začátku věty.

tokenů sloučen do jednoho, naopak složený tvar *abys* psaný jako jedno ortografické slovo je rozdělen a zpracováván jako dvojice slov, které jsou spojeny jsou až v poslední sérii úprav).

Po tokenizaci a segmentaci následuje morfologická analýza. Během ní jsou každému tokenu na základě rozsáhlého morfologického slovníku přiřazena všechna odpovídající *lemmata* (základní slovníkový tvar, který případně i sdružuje více variant jednoho slova – např. *disertace* a *dizertace* jsou opatřeny lemmatem *disertace*) a morfologické značky (*tagy*). Jejich množství je závislé na slovnědruhové (a)nebo morfologické homonymii. Rozpoznáno je téměř 100 % tvarů (Skoumalová 2011, s. 206) – kromě nesprávných interpretací (v případě homonymie) jsou tokenu vždy přiřazeny i značky a *lemmata*, které jsou pro daný kontext správné, je „pouze“ třeba z množiny všech dvojic (lemma, značka) vybrat ty správné. Slova, která se ve slovníku nevyskytují, jsou označena jako neznámá. Průměrně je přiřazeno každému 5,67 dvojic tagu a lemmatu na jeden token (bez verbtagu 4,06); jen u slovesných tvarů se přidáním verbtagu zvýšil počet z 2,49 na 11,83 kombinací tagu a lemmatu na token (Jelínek et al. 2021, s. 54–55).

Jako příklad může posloužit tvar *je*, který bude označen jako 3. os. sg. slovesa *být*, nebo *Asg neutra* či *Apl osobního zájmena oni*. Pokud by byl tvar napsán velkými písmeny (jako *JE*), mohl by značit i substantivní zkratku *jaderná elektrárna* (Jelínek a Petkevič 2011, s. 158).

Morfologické značky vyjadřují v podobě řetězce znaků slovní druh a poddruh daného tvaru a jeho morfologické vlastnosti. Pro češtinu existuje několik různých sad morfologických značek (*tagsetů*), které se navzájem liší. Nepoužívanější je tzv. pražský systém⁷² s pozičním tagsetem, který se využívá např. u korpusů ČNK a PDT, a tzv. brněnský⁷³ systém s tagsetem atributivním.

V pozičním systému má každý morfologický tag 15 pozic⁷⁴ a každé z nich je přiřazena jedna kategorie reprezentovaná sadou hodnot (značek) (Hajič 2004). Teoreticky lze vytvořit v situaci bez verbtagu více než 5000 kombinací, reálně používaných je ale zhruba 1500. V této podkapitole zmiňuji pro přehlednost jen hodnotu tagu; pozičnímu verbtagu, který obsahuje informace o gramatických kategoriích slovesa a při analýze zpracovává později, se věnuji až v následující podkapitole (3.2).

Pro již zmíněný slovní tvar *je* online verze taggeru MorphoDiTa (verze 1.3) vrací následující značky a *lemmata* (při zvoleném nastavení *Task – Analyze*, nikoliv *Task – Tag*, kde se už pokusí vybrat právě jeden tag).

⁷² Reprezentovaný zejména pracovišti ÚČNK, ÚTKL a ÚFAL.

⁷³ Reprezentovaný pracovišti MUNI.

⁷⁴ Původně měl tagset 16 pozic, ale přesunutím kategorie vidu z původní 16. pozice na původně neobsazenou 13. pozici došlo k jeho zkrácení (Křivan a Šindlerová 2022, s. 124).

lemma	tag	popis pozic ⁷⁵ (pořadí pozice)
být	VB-S---3P-AA---	sloveso (1.), tvar přítomného nebo budoucího času (2.), singulár (4.), 3. osoba (8.), prézens (9.), afirmativ (11.), aktivum (12)
on-1_^(oni/ono)	PPNS4--3-----	zájmeno (1.), osobní zájmeno (2.), neutrum (3.), singulár (4.), akuzativ (5.), 3. osoba (8.)
on-1_^(oni/ono)	PPXP4--3----- ⁷⁶	stejně jako výše, ale libovolný rod (3.) a plurál (4.)

Tabulka 5: Reprezentace slovního tvaru „je“ v pozičním tagsetu taggeru MorphoDiTa

Zdroj: Tagger MorphoDiTa, dostupné z <http://lindat.mff.cuni.cz/services/morphodita/>

V atributivním tagsetu se tag skládá z kombinace dvou znaků, kde první znak reprezentuje gramatickou kategorii a druhý její hodnotu, tagy tak mohou mít různou délku. Pro srovnání s pozičním tagsetem viz stejný tvar (*je*) v taggeru ajka, který využívá tagset atributivní.

základní tvar	číslo vzoru	kategorie	popis bigramů ⁷⁷ (bigram)
být	311-být	k5eAaImIp3n S	slovní druh – sloveso (k5), negace – afirmativ (eA), vid – imperfektum (aI), způsob – indikativ prézentu (mI), osoba – 3. (p3), číslo – singulár (nS)
on	1095-on	k3p3gMnPc4	slovní druh – zájmeno (k3), osoba – 3. (p3), rod – mužský životný (gM), číslo – plurál (nP), pád – akuzativ (c4)
		k3p3gInPc4	stejně jako výše, ale rod – mužský neživotný (gI)
		k3p3gNnPc4	stejně jako výše, ale rod – střední (gN)
		k3p3gFnPc4	stejně jako výše, ale rod – ženský (gF)
		k3p3gNnSc4	stejně jako výše, ale rod – střední (gN) a číslo – singulár (Sc)
je	210-ach	k0	slovní druh – citoslovce (k0)

Tabulka 6: Reprezentace slovního tvaru „je“ v atributivním tagsetu taggeru ajka

Zdroj: Tagger ajka, dostupné z <https://nlp.fi.muni.cz/projekty/wwwajka/WwwAjkaSkripty/morph.cgi>

⁷⁵ Detailní popis tagů viz Cvrček a Richterová (2022b).

⁷⁶ Online MorphoDiTa stále používá hodnotu tagu na třetí pozici – znak X se již v korpusech ČNK nevyužívá.

⁷⁷ Detailní popis brněnského tagsetu viz Jakubiček, Kovář a Šmerk (2011).

3.2 Přidání verbtagů

V rámci morfologické analýzy jsou v korpusu SYN2020, SYNv9 a v korpusech novějších přiřazeny možné šestimístné *verbtagy*, které doplňují gramatické kategorie⁷⁸ ke slovesům a k deverbativním adjektivům zakončených na *-ný/-tý*.⁷⁹ U sloves jsou všechny hodnoty vyznačeny u plnovýznamového slovesa, pokud jde o pomocné sloveso nese pouze hodnotu *A* (další hodnoty se neuvádějí). Verbtag značkuje konzistentně slovesa nezávisle na tom, jde-li o tvar jednoduchý (*vidím*), nebo složený (*viděl jsem*). Tato značka např. nově umožňuje jednoduše rozlišovat pomocné tvary slovesa *být* od sloves plnovýznamových nebo sponových, pracovat se slovesným způsobem a časem, rozlišovat infinitiv jako součást slovesného tvaru nebo v jiné funkci atp. V dřívějších verzích bylo k práci analytickými tvary nutné užívat poměrně komplexních dotazů. Ve verbtagu není obsažen údaj o afirmaci/negaci (nespolehlivost značkování) a také nejsou zpracovány modální a fázová slovesa nebo rezultativa (neshoda lingvistických teorií) (Jelínek et al. 2021, s. 50–52; Jelínek, Petkevič a Skoumalová 2022; Křivan a Šindlerová 2022, s. 134–139).

Kvůli přehlednosti při tvorbě korpusu a následné disambiguaci se verbtagy přidávají za původní morfologický tag (který tak dočasně vytvoří komplexní značku o délce 21 znaků, ale v rozhraní korpusového vyhledávače KonText se poté používají samostatně). Přidání verbtagů zvýší množství značek na jeden token průměrně z 4,06 na 5,67. Pokud bereme v úvahu pouze slovesa, tak dokonce z 2,49 na 11,83 (Jelínek et al. 2021, s. 54–55).

3.3 Morfologická disambiguace⁸⁰

Po morfologické analýze má většina tokenů v textu přiřazených více možných interpretací (lemmat i tagů), které jsou založené jen na izolované interpretaci daného tokenu, a z nich je potřeba vybrat správnou⁸¹ interpretaci pro daný kontext. Z toho důvodu vstupuje po ukončení morfologické analýzy do hry *automatická morfologická disambiguace* (zjednoznačnění), nejsložitější část celého morfologického značkování, která ze všech možností vybírá právě jeden tag s verbtagem, a právě jedno lemma. V současné době není systém používán pro korpusy

⁷⁸ Typ slovesa (sloveso plnovýznamové nebo sloveso *být/bývat* fungující jako pomocné) a k slovesům plnovýznamovým dále způsob, slovesný rod, osoba, číslo a čas (v pořadí tak, jak jsou používány v pozičním verbtagu), které jsou platné pro celý tvar.

⁷⁹ Ostatní slovní druhy mají verbtag „prázdný“, tj. s hodnotou -----.

⁸⁰ Srov. Jelínek (2008), Jelínek a Petkevič (2011), Petkevič (2014c, s. 540–546), Pořízka (2014, s. 226–238), Skoumalová (2011).

⁸¹ Nebo „správnější“, pokud by myslitelných interpretací bylo více, jako např. v případě následujících vět: *Jelínkovi* (Npl|Dsg) *prodávali maso*. nebo *Muži* (Npl|Dsg) *se zdáli zdraví*.

ČNK schopen pracovat s homonymními interpretacemi v možných víceznačných čteních.⁸² Za komplexností disambiguace stojí několik faktorů (srov. Petkevič 2006):

- disambiguace je relativně výpočetně náročná – jazyk se skládá z tisíců slovních tvarů a paradigmat, z čehož plyne velké množství možných kombinací, se kterými se všechny taggery musí vyrovnat
- přirozený jazyk se skládá z komplexních pravidel a množství výjimek z nich
- pro správnou disambiguaci nestačí pouze segmentovat a tokenizovat text a využít fonologických a morfologických vlastností, často je nutná i znalost syntaxe a sémantiky, která je v této fázi značkování dostupná jen v podobě dlouhých seznamů a je mnohem složitější než samotná morfologická analýza
- je nutné označkovat neznámá a cizojazyčná slova
- texty obsahují nejen spisovný jazyk, ale běžně se v nich vyskytují mluvené výrazy či nářeční varianty
- přirozený jazyk se neustále vyvíjí, což je nutné reflektovat

Disambiguaci lze v dnešní době provést způsoby uvedenými níže, díky pokrokům moderních technologií se v posledních letech přidalo i využití neuronových sítí (ve stochastickém přístupu):

- statisticky (stochasticky)
 - na základě strojového učení (*machine learning*) (Hajič 2004; Straková, Straka a Hajič 2014)
 - s využitím neuronových sítí / hlubokého učení (*neural networks / deep learning*) (Straka, Straková a Hajič 2019)
- za pomoci pravidel (tzv. rule-based)
 - automaticky vyvozených z textů (Brill 1992)
 - ručně vytvořených lingvisty (Petkevič 2006)
- hybridní metodou – spojením výše uvedených metod (anebo alespoň dvou z nich)

⁸² Již dnes lze složená slova (např. *abys*) označkovat dvěma různými tagy spojenými pomocí svislé čáry (|), v případě *abys J,-----|Vc-S---2----I--*. Podobné řešení by bylo možné použít i v případě víceznačnosti slova, jen by se zřejmě jednalo o okrajové množství případů, protože velká část víceznačných čteních je možné disambiguovat v závislosti na kontextu (např. žádná z vět uvedených v pozn. 81 se v korpusu nevyskytuje), anebo by bylo nutné číst celou větu víceznačně.

3.3.1 Statistické (stochastické) metody

U statistických metod probíhá disambiguace na základě strojového učení. Tagger se nejprve natrénuje na správně ručně označovaném vzorku dat (tzv. trénovací data)⁸³ – naučí se na něm velmi odlišné četnosti morfologických značek a jejich kombinací. Vytvoří si tak jazykový model, který je založen na pozitivních informacích o jazyce (tzn. nepracuje s tím, že v jazyce některé kombinace⁸⁴ nemohou existovat). Tuto znalost je pak schopný využít při značkování již neznámého a mnohem rozsáhlejšího textu, kdy z možných kombinací zvolí tu nejpravděpodobnější.

Disambiguace statistickou metodou má své výhody: není náročné takový tagger uzpůsobit konkrétnímu jazyku (ale je velmi nákladné získat trénovací data), zároveň je tagger poměrně robustní a schopný rychle označovat libovolný text, a to včetně textů komplikovaných či obsahujících chyby. Z těchto důvodů byl statistický tagger dlouhou dobu nejrozšířenějším typem disambiguace a až v posledních letech ho vytlačily taggery založené na neuronových sítích. Nevýhodou této metody naopak je, že může vytvořit i gramaticky nesprávné interpretace a je závislá na velikosti a kvalitě trénovacích dat. Kvůli využití trénovacích dat využívá pouze pozitivní informace (tj. informace přítomné a platné v dané větě), na rozdíl od ručně psaných pravidel, která mohou pracovat i na základě nepřítomných rysů nebo slov. Omezená je také schopnost se zlepšovat na základě odhalených chyb, protože samotný proces statistické disambiguace se chová jako černá skříňka a není možné do něj zasahovat (Oliva et al. 2000; Oliva a Květoň 2002). Chyby u starších taggerů (např. tvar *se* byl v 1 % případů značkován jako předložka v případech, kdy za ním následovalo slovo začínající na samohlásku – *díval se oknem*), které jsou jednoduše opravitelné, vedly k myšlence značkování za pomoci pravidlové metody (viz oddíl 3.3.2).

Mezi nejznámější mělké statistické taggery pro češtinu z dílny ÚFAL patří tzv. *Hajičův tagger* (Hajič 2004), *MorČe*⁸⁵ (Votrubec 2005) a nejnovější tagger *MorphoDiTa*⁸⁶ (Spoustová et al. 2009; Straková, Straka a Hajič 2014); z dílny Centra zpracování přirozeného jazyka

⁸³ V pražském kontextu je využíván korpus PDT, který se skládá z 1,95 milionu ručně anotovaných tokenů; zdrojová data k poslední verzi PDT 3.5 na <http://hdl.handle.net/11234/1-2621> (odkaz na webové rozhraní a další zdroje tamtéž).

⁸⁴ Takovým kombinacím se v počítačové lingvistice obecně říká *n-gramy* (*bigramy*, *trigramy* etc.), obvykle se jedná o kombinace forem, lemmat a značek. Pokud tvoří nepřerušovanou posloupnost, jedná se o *spojitý n-gram*, v případě posloupnosti přerušované *nespojité n-gram* (Petkevič 2014c, s. 541).

⁸⁵ Zkratka pochází z názvu MORfologie Češtiny; web projektu a program ke stažení na <http://ufal.mff.cuni.cz/morce/index.php>.

⁸⁶ Zkratka pochází z *Morphological Dictionary and Tagger*, web projektu na <http://ufal.mff.cuni.cz/morphodita/>.

FI MU jsou to taggery *ajka*⁸⁷ (Sedláček 1999) a nejnovější verze *majka*⁸⁸ (Šmerk 2009).⁸⁹ Obě zmíněná pracoviště používají pro anotaci textů vlastní sadu morfologických značek, které se navzájem liší (odlišné tagsety, viz výše), ale existuje snaha o jejich sblížení: projekt NovaMorf (viz např. Osolsobě et al. 2017; Petkevič et al. 2019), který přináší nové pojetí morfologického značkování.

Stejně jako v jiných odvětvích se i v lingvistice projevila dostupnost a síla neuronových sítí / hlubokého učení (*neural networks / deep learning*) a během psaní této práce se tyto technologie začaly ve velké míře používat i pro morfologickou disambiguaci. Právě na nich staví výrazně vylepšená verze původně statistického nástroje MorphoDiTa (Straka, Straková a Hajič 2019) zmiňovaného výše. I díky tomu⁹⁰ se úspěšnost disambiguace a celého morfologického značkování výrazně zvýšila – v korpusu SYN2020 byla správnost přiřazení lemmatu, tagu a verbtagu 97,32 % (Jelínek et al. 2021, s. 55). Na obecné rovině strojového učení tyto nástroje fungují podobně jako výše zmíněné starší taggery.

MorphoDiTa využívá podkladů vytvořených programem FastText⁹¹ (Bojanowski et al. 2017), který pro natrénování využil neanotovaná data z korpusu SYNv7 a české verze Wikipedie⁹². Pro samotné učení neuronového taggeru byl využit korpus Etalon⁹³ obsahující 2,27 milionů tokenů ručně doplněných o lemmata, tagy a verbtagy a skládající se z publicistických, odborných i beletristických textů.

3.3.2 Pravidlové metody

Disambiguace za pomoci pravidel se od výše uvedených metod odlišuje tím, že nevyvozuje pravidla z omezeného kontextu testovacího korpusu, ale využívá k odstranění homonymních tvarů pravidla vytvořená lingvisty pro konkrétní jazyk. Tento způsob uvažování o značkování syntaktickém a morfologickém byl dříve navrhnout a implementován pro závislostní teorii

⁸⁷ Zkratka pochází z Analyzátor JazyKA, web projektu: <http://nlp.fi.muni.cz/projekty/ajka/ajkacz.htm>.

⁸⁸ Zkratka z Morfologický Analyzátor JazyKA, stránka projektu: <http://nlp.fi.muni.cz/czech-morphology-analyser/>.

⁸⁹ Přehled nástrojů ÚFAL MFF UK a Centra zpracování přirozeného jazyka FI MU včetně stručné charakteristiky dalších existujících tagsetů viz Pořízka (Pořízka 2014, s. 226–238).

⁹⁰ Za zmínku zde stojí skutečnost, že verbtag naopak značkují v drtivé míře pravidla (za poznámku děkuji Vladimíru Petkevičovi).

⁹¹ Software včetně předtrénovaných modelů ke stažení na <https://fasttext.cc/>.

⁹² Dostupné ke stažení zde: <https://dumps.wikimedia.org/cswiki/>.

⁹³ Dostupný v repozitáři LINDAT na adrese: <http://hdl.handle.net/11234/1-3698>; další informace ke korpusu jsou dostupné z <https://wiki.korpus.cz/doku.php/cnk:etalon>.

(*constraint grammar*) F. Karlssona (Karlsson et al. 1995), k myšlence aplikovat ho i na češtinu viz Oliva et al. (2000).

Jednotlivá pravidla představují precizně formulované a formalizované poznatky o jazykovém systému, které jsou založeny na syntaktických pravidlech. Např. v češtině je při vytváření pravidel třeba brát ohled na složitou morfologii, velkou homonymii výrazů, množství výjimek a nepravidelností, a především volný slovosled. Pravidlové systémy přitom umějí využívat pozitivní i negativní fakta o jazyce, které mohou čerpat z de facto neomezeného kontextu, ale jen v rámci věty.

Při negativním přístupu vytvořená sada pravidel odstraňuje morfologické značky, které neodpovídají gramatickým konstrukcím v daném textovém kontextu. Pravidla se aplikují do té doby, než má každé slovo právě jedno lemma a jednu značku, nebo než se všechna pravidla „vyčerpají“ a již neexistuje žádné, které by se mohlo aplikovat. Doplnkem k těmto pravidlům je kolokační komponenta, která naopak identifikuje správnou značku rovnou (pozitivní přístup).

Zásadní nevýhodou disambiguace založené na pravidlech je jednak velká intelektuální náročnost vytváření pravidel, jednak citlivost na chyby v textu či slova nerozpoznaná při morfologické analýze; problémem jsou i slova chybně anotovaná, chybějící interpunkce či nesprávné rozdělení věty. Navíc tato metoda nedokáže většinu vět disambiguovat úplně – i po aplikaci pravidel zůstanou slova disambiguována jen částečně (zbývá jim více než jeden tag nebo lemma). V posledních letech jí navíc začínají konkurovat nástroje založené na neuronových sítích (viz předchozí oddíl).

Výhodou je malá chybovost a možnost případné chyby poměrně snadno napravit (úpravou pravidla, vytvořením výjimky apod.). Nutnost pravidla přesně zformulovat vede k přesnějšímu popisu gramatického systému a jeho hlubšímu pochopení – bez toho by vůbec nebylo možné taková pravidla sestavit. Pravidla mohou sloužit jako základ pro budoucí syntaktickou a sémantickou analýzu věty, resp. jednotlivých slov, a také je možné je využít pro tvorbu dalších nástrojů, např. pro kontrolu pravopisu. Toho se reálně využilo při tvorbě nástroje pro kontrolu české gramatiky v programu Microsoft Office™ – Word verze 2003, 2007, 2010 a 2017 (Petkevič 2014b). Přístupu založeného na pravidlech se využilo při vytváření nového nástroje pro kontrolu pravopisu Opravidlo.⁹⁴

⁹⁴ Dostupné z: <https://www.opravidlo.cz/>.

Pro češtinu existuje takový pravidlový tagger pouze jeden, a to počítačový systém pravidly řízené redukční automatické morfologické disambiguace LanGr⁹⁵ (srov. Květoň 2006a), který je rozvíjen na ÚTKL. Detailněji se jím budu zabývat v následující kapitole.

3.3.3 Hybridní metody

Kombinace metody založené na neuronových sítích nebo starších metodách strojového učení a pravidlové metody disambiguace se nazývá *metoda hybridní*. Spojuje výhody disambiguace podle pravidel (malá chybovost) a statistických taggerů (dopouštějí se minima chyb a provedou úplnou disambiguaci). Taková kombinace taggerů dosahuje prozatím nepatrně lepších výsledků než samostatná MorphoDiTa založená na neuronových sítích. V případě kombinace systému LanGr a MorphoDiTy použité na korpusu SYN2020 se jednalo už o statisticky nevýznamný rozdíl – 97,36 % při využití obou modulů oproti 97,32 % při použití samotné MorphoDiTy. I přesto je hybridní metoda při značkování korpusů SYN preferována, a to z následujících důvodů (Jelínek et al. 2021, s. 57):

- hybridní metoda je stále úspěšnější než samostatný tagger (viz výše)
- chyby v pravidlech je možné snadno identifikovat a opravit, což u neuronových sítí není možné
- tagger MorphoDiTa občas chybuje při disambiguaci relativně jednoduchých syntaktických struktur, jako je předložková rekece a shoda přívlastku s podstatným jménem, což se v pravidlové disambiguaci téměř nikdy nestává.

Jak již bylo uvedeno, hybridní metoda se používá pro automatickou disambiguaci korpusů řady SYN ČNK. Od prvního použití hybridního modelu na korpusu SYN2005, kdy byl pravidlový tagger využit jen minimálně (zjednodušoval pádovou homonymii), uplynula poměrně dlouhá doba, během níž se vyzkoušela různá míra spolupráce se stochastickými taggery (srov. Jelínek 2008; Skoumalová 2011; Spoustová et al. 2007) a poslední době s taggery založenými na neuronových sítích (Jelínek et al. 2021). Aktuálně se používá kombinace následujících modulů:

- (a) pravidlový systém *LanGr* (Jelínek 2008; Jelínek a Petkevič 2011; Květoň 2006a; Petkevič 2006), který obsahuje přibližně 4500 ručně psaných pravidel
- (b) kolokačně-frazémový modul *Phras* (Hnátková 2002; Hnátková a Petkevič 2017), který identifikuje a disambiguuje členy frazémů a kolokací

⁹⁵ Zkratka pochází ze spojení LANguage GRammar.

(c) tagger *MorphoDiTa*⁹⁶ založený na neuronových sítích (Straka, Straková a Hajič 2019)

3.4 Syntaktická anotace

V případě psaných korpusů řady SYN je přidána ještě vrstva syntaktické anotace, a to pomocí parseru ze skupiny NeuroNLP2 (Ma et al. 2018) založeného na neuronových sítích. Tento parser využívá pro učení podklady z analytické vrstvy v PDT (Bejček et al. 2012) a beletristického korpusu FicTree,⁹⁷ který je také syntakticky anotován. Dále je stejně jako MorphoDiTa natrénován na datech SYNv7 a české Wikipedii (Jelínek et al. 2021, s. 55).

S ohledem na to, že pravidla vytvořená v rámci práce budou kvalitu syntaktického parsování ovlivňovat pouze nepřímo (lepší morfologická anotace ústí v potenciálně lepší syntaktickou analýzu), nebudu v této podkapitole zabíhat do větších detailů.

3.5 Dodatečné úpravy a finalizace

Úpravy prováděné po dokončení samotného procesu morfologické disambiguace (případně i syntaktické analýzy) se mohou lišit v závislosti na značkováném korpusu. Např. jednou z fází při anotaci korpusů řady SYN je přidání dodatečných informací, jako jsou sublemmata⁹⁸ (podskupina lemmatů zohledňující různý zápis nebo stylovou variantu – např. *filosofie* a *filozofie* nebo *okénko*, *okýnko* a *vokno*) a spojení víceslovných jednotek (*multiword tokens*),⁹⁹ které byly rozděleny během tokenizace, zpět do jednoho tokenu (více lemmat i tagů je zachováno rozdělením hodnot pomocí svislé čáry – token *abys* je reprezentován složeným lemmatem *aby|být* a složeným tagem *J,-----|Vc-S---2----I--*) (Jelínek et al. 2021, s. 49–50, 52–53, 56).

⁹⁶ Web projektu na <http://ufal.mff.cuni.cz/morphodita/>.

⁹⁷ Dostupný v repozitáři LINDAT na adrese: <http://hdl.handle.net/11234/1-2517>.

⁹⁸ Tzv. dvojúrovňová lemmatizace se používá v korpusech SYN2020 a novějších; k těmto změnám více Křivan a Šindlerová (2022) nebo *Příručka ČNK* (Cvrček a Richterová 2022a).

⁹⁹ Dle standardu Universal Dependencies (Nivre et al. 2020).

4 Systém LanGr¹⁰⁰

V této kapitole se budu detailněji věnovat systému pravidly řízené redukční automatické morfologické disambiguace LanGr (Jelínek 2008; Jelínek a Petkevič 2011; Květoň 2006a; 2006b; Petkevič 2006; 2014a; 2014c; 2014d), který se používá v ÚTKL¹⁰¹ pro značkování korpusů ČNK (na jiných pracovištích se využívá jen stochastických taggerů).

Analýza poměrně triviálních chyb statistických taggerů (např. *díval se oknem*, kde je *se* označeno jako předložka) vedla v roce 2000 k započetí práci na ručně psaných pravidlech pro morfologickou disambiguaci (Oliva et al. 2000) a systém se neustále vyvíjí a rozšiřuje. V dnešní době čítá přes 4500 ručně vytvořených pravidel (Hnátková et al. 2014, s. 162).

4.1 Pravidla

Jak již bylo uvedeno v oddíle 3.3.2, jednotlivá pravidla představují precizně formulovaný a formalizovaný poznatek o úzce specifikovaném jevu nebo vlastnosti. Příkladem může být např. pravidlo, že po vokalizované předložce nestojí slovo začínající na samohlásku (51) či že po předložce bezprostředně nestojí slovesný tvar (52) (Petkevič 2006, příklady s. 11 a 17).

(51) Včera jsme **se** (~~Prep~~|Pron) vážně snažili nakoupit nějaké jídlo. (VP)

(52) Jsem **pro rozhodnou** (Adj|~~Verb~~) odpověď válečným štváčům. (VP)

Vytváření pravidel pro češtinu ztěžuje mj. složitá morfologie, velká tvarová homonymie, množství výjimek a nepravidelností a volný slovosled, z něž vyplývá poměrně malý počet pevných bodů ve větě, které lze pro disambiguaci použít. Opřít se lze naopak o přísná interpunkční pravidla v českém pravopisu, která je možné využít zejména tehdy, je-li text psán spisovnou češtinou.

Ke konstrukci pravidel dochází buď introspekci lingvisty (ta je pak ověřena na korpusových datech), anebo studiem korpusových dat a jejich dosavadního značkování a vyvozováním pravidel (metoda *corpus-driven*); pravidla dosud nejsou tvořena zkoumáním chyb statistických taggerů. Využívá se i dalších informací, jako jsou soupisy slov se stejnou vlastností,

¹⁰⁰ Srov. Hnátková et al. (2014), Jelínek a Petkevič (2011b, s. 159–168), Květoň (2006), Petkevič (2014b, s. 542–558, 2014c), Spoustová et al. (2007).

¹⁰¹ Systém vyvíjejí členové ústavu – Milena Hnátková, Tomáš Jelínek, Vladimír Petkevič a Hana Skoumalová – a podíleli se na něm i externisté Pavel Květoň s Karlem Olivou (Petkevič 2014c, s. 542–543; aktualizováno 2022).

valenční slovníky a samozřejmě dostupná literatura atp. Všechna nově implementovaná pravidla jsou před zařazením do ostrého provozu intenzivně testována na testovacím korpusu.

Cílem pravidel je odstranit co nejvíce nesprávných morfologických značek a lemmat a zároveň neodstranit údaje správné. Pravidla je proto nutné formulovat s rozmyslem: pokud budou neopatrná a bezohledná, odstraní sice značek více, ale riskujeme, že odstraní i ty správné, na druhou stranu příliš opatrná pravidla odstraní méně správných značek, ale zároveň i méně značek nesprávných. Pravidla jsou ještě doplněna speciálním kolokačním modulem *Phras*, který zpracovává kolokace, frazémy a idiomy (srov. Hnátková 2002; Hnátková a Petkevič 2017) a během disambiguace se spouští mezi jednotlivými kroky. Skupiny pravidel jsou uspořádány do dalších skupin a podskupin, při jejichž aplikaci záleží na pořadí – nejprve se aplikují nejspecifičtější pravidla a až pak pravidla obecná, zároveň úzce spolupracují s frazémovým komponentem. Pravidla jsou téměř 100% spolehlivá a reprezentují syntaktické pravidelnosti češtiny – např. slovní tvar v nominativu nemůže být za jistou předložkou:

(53) **Prohřešek** (nom|~~ae~~) ho dovedl až **před soud** (~~nom~~|acc). (VP)

Samotný zápis pravidla se skládá ze dvou, resp. čtyř částí (Květoň 2006a; Petkevič 2004; Petkevič a Hnátková 2002); k jejich konstrukci a implementaci více v oddíle 4.3, zde jen stručně:

(a) **kontext a konfigurace** (context / configuration part)

(a1) **kontext** – jednoznačná definice kontextu (vstupní podmínky), který musí být bezvýhradně splněn, aby se vůbec pravidlo uplatnilo; není změněn provedením výkonného příkazu

(a2) **prostor pro disambiguaci** – charakteristika kontextu, na němž se akce (změna dat) provede

(b) **akce / výkonný příkaz** (action / executive part)

(b1) **nepovinný report** – informace o provedené akci

(b2) **vykonání akce** – samotné odebrání jedné nebo více nesprávných značek

Kontext (a1) je vždy jednoznačně specifikovaný a pravidlem se nemění. Naopak prostor pro disambiguaci (a2) je předmětem změny, tj. je vybírán takový token (či tokeny), který obsahuje morfologickou značku, kterou bude akce upravovat. U vybraných tagů a lemmat je po splnění podmínek pravidla prováděna akce (b2) – nejčastěji odstranění nesprávné značky nebo ponechání pouze správné interpretace, doplňkově k tomu se mohou unifikovat značky napříč tokeny

(např. u shody adjektiva a substantiva apod.). Report (b1) uloží do logu zprávu popisující provedenou změnu.

4.2 Průběh disambiguace

Jak napovídá slovo „redukční“ v názvu celého disambiguačního systému, jedná se zejména o negativní přístup k problematice značkování, pracuje se s negramatickými kombinacemi¹⁰² tagů a lemmat (srov. Petkevič 2004). Platí, že pokud pravidlo značku odstraní, je si téměř 100% jisté, že je nesprávná. Pokaždé, když se konkrétní pravidlo úspěšně aplikuje a odebere se morfologická značka, zpřesní se znalost o dané větě. Díky tomu mohou být použita další pravidla a tím se odstraňují další značky; prostřednictvím stále jednoznačnějších dat tak pravidla spolupracují. V případě, že pravidlo odstraní značku nesprávnou, je možné při kontrole chybu lokalizovat a následně napravit (na rozdíl od přístupu založeného na strojovém učení).

Průběh disambiguace vypadá zhruba následovně: nejprve se spustí pravidla, pokud se vyčerpají, zapojí se modul Phras (Hnátková 2002; Hnátková a Petkevič 2017), který rozpoznává kolokace a frazémy. Pravidla a kolokační modul se aplikují stále dokola, dokud už žádnou značku odstranit nedokážou. Pokud věta není zcela disambiguována, práci převezme statistický tagger (při značkování korpusů ČNK tagger MorphoDiTa). Ten práci dokončí a zajistí, že každé slovo bude mít právě jednu značku a lemma (z homonymních struktur je vybrána jen jedna) (Petkevič 2014d, s. 199).

Podle výzkumu Skoumalové (2011), která měřila účinnost morfologické analýzy a následné disambiguace založené na pravidlech na testovacích datech z Pražského závislostního korpusu, systém chybně odstraní správnou značku zhruba v 1,18 % případů, přitom odstraní zhruba 70 % značek nesprávných. Novější měření provedené Jelínkem et al. (2021, s. 55) účinnost potvrzuje – při anotaci korpusu SYN2020 bylo odstraněno 69 % nesprávných značek.

4.3 Implementace pravidel LanGr

Pravidla samotná jsou zapsána ve speciálním programovacím jazyce *LanGr* (Květoň 2006a; 2008), který byl vytvořen právě za účelem tvorby a implementace disambiguačních pravidel. V pravidlech je možné využít širokého spektra již dostupných lingvistických zdrojů – homonymních tříd, frazeologických slovníků, valenčních seznamů a slovníků a jiných (Petkevič 2014c, s. 545). Co se týká samotné podoby jazyka LanGr, jde o programovací jazyk pro popis

¹⁰² Pokud takové kombinace následující za sebou, jedná se o tzv. *negativní n-gramy* (*negative n-grams*); podrobnější informace viz např. Květoň (Květoň 2006a, s. 27–38).

pravidel přirozeného jazyka, který byl vytvořen jako jednodušší programovací jazyk přímo na míru lingvistům.

Dnes je termínem LanGr označován jak programovací jazyk, tak celý disambiguační systém. V původní koncepci a s ní související literatuře byl programovací jazyk označován termínem *Rule Language* a rozlišoval dvě vrstvy (Květoň 2008, s. 2):

- *RL jádro (core)*, což je procedurální jazyk podobný programovacím jazykům C/C++ a Pascal; slouží k vytváření funkcí a interpretaci pravidel psaných v RL aplikaci
- *RL aplikace*, soubor funkcí, datových typů a proměnných definovaných v RL jádře; hlavní nástroj pro lingvistický zápis pravidel; součástí aplikace jsou také jazykově závislé informace, jako slovníky, valence, soubor morfologických kategorií.

Samotné jádro programu (Květoň 2006a, s. 52–56; 2008) ponechám stranou, protože se práce lingvistů převážně netýká a v pozdější literatuře již není zmiňováno. Budu se věnovat té části programovacího jazyka LanGr, kterou lingvisté ve své práci používají, a to konkrétně zápisu jednotlivých pravidel. Pro „polidštěnou“ formu zápisu pravidel včetně detailního lingvistického rozboru srov. např. Jelínek a Petkevič (2011, s. 162–167) a Petkevič (2014c, s. 546–558).

Každé pravidlo zapsané v programovacím jazyce LanGr se skládá ze dvou formálních částí: *kontextu/konfigurace* a *akce / výkonného příkazu*, viz podkapitola 4.1. Kontext sestává z řady tokenů, které jsou vyhledány ve vstupní větě (svou strukturou připomínají regulární výrazy). Na nich (nebo v závislosti na nich) se provede cílová akce, tedy odstranění morfologických značek v případě, že pravidlo bylo splněno. Délka kontextu je z praktických důvodů omezena hranicí věty, obvykle se jedná pouze o několik slov, tzv. n-gramy.

Samotný programovací jazyk LanGr má velmi striktní syntaktická pravidla. Jako každý programovací jazyk umožňuje i tento psát komentáře, v tomto případě je použit styl známý například z jazyků C++, PHP či JAVA. Příloha 1 obsahuje (snad kompletní) přehled klíčových slov, operátorů, výkonných příkazů, komentářů a jejich formátování, které mohou být v LanGr použity a který jsem sestavila během psaní této práce, jelikož nikde kompletní seznam publikován nebyl.

V následujících oddílech pro čtenářovu představu stručně projdu dva úvodní popisy pravidel tak, jak je uvádí ve své disertační práci Květoň (2006a, s. 56–75), veškerý přejatý kód pochází z jeho práce.

4.3.1 Negativní bigram¹⁰³

```
1. rule TwoFiniteVerbs {
2.
3.   rulevariant VarI {
4.     ITEM IsSafe FiniteVerb;
5.     ToDelete = ITEM Possible not FiniteVerb;
6.     DELETE FiniteVerb FROM ToDelete;
7.   };
8.
9.   rulevariant VarII {
10.    ToDelete = ITEM Possible FiniteVerb;
11.    ITEM IsSafe FiniteVerb;
12.    DELETE FiniteVerb FROM ToDelete;
13.    // Případně je možno i též:
14.    // LEAVE ONLY not FiniteVerb
15.
16.   };
17.
18. };
```

Zdrojový kód 1: Zápis negativního bigramu po sobě následujících finitních sloves v LanGru
Zdroj: Květoň (2006a, s. 56); aktualizováno dle současně používané verze LanGru

Výše uvedený kód reprezentuje jednoduchý negativní bigram – v češtině nemohou existovat dvě po sobě následující **finitní** slovesa. Jako příklad můžeme uvést věty:

(54) **Boj** (Subst|Verb) **začal** (Verb) znenadání. (PK)

(55) Muži **vyčistili** (Verb) **ústí** (Subst|Verb) řeky. (PK)

Každé pravidlo v RL aplikaci má danou strukturu – začíná na klíčové slovo *rule*, po němž následuje název pravidla. Pravidlo se vždy skládá alespoň z jedné pojmenované varianty (*rulevariant*), v našem případě jsou dvě, protože možný finitní slovesný tvar může být v rámci kontextu jako první nebo druhý. Uvnitř varianty je samotný příkaz.

ITEM odpovídá pozici, která splňuje po něm následující podmínku. Aby byla daná pozice vybrána, musí být podmínka splněna. V tomto případě jde o podmínku *IsSafe*, která znamená, že všechny značky daného tokenu musí splňovat danou podmínku, tj. odpovídat definovanému kontextu. Token, s nímž se dále pracuje, je třeba uložit do proměnné (zde *ToDelete*). Ta je předána výkonné části, v tomto případě příkazu *DELETE*, který odstraní danou vlastnost (vyjádřenou *globálním identifikátorem*) z vybraného tokenu. *Globální identifikátor* je nadefinovaný v jádru RL a reprezentuje nejčastěji vlastnost nebo seznam, v LanGru je takových identifikátorů dostupné velké množství, proto není třeba znovu a znovu vypisovat tytéž příkazy

¹⁰³ Květoň (2006a, s. 52–56).

(globální identifikátor se chová v podstatě jako pojmenovaná funkce). V případě *FiniteVerb* algoritmus zjistí, jestli se jedná o minulé participium, přítomný nebo budoucí čas, či imperativ.

4.3.2 Sjednocení sekvence s omezením¹⁰⁴

```
19. rule Jehoz {
20.   rulevariant v3 {
21.
22.     PronounJehoz = ITEM (lower form == "jehož") or (IsSafe PronounRelativePoss);
23.
24.     SEQUENCE of IsSafe ((Adjective or PronounSyntAdj or NumeralSyntAdj)
25.     and not lemma Member of ["všechno", "ten", "všichni", "to"]);
26.
27.     Noun1 = ITEM IsSafe Noun;
28.
29.     ITEM IsSafe not (AdjectiveVerbal);
30.
31.     REPORT "Vztažné zájmeno přivlastňovací "
32.     emphasize(wordformonposition(PronounJehoz))
33.     " se musí shodovat s následujícím substantivem "
34.     emphasize(wordformonposition(Noun1))
35.     " v rodě, čísle a pádě!";
36.
37.     UNIFY UNILATERALLY PronounRelativePoss
38.     FROM PronounJehoz WITH Noun1
39.     IN [number, gender, case];
40.
41.   };
42. }
```

Zdrojový kód 2: Zápis v LanGru pro tvar *jehož* využívající sjednocené sekvence s omezením
Zdroj: Květoň (2006a, s. 68); aktualizováno dle současné používané verze LanGru

V češtině se přivlastňovací vztažná zájmena (*jehož*, *jejíž*) musí shodovat s následujícím podstatným jménem, ale je třeba dát pozor na to, že tvar slova *jehož* může být zároveň i zájmenem osobním vztažným. Právě tento problém řeší výše uvedené pravidlo *Jehoz*.

V kódu vidíme nový příkaz *SEQUENCE OF*, který specifikuje řadu tokenů splňující danou podmínku (adjektivum, zájmeno nebo číslovka s funkcí syntaktického adjektiva).

Samotná výkonná část je jednostranné sjednocení, které vybere z *PronounJehoz* pouze *PronounRelativePoss* a sjednotí ho v čísle, rodě a pádu s *Noun1* (*UNILATERALLY* znamená, že bude změněn pouze první člen sjednocené pozice).

¹⁰⁴ Květoň (2006a, s. 68–69).

5 Praktická část

Hlavním záměrem empirické části této disertace je spojit poznatky z předchozích kapitol a využít zjištění nabytá studiem APKS (kapitola 2) k vylepšení značkování korpusů (kapitola 3) pomocí pravidlově řízené morfologické disambiguace LanGr (kapitola 4). Předpokládám také, že nový pohled na problematiku APKS pomůže rozšířit lingvistické poznání těchto spojení.

V této kapitole nejprve objasním svoji motivaci a předpoklady (5.1), se kterými přistupuji k výzkumu. Představím také jazykový materiál a to, jak s ním budu dále pracovat (5.2). Poté se ve dvou případových studiích zaměřím na APKS v akuzativu a poznatky ze studií implementuji do pravidel pro morfologickou disambiguaci (5.3). V závěru se ještě vrátím k popisu používaných skriptů (5.4).

5.1 Motivace a předpoklady

Petkevič (2014d, s. 199–200) ve svém starším článku věnujícím se nedostatkům automatické morfologické disambiguace češtiny zmiňuje několik okruhů, jejichž řešení bude mít velký vliv na kvalitu tohoto procesu. Na prvním místě stojí problematika morfologické homonymie (zahrnující i pádový synkretismus; detailněji k termínu níže), dále autor uvádí např. struktury s krátkým kontextem, nedostatečné využití valenčních vlastností lexému a malý počet zpracovaných kolokací.

Domnívám se, že by k řešení výše zmíněných okruhů mohlo přispět právě zformulování obecných povrchově syntaktických pravidel o strukturách APKS. Předpokládám, že stejně jako jiné jazykové jevy budou mít i APKS jakousi nejčastější reprezentaci, o kterou se půjde při značkování opřít. Níže nastíním, jak by takové řešení mohlo fungovat.

Nejpalčivějším problémem je již zmíněná morfologická homonymie. Tímto termínem označuji morfologickou totožnost formy (Petkevič 2014c, s. 27–31; 2015, s. 129–132; Těšitelová 1966, s. 8), a to v těchto případech:

- (a) v rámci jednoho tvarového systému jednoho lexému – **morfologická homonymie vlastní**, vnitroparadigmatická (např. tvar *soudce* vyjadřuje Nsg, Gsg, Vsg, Npl, Apl a Vpl lexému *soudce*)
- (b) v rámci více tvarových systémů – **morfologická homonymie nevlastní**

- (b1) jeden lexém s různým jmenným rodem (např. substantivní tvar *model* je tvarem Nsg, Asg maskulina životného i neživotného *model* a tvarem Gpl feminina *modla*)
- (b2) více lexémů patřících k jednomu slovnímu druhu (např. tvar *tancích* je Lpl substantiv *tank* i *tanec*)
- (b3) více slovních druhů (např. tvar *smlouvám* je Dpl substantiva *smlouva*, ale i 1. os. sg. přítentu slovesa *smlouvat*)

Jedním z mnoha případů vlastní i nevlastní morfologické homonymie je pádový synkretismus, často tematizovaný ve spojitosti se substantivy – v každém z paradigmat najdeme alespoň jeden tvar, který má více morfologických interpretací (Petkevič 2015, s. 131). V případě verbálních substantiv, která do APKS často vstupují, jsou dokonce homonymní všechny tvary kromě Lpl a Ipl (Isg se shoduje s Dpl, Nsg je shodný se všemi zbývajícími pády). Pokud jde o další typy morfologické homonymie, extenze tohoto jevu vynikla až s rozvojem počítačového zpracování dat.¹⁰⁵ Jelikož je třeba se s tímto fenoménem vypořádat při disambiguaci, vybudoval se rozsáhlý morfologický slovník (viz kapitola 3). Kromě toho má např. homonymie nominativu a akuzativu vliv i na syntaktickou anotaci, a to při rozpoznávání podmětu a předmětu.

Výše zmíněné můžeme jednoduše ilustrovat níže na příkladu vlastní morfologické homonymie mezi tvary nominativu a akuzativu. Pro jejich správnou disambiguaci potřebujeme jemně rozpracovanou sémantiku, protože se tvary neliší formou a slovosled češtiny není pevný.

(56) **Soudce** (Nsg|Gsg|Asg|Vsg|Apl) vynesl **rozsudek** (Nsg|Asg). (vl)

(57) **Rozsudek** (Nsg|Asg) vynesl **soudce** (Nsg|Gsg|Asg|Vsg|Apl). (vl)

To, že *soudce* je v příkladech (56) a (57) podmětem v nominativu a *rozsudek* předmětem v akuzativu, z věty nelze syntaktickými nebo morfologickými pravidly bezpečně určit; i taggery v některých případech ve volbě správných pádů chybují. Pokud ale víme, že substantivum *rozsudek* tvoří APKS mj. se slovesem *vynést* (a sloveso netvoří žádné akuzativní spojení se substantivem *soudce*), můžeme předpokládat, že jsou-li obě jednotky v jedné větě/klauzi v rámci předem specifikované struktury, je substantivum *rozsudek* v akuzativu. Takto lze věty disambiguovat – nejprve přiřadit Asg tokenu *rozsudek*, a díky tomu v dalším kroku disambiguovat i *soudce*, kterému už žádná jiná interpretace nezbývá:

¹⁰⁵ Na základě korpusových dat byla také publikována rozsáhlá monografie *Morfologická homonymie v současné češtině* (Petkevič 2014c), která se problematice zevrubně věnuje.

(58) **Soudce** (Nsg|Gsg|Asg|Vsg|Apl) vynesl **rozsudek** (Nsg|Asg). (vl)

(59) **Rozsudek** (Nsg|Asg) vynesl **soudce** (Nsg|Gsg|Asg|Vsg|Apl). (vl)

Pokud by ale ve vzorové větě (56) bylo substantivum *soudce* nahrazeno substantivem *soud* jako v příkladu (60), uvedený postup by nefungoval, jelikož bohužel existuje, jak APKS *vynést rozsudek*, tak APKS *vynést soud*, které je sice méně frekventované, ale přesto se v korpusových datech jeho výskyty najdou (61).

(60) **Soud** (Nsg|Asg) vynesl **rozsudek** (Nsg|Asg). (vl)

(61) Od otce nebylo správné, že **vynesl soud** a přitom nevěděl, oč ve sporu jde. (SYNv10)

Mezi APKS se vyskytují samozřejmě i případy nevlastní homonymie. Příkladem může být tvar *víru*, který označuje Asg substantiva *víra*, ale zároveň i Gsg substantiva *vír*. Do APKS se slovesem *ztratit/ztrácet* nicméně vstupuje jen substantivum *víra*. Dalším příkladem je tvar *růst*, který může být Nsg i Asg substantiva *růst* stejně jako infinitivem slovesa, součástí APKS *znamenat růst* je ale pouze substantivum.

Seznamy založené na podobné bázi dvojic (příp. trojic) slov, která se vyskytují společně, již součástí pravidlové disambiguace systému LanGr jsou a jejich aplikace dosahuje velmi dobrých výsledků; navíc je jejich využití poměrně výpočetně nenáročné. Pro tvorbu seznamů se využívá např. zmíněná databáze víceslovných lexikálních jednotek LEMUR (Petkevič et al. 2020). APKS samým nicméně doposud nebyla věnována systematická pozornost.

5.2 Data a jejich zpracování

Veškerá jazyková data, se kterými pracuji, pocházejí z korpusu SYNv10 (Hnátková et al. 2014), pokud není uvedeno jinak. Jedná se o momentálně největší psaný korpus dostupný na ČNK (4882 milionů slov bez interpunkce). Korpus je referenční, ale není reprezentativní – tvoří jej téměř výhradně publicistické texty. To může ovlivňovat obsažený podíl APKS, protože dle předchozích výzkumů jsou APKS doménou zejména odborných textů (srov. Jelínek 2002; 2003; 2007; Macháčková 1979; 1983; 1994).¹⁰⁶

Práci s korpusem řady SYN ovlivňuje, že jde o velmi rozsáhlý zdroj. Tato skutečnost se neodráží jen ve velikosti výsledných dat; je třeba počítat s tím, že zpracování samotného dotazu

¹⁰⁶ Přestože zdrojové texty korpusu SYNv10 pocházejí téměř výhradně z publicistických zdrojů, u první mnou provedené případové studie pro APKS s kategoriálním slovesem *chovat/chovávat* připadá zhruba 16 % výskytů na beletrii, 10 % na oborovou literaturu a jen zbylých 76 % na texty publicistické. Toto pozorování je založené na korpusu, který není referenční, nelze ho tedy vztáhnout na užívání APKS v jazyce obecně, je nicméně zajímavé, že rozložení výskytů nekopíruje poměrově zastoupení žánrů v korpusu.

zabere i deset minut a příprava značného množství dat k exportu další vyšší jednotky minut. Jelikož se hledání a export v rámci jedné studie opakují mnohokrát,¹⁰⁷ může výzkumník čekáním na data v součtu strávit i dlouhé hodiny. Je-li pravděpodobné, že se budeme k výsledkům vracet, je praktické si uložit trvalý odkaz na výsledky vyhledávání – netrávíme čas opětovným zpracováváním dotazu a zároveň máme stejné pořadí dat jako u původního dotazu (korpus výsledky pokaždé náhodně promíchá).

Používané vyhledávací rozhraní KonText (Machálek 2014; 2020) má i další limity, na které jsem narazila až během samotné analýzy:

- bohužel není možné dané sloveso odlišit od tvaru se zvrtným zájmenem *se*¹⁰⁸ (např. *chovat* vs. *chovat se*); do budoucna by bylo přínosné, kdyby byla informace o tom, že je sloveso zvrtné, již součástí sublemmatu nebo verbtagu
- některá nastavení (např. volba *Zobrazení* → *KWIC/Věta*) se při zadání nového dotazu vrátí zpět na výchozí hodnotu a musí se u každého nového hledání opětovně nastavovat
- nelze dodatečně vyfiltrovat kolokace (např. na základě slovního druhu) ani k nim zobrazit doplňující informace (např. rozlišit slovní druh, pokud filtruji podle tvaru slova, který je homonymní), přestože v korpusu tyto informace dostupné jsou
- občas dochází k problémům s exportem dat (zcela chybějící nebo špatné hlavičky, omezené počty řádků apod.), který navíc v některých případech ani neproběhne kvůli vytížení a časovému limitu serveru a je třeba ho provádět opakovaně

Jako nejfunkčnější se nakonec ukázalo omezit použití KonTextu na nejnütnější úkony: zadání dotazu a stažení dat ve formátu celých vět obsahujících KWIC doplněných o lemma a morfologický tag u každého slova. To mi umožnilo stáhnout obsáhlé množství dat naráz, kdykoliv se k nim jednoduše vracet a pracovat s nimi dle potřeby. Tak objemný dataset by se nicméně jen stěží zpracovával v tabulkovém editoru, výhodnější bylo vytvořit vlastní skripty v programovacím jazyce Python.

Skriptování zjednodušilo vyhodnocování dat a snížilo množství chyb potenciálně způsobených jejich manuálním zpracováváním, u anotace zajistilo konzistentnost a zamezilo vynechání některého z kroků a konečně u analýzy umožnilo snadno přepočítat výsledky, pokud

¹⁰⁷ Např. pro druhou případovou studii z oddílu 5.3.2 bylo potřeba téměř 100 dotazů.

¹⁰⁸ Až po provedení první empirické studie jsem zjistila, že alespoň částečným řešením by mohla být možnost „Specifikovat kontext“, která umožňuje vybrat nebo vyloučit lemma (např. *se*) v rámci rozsahu kontextu a je dostupná na úvodní obrazovce při zadávání korpusového dotazu přes KonText.

se vstupní data změnila. Ulehčilo mi také perné chvílky, kdy jsem zjistila, že některé informace jsem se zdrojovými daty opomněla exportovat a zároveň jsem data již stihla zpracovat – stačilo „jen“ dopsat skript, který spojil stará data (ale opatřená anotací) s novými a přidal chybějící verbtagy. Tento přístup mi navíc umožnil připravit si pro anotování vhodnější prostředí: zobrazovat věty k anotaci jednu po druhé, přidávat k nim pouze potřebné množství informací, kontrolovat vstupy a jiné. Detailnější poznámky k jednotlivým skriptům jsou doplněny v podkapitole 5.4.

Zde je nutné podotknout, že bez znalosti programování by v praxi nebylo možné takovou studii vůbec realizovat. Ruční anotace dat v tabulce stažené z webového rozhraní by byla výrazně časově náročnější, ale stále proveditelná, byť se všemi údaji, jako je lemma nebo morfologický tag, značně nepřehledná. Samotné vyhodnocování výsledků, vyhledávání vzorců chování a vůbec práce s daty napříč různými tabulkami a složkami by ale byla již nad lidské síly. Ukazuje se, nakolik je taková znalost důležitou výbavou moderního lingvisty, který chce pracovat s rozsáhlými jazykovými daty.

5.3 APKS s predikativním jménem v akuzativu

Zaměřím se nyní na APKS s bezpředložkovým akuzativem, stejně to činí jako většina existujících studií věnovaných APKS. Tyto konstrukce tvoří centrální typ APKS a jsou nejfrekventovanější; zároveň, jak již bylo nastíněno v části věnované morfologické homonymii (5.1), je jedním z nejpálčivějších problémů homonymie nominativu a akuzativu. Ta se může dále promítat i na rovinu syntaktickou (podmět vs. předmět); homonymní tvar ve spojení s předložkou je z hlediska disambiguace o něco menším oříškem, i když samozřejmě existují předložky rekčně i slovnědruhově víceznačné (např. předložka *za* se pojí s genitivem, akuzativem i instrumentálem).

V rámci této podkapitoly provedu pilotní studii APKS s kategoriálním slovesem *chovat/chovávat*, ve které se snažím analyzovat zejm. pravidelnosti v povrchové struktuře, které budu moci dále využít. V druhé případové studii se zaměřím na ověření těchto pozorování a třetí část je věnována jejich implementaci pro pravidlovou disambiguaci LanGr.

5.3.1 Případová studie: *chovat/chovávat*

Pro první sondu do jazykových dat jsem zvolila kategoriální sloveso *chovat/chovávat*. Důvodem byl předpoklad, že množství výskytů bude při manuální analýze lépe obsáhnutelné než např. spojení se slovesem *dát/dávat*, která jsou mnohanásobně frekventovanější. Zároveň má

zvolené sloveso výhodu poměrně jasné hranice mezi predikativním a kategoriálním významem, aniž by bylo nutné provádět testy nad každou jednotlivou větou. Navíc se jedná o sloveso s pouze jedním obligatorním doplněním (predikativním jménem), ostatní jsou pouze fakultativní, spadá tedy mezi nejrozšířenější skupinu kategoriálních sloves podle valenčního rámce (Kettnerová a Lopatková 2017, s. 21).

VALLEX 4.0¹⁰⁹ u slovesa *chovat/chovávat* uvádí šest různých významů a valenčních rámců. První a nejčastější jsou významy predikativní – držet v náruči (*Chová dítě v náruči.*) a pěstovat (*Chová na farmě prasata.*). Dále slovník zmiňuje *chovat/chovávat* jako součást komplexních predikátů (APKS) (*Proč k němu chová stále stejné něžné city?*). Nakonec slovník uvádí i významy idiomatické – čekat dítě (*Bude chovat.*), starat se / opatrovat (*Chová někoho jako v bavlnce.*) a vážit si (*Komu je přes dvacet, toho chovají v úctě.*). Poslední jmenovaný význam by mohl patřit mezi APKS, ale kolektiv autorů stojící za anotací APKS ve VALLEXu považuje předložkovou alternativu k existujícímu nepředložkovému APKS za idiomatictější a vyčleňuje ji (*chovat úctu vs chovat v úctě*; viz Kettnerová, Kolářová a Vernerová (2017, s. 213–216)). V této studii nicméně není potřeba stanovovat hranici mezi APKS a jinými spojeními příliš striktně, protože se snažím pokrýt spíše prototypická užití než hraniční příklady.

Níže následuje podrobný popis zvoleného postupu i výsledné analýzy. Mým záměrem bylo, aby byla studie dobře replikovatelná a využitelná pro případný další výzkum, což bohužel není případ všech studií k tématu dostupných. U nových studií jsou velice často (patrně z rozsahových důvodů jednotlivých periodik a konferenčních příspěvků) prezentovány jen výsledky a stručný popis průběhu, ale ani postupy, ani zpracovaná data dostupná až na výjimky nejsou. K obecné práci s daty viz předchozí podkapitola 5.2.

5.3.1.1 Získání seznamu potenciálně predikativních jmen

Tato pilotní studie je pojata co nejotevřeněji (na rozdíl od studie následující, která vytěží již existující materiál), neboť jejím cílem je ke slovesu *chovat/chovávat* doplnit predikativní jména. Základní přehled vznikl sloučením seznamů ze tří různých zdrojů¹¹⁰ (počet substantiv je uveden v závorce, podtrženě vždy substantiva přidaná v dané odrážce oproti předchozímu seznamu):

¹⁰⁹ Dostupné z: <https://ufal.mff.cuni.cz/vallex/4.0/#/lexeme/chova1/0>.

¹¹⁰ Při realizaci případové studie jsem bohužel ještě neměla k dispozici některé seznamy z podkapitoly 2.6. Seznam Macháčkové (1979) by nicméně seznam predikativních jmen doplnil o pouhá 4 další substantiva: *obavy, úmysl, názor a přesvědčení*, ta jsem doplnila při dalším zpracování později já. Seznam Kettnerové a Lopatkové (2017) se kryje s VALLEXem 4.0 (chybí pouze *přátelství*).

- VALLEX 4.0:¹¹¹ *cit, naděje, náklonnost, nedůvěra, nenávist, obdiv, podezření, přátelství, respekt, sympatie, úcta, zášť* (12)
- pracovní verze slovníku VALLEX:¹¹² *cit, důvěra, naděje, náklonnost, nedůvěra, nenávist, nepřátelství, obdiv, podezření, přátelství, respekt, sympatie, úcta, zášť* (14)
- výsledky hledání¹¹³ v elektronické verzi práce Radimského (2010, s. 30, 42, 72): *cit, láska, naděje, náklonnost, úcta, podezření, sympatie, obdiv, nenávist* (9)

Po sloučení do jednoho seznamu získáme 15 predikativních jmen: *cit, důvěra, láska, naděje, nedůvěra, nenávist, nepřátelství, náklonnost, obdiv, podezření, přátelství, respekt, sympatie, úcta, zášť*. Domnívám se, že se zdaleka nejedná o všechna predikativní jména, která se slovesem *chovat/chovávat* mohou APKS vytvářet. Proto se pokusím seznam ještě rozšířit.

V korpusu jsem vyhledala taková spojení, která v rámci jedné věty obsahují lemma *chovat/chovávat* a substantivum v akuzativu a zároveň jsou mezi těmito výrazy maximálně čtyři slova, která nejsou ani interpunkčním znaménkem, ani lemmatem *se* a *si*.¹¹⁴

Výsledky jsem uložila do CSV souboru pro další zpracování. V následujícím kroku jsem z korpusového prohlížeče navíc vyexportovala seznam podstatných jmen, se kterými se sloveso *chovat/chovávat* v daném dotazu v akuzativu pojí. Takový soupis je možné získat setříděním podle frekvence s vlastním nastavením, a to podle atributu *lemma* začínajícího od slova *KWIC nejvíce vpravo*. Jako minimální frekvenční limit jsem zvolila 10 výskytů. Výsledkem je seznam, který obsahuje 541 unikátních substantiv k dalšímu manuálnímu zpracování (v elektronické příloze soubor *prvni_studie_substantiva.csv*).

Substantiva, která nefigurovala v seznamu vytvořeném kombinací VALLEXu a Radimského, bylo potřeba vyhodnotit ručně. Pomohl mi vlastní skript (viz 5.4.1), který mi umožnil procházet jedno substantivum za druhým a zobrazovat si k nim relevantní příklady (tj. výskyty s daným substantivem v KWIC). Využila jsem přitom rozdělení významů jednotlivých lexikálních jednotek ve slovníku VALLEX a při anotaci jsem spojení prisuzovala zjednodušeně buď predikativní, nebo nepredikativní význam. Další testy (viz 2.1.2) jsem prováděla až ex post a

¹¹¹ Dostupné z <https://ufal.mff.cuni.cz/vallex/4.0/#/lexeme/choval/0>.

¹¹² Dostupné z <https://quest.ms.mff.cuni.cz/vallex/> verze 9.9, aktualizace 4. dubna 2022.

¹¹³ Jedná se pouze o fulltextové vyhledávání v elektronické verzi publikace. Radimský málokdy uvádí podrobnější výčet, přesto se v některých případech jedná o spojení, která VALLEX neregistruje.

¹¹⁴ Použitý korpusový dotaz v modu „pokročilý dotaz“: `[lemma="chovat"|lemma="chovávat"] [tag!="Z.*" & lemma!="se" & lemma!="si"]{0,4} [tag="N...4.*"] within <s/>`. Uložené výsledky vyhledávání jsou dostupné na: <https://www.korpus.cz/kontext/view?q=~sSeq8qQe6Uk2>.

jen u potenciálně kategoriálních spojení, což celou anotaci významně zrychlilo. Pro zařazení jména do seznamu pro další analýzu stačil jeden výskyt APKS.

Popsaným způsobem jsem výše uvedený seznam rozšířila na 47 potenciálně predikativních jmen, která se s kategoriálním slovesem *chovat/chovávat* mohou pojít (podtržená slova byla součástí výchozího seznamu): *ambice*, *antipatie*, *cit*, *despekt*, *důvěra*, *iluze*, *láska*, *myšlenka*, *naděje*, *náklonnost*, *názor*, *nedůvěra*, *nechuť*, *nenávist*, *nepřátelství*, *nevraživost*, *obava*, *obdiv*, *obraz*, *odpor*, *ohled*, *opovržení*, *plán*, *pocit*, *podezření*, *postoj*, *potřeba*, *přání*, *přátelství*, *představa*, *předsudek*, *přesvědčení*, *přízeň*, *respekt*, *sen*, *skepse*, *slabost*, *sympatie*, *touha*, *úcta*, *úmysl*, *vášeň*, *víra*, *vztah*, *záměr*, *zášť*, *zloba*.

5.3.1.2 Anotace vět obsahujících hledané sloveso i podstatné jméno

Pořízením seznamu potenciálně predikativních jmen pojících se s kategoriálním slovesem *chovat/chovávat* práce nekončí. Pro samotné poznání struktur je nutné projít jednotlivé věty, posoudit, zda se jedná nebo nejedná o konstrukce s APKS, a srovnat je s čistě predikativními strukturami. Níže shrnu, jak anotace probíhala, a v dalším oddíle se zaměřím na analýzu spojení.

V tomto bodě se bohužel nedá vyhnout poměrně zdlouhavé ruční anotaci jednotlivých vět. Pro každé podstatné jméno ze seznamu (viz předchozí oddíl 5.3.1.1) jsem zkonstruovala vždy dva dotazy obsahující v jedné větě zároveň lemma *chovat/chovávat* a lemma vybraného podstatného jména, přičemž v jednom případě sloveso předchází podstatnému jménu (V-N), v druhém případě je naopak následuje (N-V).¹¹⁵

Při ruční anotaci stažených konkordancí prováděné pomocí malého anotačního skriptu napsaného v jazyce Python (detaily viz 5.4.1) jsem každému KWIC přiřadila jednu z následujících hodnot:

hodnota anotace	druh spojení	příklad
0	jiný druh spojení	Má neustálou potřebu se chovat a mazlit.
1	APKS vhodné k další analýze	Bodové ambice chová především krnovské mužstvo.
3	APKS nevhodné k další analýze (konstrukce, která rozvíjí	Došlo mu, že zradil city , jež k němu Jane chovala .

¹¹⁵ Příklad dvou dotazů na spojení *chovat/chovávat cit* (další dotazy jsou vytvářeny analogicky):
`[lemma="cit"&tag="N...4.*"] [* [lemma="chovat|chovávat"] within <s/>, [lemma="chovat|chovávat"] [* [lemma="cit"&tag="N...4.*"] within <s/>`

	predikativní jméno vedlejší větou), dále označované zkráceně „vztažné“	
--	---	--

Tabulka 7: Hodnoty přiřazené v rámci anotace jednotlivých vět při analýze potenciálních APKS se slovesem *chovat/chovávat*

Spojení označená hodnotou 1 jsou APKS, které se vyskytují v rámci jedné klauze a budou v další části podrobena analýze. Spojení s hodnotou 3 sice mezi APKS patří, ale bylo potřeba je vydělit do samostatné kategorie, protože se jako podklad pro tvorbu pravidel v další části nehodí. Ve většině případů jde o rozvití substantiva vedlejší větou, oba členy nejsou tudíž přítomny v rámci jedné klauze – a pravidlové zachycení podobných konstrukcí je poměrně složité. Hodnotou 0 byla označena zbývající spojení.

Zadanému dotazu odpovídalo více než 18 tisíc spojení, která jsem postupně ručně prošla a provedla jejich anotaci. Tato data tvoří elektronickou přílohu práce (jako příloha *prvni_studie_apks_vztazne.csv* a *prvni_studie_ostatni.csv*), a to nejen kvůli rozsahu. V průběhu analýzy jsem totiž přidala další informace, které umožňují filtrování a řadění podle různých kritérií, přičemž klasická příloha v tištěné nebo PDF verzi takovou funkcionalitu zcela postrádá a data v takové podobě by byla téměř nevytěžitelná.

Jelikož jsem data anotovala sama, tj. bez pomoci dalších anotátorů, se kterými bych mohla porovnat výsledky, jevílo se mi vhodnější jít kvantitativní cestou a opatřit anotaci všechny věty, nicméně pouze jednou. Předpokládala jsem, že případné chyby a nedostatky odhalím při kvalitativní analýze a ověřování jednotlivých tvrzení. To se nakonec ukázalo jako plodný přístup – pokud např. všechny APKS odpovídaly určitému pravidlu a objevilo se jen několik málo výjimek. Obvykle se ukázalo, že se jedná o chybu způsobenou lidským faktorem při anotaci, a bylo relativně jednoduché provést patřičnou korekci. Přes veškerou snahu mohly v anotaci stále zůstat drobné nedostatky, přesto věřím, že jich je minimum a na celkový výsledek mají zanedbatelný vliv.

5.3.1.3 Analýza dat

V tomto pododdíle se zaměřím na to, co je možné z ručně analyzovaných spojení obsahujících sloveso *chovat/chovávat* a predikativní jméno v akuzativu odvodit pro chování struktur s APKS. Záměrem je nasbírat podklady k formulaci obecných pravidel pro rozpoznávání struktur s APKS a zlepšit tak morfologickou disambiguaci prováděnou pravidly, jak bylo nastíněno výše v podkapitole 5.1.

Ráda bych zde zdůraznila, že v této analýze pojem *struktura* striktně odpovídá výsledku hledaného dotazu (KWIC). Nezahrnuje tedy členy spojení, které jsou realizovány mimo rozsah

dotazu, a naopak může zahrnovat tokeny, které členy APKS nejsou. V tomto ohledu jde jen o formální pozorování na rovině povrchové, nikoliv syntaktické.

Přestože všechny dotazy zahrnovaly jak nedokonavé sloveso *chovat*, tak jeho iterativní variantu *chovávat*, kterou VALLEX uvádí jako specifickou podtřídu nedokonavých sloves, v datech se nakonec iterativní varianta *chovávat* reálně vyskytla pouze jednou (62).¹¹⁶ V další analýze tedy mluvím jen o APKS se slovesem *chovat*.

(62) Neodpustil si aférky, leč po rozvodu **chovával náklonnost** k bývalým chotím.

5.3.1.3.1 Obecná charakteristika dat

Z analyzovaných vět téměř polovina tvořila APKS. Rozdělení shrnuje následující Tabulka 8:

pořadí členů	APKS (%)	vztažné (%)	jiné (%)	celkem (%)
N-V	1012 (5,48)	505 (2,74)	5242 (28,4)	6759 (36,62)
V-N	7300 (39,55)	0 (0)	4398 (23,83)	11698 (63,38)
celkem	8312 (45,03)	505 (2,74)	9640 (52,23)	18457 (100,0)

Tabulka 8: Shrnutí analyzovaných APKS a jiných spojení se slovesem *chovat*, procenta se vztahují k celkovému počtu analyzovaných spojení¹¹⁷

Ze sondy vyplývá, že APKS s kategoriálním slovesem *chovat* se vyskytují zejména v nepříznačkovém slovosledu (tj. v pořadí V-N), viz shrnutí jen pro případy APKS:

typ APKS	celkem (%)
N-V	1012 (11,48)
N-V vztažné	505 (5,73)
V-N	7300 (82,79)
celkem	8817 (100)

Tabulka 9: Shrnutí analyzovaných APKS se slovesem *chovat*

V datech se objevují APKS v rámci jedné klauze (příklady (63) a (64) níže, tučně zvýrazněno hledané KWIC), které mě zajímají primárně, protože jejich struktura může být použita jako podklad pro vytvoření pravidel pro disambiguaci. Jak již bylo zmíněno v předchozím pododdíle, do samostatné kategorie („vztažné“) vyčleňuji spojení, která mezi APKS patří, ale vyjádřené členy jsou rozdělené mezi několik klauzí, tj. predikativní jméno je rozvíté vedlejší větou

¹¹⁶ Sloveso *chovávat* má ostatně velmi řídké zastoupení v celém korpusu SYNv10 – vyskytuje se pouze 84krát.

¹¹⁷ Data v této tabulce i v tabulkách následujících jsou na základě anotovaných dat dynamicky generována skriptem, bohužel není v mých možnostech správně oddělit tisíce mezerou.

(65) (sem by patřilo např. i nahrazení podstatného jména zájmenem jako v (66)). Tato spojení samozřejmě v kvantitativním pohledu mezi APKS řadím, nicméně pro formulaci pravidel jsou nevhodná, jelikož jsou velmi heterogenní. V předposledním sloupci Tabulky 8 („jiné“) shrnuji výskyty jiných konstrukcí (67) a (68).

(63) Bodové **ambice chová** především krnovské mužstvo.

(64) K šéfovi zástupci **chovám nejvyšší respekt**.

(65) Došlo mu, že zradil city, jež k němu Jane chovala.

(66) Mužstvo mělo vysoké **ambice**. **Ty** v sobě **chovalo**, protože v sezóně nedostalo ani jeden gól. (vl)

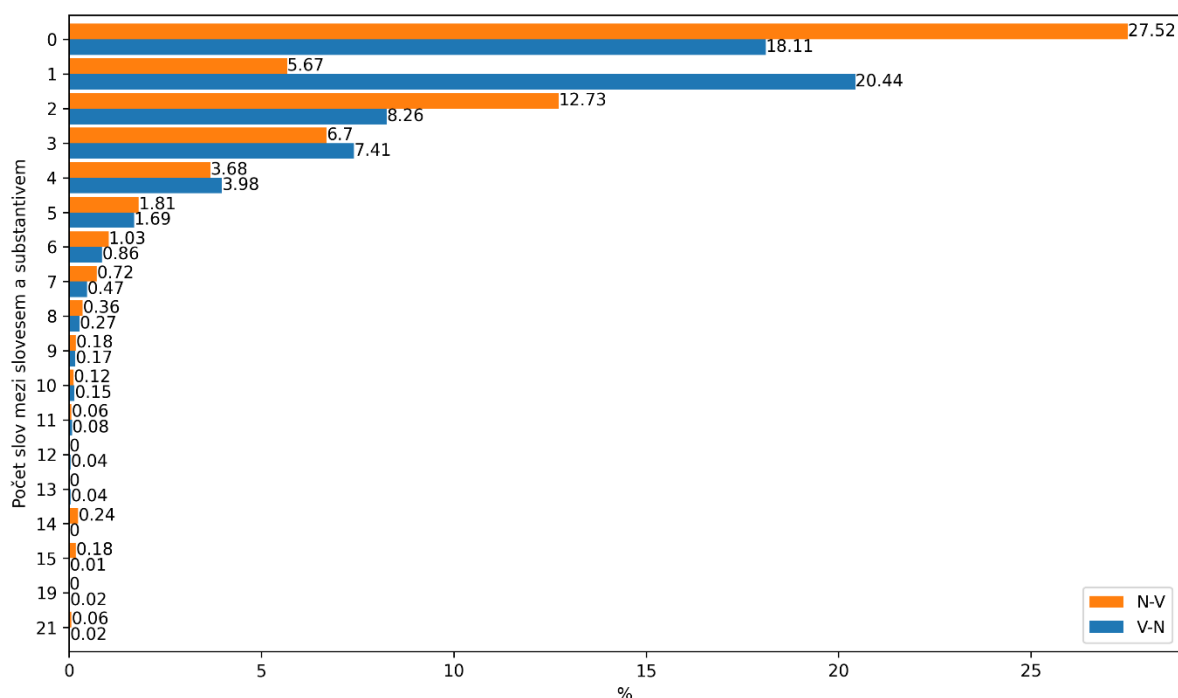
(67) Má neustálou **potřebu se chovat** a mazlit.

(68) V roce 1579 mu Rudolf II. udělil statut císařského dvorního hřebčína, který **choval koně pro potřeby dvora**.

Během analýzy jsem narazila i na jeden případ, kdy v rámci jazykové hry sloveso *chovat* vystupuje současně ve dvojí funkci – jak predikativní, tak kategoriální:

(69) **Chová** krátkosrstou **kolii** a přívětivou **skepsi** k sobě i ke světu.

Dále se budu zabývat pouze APKS, ve kterých jsou oba členy spojení přítomny v téže klauzi (tj. příklady (63) a (64) výše). Nejprve se zaměříme na délku těchto spojení. Z Grafu 1 níže (zdrojová data viz Příloha 2) plyne, že je pro tyto struktury typické velmi těsné spojení. Vzdálenost s hodnotou 0 znamená, že mezi členy spojení není žádné další slovo, s rostoucí hodnotou počet tokenů mezi členy přibývá.



Graf 1: Počet slov mezi kategoriálním slovesem a predikativním jménem, procentuální srovnání mezi N-V a V-N

Případy (např. (70) a (71)) s vysokým počtem slov mezi členy spojení nejsou časté, přesto nepůsobí nepatříčně.

(70) **Naději**, že by Roman Hubník své hostování v mateřské Olomouci prodloužil i pro jarní část sezony, v Sigmě příliš velkou **nechovali**.

(71) Avšak na základě těchto zkušeností se mi nikdo nemůže divit, že **chovám** k novým metodám, které významně zasahují do samé podstaty živých organismů, a k tvrzením odborníků odůvodněnou a nikoli iracionální **nedůvěru**.

Mnohem zajímavější informace ale získáme, zaměříme-li se na jednotlivá lemmata substantiv. Přestože rozsáhlé tabulky obvykle do těla prací nepatří, domnívám se, že zařazení až do přílohy by bylo v tomto případě kontraproduktivní a následná analýza by byla pro čtenáře nepřehledná, tudíž je v takových případech vkládám rovnou.

Tabulka 10 níže obsahuje všechna zkoumaná potenciální predikativní jména seřazená podle procentuální frekvence výskytů v APKS (včetně hodnoty „vztažné“, kdy jsou členy rozděleny do dvou klauzí) – sloupeček zcela vpravo. Hvězdičkou jsou označena lemmata, která zmiňuje VALLEX nebo Radimský (viz oddíl 5.3.1.1). Data rozdělená do dvou tabulek podle pořadí členů APKS (V-N a N-V) jsou v tabulkách přiložených v Příloze 3 a Příloze 4.

lemma	celkem	jiné (%)	vztažné (%)	APKS (%)	APKS + vztažné (%)
zášť*	351	17 (4,84)	2 (0,57)	332 (94,59)	334 (95,16)
sympatie*	719	71 (9,87)	30 (4,17)	618 (85,95)	648 (90,13)
nedůvěra*	295	32 (10,85)	24 (8,14)	239 (81,02)	263 (89,15)
obdiv*	471	59 (12,53)	21 (4,46)	391 (83,01)	412 (87,47)
antipatie	62	8 (12,9)	1 (1,61)	53 (85,48)	54 (87,1)
náklonnost*	298	41 (13,76)	16 (5,37)	241 (80,87)	257 (86,24)
úcta*	1643	242 (14,73)	82 (4,99)	1319 (80,28)	1401 (85,27)
skepse	41	7 (17,07)	2 (4,88)	32 (78,05)	34 (82,93)
nenávisť*	376	65 (17,29)	14 (3,72)	297 (78,99)	311 (82,71)
cit*	880	153 (17,39)	68 (7,73)	659 (74,89)	727 (82,61)
naděje*	834	155 (18,59)	12 (1,44)	667 (79,98)	679 (81,41)
despekt	31	6 (19,35)	0 (0,0)	25 (80,65)	25 (80,65)
nepřátelství*	74	15 (20,27)	2 (2,7)	57 (77,03)	59 (79,73)
respekt*	1574	356 (22,62)	29 (1,84)	1189 (75,54)	1218 (77,38)
opovržení	41	13 (31,71)	4 (9,76)	24 (58,54)	28 (68,29)
nechuť	35	12 (34,29)	3 (8,57)	20 (57,14)	23 (65,71)
nevraživost	35	13 (37,14)	2 (5,71)	20 (57,14)	22 (62,86)
zloba	33	13 (39,39)	1 (3,03)	19 (57,58)	20 (60,61)
předsudek	86	34 (39,53)	15 (17,44)	37 (43,02)	52 (60,47)
láska*	640	266 (41,56)	59 (9,22)	315 (49,22)	374 (58,44)
důvěra*	609	293 (48,11)	21 (3,45)	295 (48,44)	316 (51,89)
úmysl	98	48 (48,98)	4 (4,08)	46 (46,94)	50 (51,02)
podezření*	425	224 (52,71)	2 (0,47)	199 (46,82)	201 (47,29)
vášně	95	53 (55,79)	7 (7,37)	35 (36,84)	42 (44,21)
víra	133	75 (56,39)	1 (0,75)	57 (42,86)	58 (43,61)
přátelství*	98	56 (57,14)	8 (8,16)	34 (34,69)	42 (42,86)
iluze	139	81 (58,27)	3 (2,16)	55 (39,57)	58 (41,73)
odpor	307	185 (60,26)	5 (1,63)	117 (38,11)	122 (39,74)
presvědčení	125	78 (62,4)	0 (0,0)	47 (37,6)	47 (37,6)
přízeň	93	60 (64,52)	10 (10,75)	23 (24,73)	33 (35,48)
ambice	125	81 (64,8)	0 (0,0)	44 (35,2)	44 (35,2)
touha	151	101 (66,89)	0 (0,0)	50 (33,11)	50 (33,11)
vztah	1314	1013 (77,09)	14 (1,07)	287 (21,84)	301 (22,91)
slabost	72	58 (80,56)	1 (1,39)	13 (18,06)	14 (19,44)
myšlenka	261	217 (83,14)	1 (0,38)	43 (16,48)	44 (16,86)
sen	255	219 (85,88)	1 (0,39)	35 (13,73)	36 (14,12)
přání	166	144 (86,75)	2 (1,2)	20 (12,05)	22 (13,25)
představa	526	457 (86,88)	3 (0,57)	66 (12,55)	69 (13,12)
záměr	134	118 (88,06)	2 (1,49)	14 (10,45)	16 (11,94)
obava	429	380 (88,58)	2 (0,47)	47 (10,96)	49 (11,42)

postoj	362	327 (90,33)	3 (0,83)	32 (8,84)	35 (9,67)
pocit	1844	1670 (90,56)	24 (1,3)	150 (8,13)	174 (9,44)
obraz	208	194 (93,27)	2 (0,96)	12 (5,77)	14 (6,73)
plán	222	212 (95,5)	1 (0,45)	9 (4,05)	10 (4,5)
ohled	167	162 (97,01)	0 (0,0)	5 (2,99)	5 (2,99)
názor	956	938 (98,12)	1 (0,1)	17 (1,78)	18 (1,88)
potřeba	624	621 (99,52)	0 (0,0)	3 (0,48)	3 (0,48)

Tabulka 10: Predikativní jména tvořící APKS se slovesem *chovat*. Seřazeno podle procentuálního výskytu APKS + vztažné. Hvězdičkou jsou označena spojení uvedená ve VALLEXu a Radimském.

Z výše uvedených dat je zřejmé, že některá substantiva ve spojení se slovesem *chovat* tvoří APKS téměř vždy (*zášť, nedůvěra, sympatie, obdiv...*), jiná naopak velmi zřídka až téměř nikdy (*obraz, plán, ohled, názor, potřeba...*) a patrně bylo lepší je do analýzy vůbec nezařazovat.

Predikativní jména uvedená ve VALLEXu a Radimském obsadila přední příčky – více než 75 % všech jejich výskytů se řadí mezi APKS (menší podíl má jen *láska, důvěra, podezření a přátelství*). Z mnou dohledaných konstrukcí se mezi ně vklínila pouze substantiva *antipatie, skepse* a *despekt*, která ani jeden ze zdrojů neuvádí.

Zaměříme se nyní krátce na první predikativní jméno z tabulky výše, totiž substantivum *zášť*. Při pohledu na jednotlivé výskyty v pořadí V-N (viz Příloha 3) zjistíme, že není-li spojení *chovat zášť* součástí APKS, jedná se buď o záměnu za sloveso *chovat se* (72), nebo jsou slova *chovat* a *zášť* součástí rozdílných vět (73), případně je sloveso *chovat* ve větě dvakrát a chamtivý (*greedy*) regulární výraz¹¹⁸ použitý v korpusovém dotazu se snaží vybrat nejdelší KWIC vyhovující podmínce (74), aby obsáhl co nejdelší spojení, a přitom vybere „nesprávné“ sloveso. V některých případech se kategoriální sloveso pojí s těsně následujícím predikativním jménem a k vyhledání korpusovým dotazem dojde, protože je kontext příliš široký a ve stejné větě se vyskytuje i aktuálně dotazované substantivum (75). Takové případy analyzují v rámci konstrukce se správným predikativním jménem.

(72) Pokud bude Berlín platit účty a diktovat zbytku Evropy, jak se má **chovat, riskuje, že si vypěstuje destruktivní nacionalistickou zášť** vůči Německu.

¹¹⁸ Příště by možná bylo vhodné problému rovnou předejít, a to vyloučením lemmatu stejného slovesa mezi hledanými členy. Dotaz by pak vypadal následovně: [lemma="cit"&tag="N...4.*"] [lemma!="chovat|chovávat"]* [lemma="chovat|chovávat"] within <s/>, [lemma="chovat|chovávat"] [lemma!="chovat|chovávat"]* [lemma="cit" & tag="N...4.*"] within <s/>.

- (73) Kristinina láska ke Karlu Gustavovi se ale nezměnila v **zášť** nebo něco podobného: **chovala** ho i nadále v úctě a zacházela s ním vlídně.
- (74) Choval se sice bezproblémově, ale v duši proti svému živiteli choval velikou zášť, protože mu záviděl jeho majetek.
- (75) Usmála jsem se na něj, abych mu dala najevo, že k němu **nechovám nepřátelství, ale taky abych zjistila, jestli on cítí zášť** ke mně.

Podobné případy se opakují napříč dalšími predikativními jmény uvedenými v tabulce. Část případů by možná odfiltroval mnohem detailnější dotaz, dalším poměrně jednoduchým řešením by bylo zkrátit délku KWIC. Zpočátku jsem nicméně chtěla obsáhnout co nejpestřejší paletu spojení i za cenu tohoto druhu chyb. Čím delší totiž KWIC je, tím více jiných spojení se do výsledků dostane (viz strmě klesající zastoupení APKS v závislosti na počtu slov mezi predikativním jménem a kategoriálním slovesem v Grafu 1 výše).

U případů APKS s obráceným, příznakovým slovosledem N-V (viz Příloha 4) musíme vést v patnosti, že jich je výrazně méně než APKS se slovosledem V-N. Zároveň je nutné případy, u kterých se sloveso a substantivum vyskytuje v rámci jedné klauze (76), uvažovat společně s případy vedlejší věty (77) rozvíjející predikativní jméno. Do této kategorie (v tabulkách označované jako „vztažné“) by patřily i případy, kdy je podstatné jméno zastoupené zájmenem nebo je součástí nevyjádřeného podmětu (78). Můj korpusový dotaz je nicméně nepokrývá systematicky, protože neobsáhl případy, kdy se jednotlivé členy APKS nacházejí přes hranici věty (79).

- (76) Ke Catherine Larssonové a jejím dětem žádnou **zášť** **nechoval**.
- (77) Ani rozpad jejich vztahu však nesnížil **sympatie, které k němu prezident chová**.
- (78) Já k vašim lidem nechovám žádnou **zášť** a **většina z nich ji nechová** ani ke mně.
- (79) Mezi sezvanými odborníky panoval vzájemný **respekt**. **Ten** k nim **chovala** ostatně i široká veřejnost, která [...]. (vl)

5.3.1.3.2 Kolokace

Analýza ukázala existenci mnohých dvojic substantiva a slovesa, které APKS vytvářejí v naprosté většině případů, což na první pohled poukazuje na silný vzájemný vztah těchto slov. Nabízí se proto vyhledat kolokace¹¹⁹ lemmatu *chovat*. Při filtrování lemmatu v kontextu tří slov před hledaným slovesem a tří slov po něm se jako nejvhodnější jeví kolokační míra logDice.

¹¹⁹ Prozkoumat úzkou kolokační vazbu slovesa na predikativní jméno navrhuji jak Čermák (1974), tak Radimský (2017a).

Ta kolokace zachycuje vyrovnaně a není závislá na velikosti korpusu, protože vychází z poměrování frekvence jednotlivých slov a frekvence kolokace jako takové (Čermák 2017b), což ovšem favorizuje frekventovaná spojení (*úcta, respekt*).

Výstupy z hledání kolokací bohužel nelze v KonTextu filtrovat pomocí slovního druhu (POS). Obsahují tak velké množství šumu, jelikož přední příčky zabírají zejména adverbia (*slušně, agresivně, zodpovědně, jednak* apod.), zatímco nás zajímají pouze substantivní kolokáty. Z toho důvodu je mezi prvními 300 nejčastějšími kolokacemi pouze sedm APKS (*úcta, respekt, sympatie, cit, obdiv, naděje, zášť*; řazeno podle síly asociační míry). Tento nedostatek jsem opravila pomocí vlastního skriptu, který k jednotlivým lemmatům naivně¹²⁰ doplní slovní druh z konkordancí, které jsem analyzovala dříve. Data stažená z korpusu a obohacená o slovní druh a příp. informaci, zda substantivum tvoří APKS, jsou součástí elektronické přílohy jako *prvni_studie_kolokace.csv*.

Díky skriptu je možné vyfiltrovat kolokace se substantivy (pro představu viz Tabulka 11), což mezi prvních 300 výsledků posune ještě substantiva *ohled, láska, nenávisť, vztah, nedůvěra, náklonnost, důvěra, představa* a *podezření*. Přední příčky zabírají substantiva pojící se s predikativním významem slovesa (*chovat zvíře*), která označující nejrůznější zvířata. Abychom tomu předešli, museli bychom se např. omezit pouze na abstraktní podstatná jména, což ale v jiných případech může být kontraproduktivní (srov. *dát/vstřelit branku*), navíc se abstrakta v korpusu neznačí, bylo by tedy nutné nejdříve sestavit jejich seznam.

pořadí	pořadí bez ohledu na POS	lemma	frekvence	MI-score	t-score	logDice
1	5	zvíře	7 132	7,086	83,83	7,661
2	6	zoo	4 107	7,699	63,777	7,592
3	15	ovce	2 094	8,543	45,638	7,1
4	17	úcta	2 124	8,285	45,939	7,073
5	19	prase	2 006	7,964	44,609	6,944
6	21	kráva	1 908	7,922	43,501	6,878
7	30	slepice	1 493	8,768	38,551	6,703
8	36	pes	3 625	5,929	59,22	6,568
9	37	skot	1 322	8,465	36,256	6,514
10	39	kůň	2 672	6,112	50,944	6,509

Tabulka 11: 10 nejčastějších substantivních kolokátů slovesa *chovat* podle míry logDice

¹²⁰ V případě synonymních lemmat skript vybere to, které je v datech ze zpracovávaných konkordancí první. Věřím, že omezím-li se hledání na podstatná jména, toto zjednodušení nehraje roli.

Zjištěnou hodnotu kolokační míry logDice nyní doplníme k seznamu potenciálních predikativních jmen z Tabulky 10. Zároveň jsem přidala i informaci o pořadí v rámci substantivních kolokací zmíněných výše (sloupeček „pořadí N kolokace“). Intuitivní by patrně bylo předpokládat, že pořadí predikativních jmen seřazených podle hodnoty logDice bude alespoň na začátku kopírovat pořadí toho, jak často tato jména tvoří APKS (např. dvojice *chovat zášť* se jako APKS vyskytuje v 95 %), protože to naznačuje silnou vazbu těchto dvou slov. Přesto tomu tak není. Důvodem je, že kolokační míry počítají (v různé míře) i s jinými spojeními či frekvencí daných slov, zatímco tabulka níže vychází pouze z jednoho přesně definovaného kontextu.

lemma	celkem	APKS + vztažné (%)	jiné (%)	logDice	logDice N pořadí
zášť*	351	334 (95,16)	17 (4,84)	4,351	82.
sympatie*	719	649 (90,26)	70 (9,74)	5,42	30.
nedůvěra*	295	263 (89,15)	32 (10,85)	3,812	154.
obdiv*	471	412 (87,47)	59 (12,53)	4,49	70.
antipatie	62	54 (87,1)	8 (12,9)	1,755	935.
náklonnost*	298	257 (86,24)	41 (13,76)	3,739	170.
úcta*	1643	1401 (85,27)	242 (14,73)	7,073	4.
skepse	41	34 (82,93)	7 (17,07)	0,77	1831.
nenávisť*	376	311 (82,71)	65 (17,29)	4,008	126.
cít*	880	727 (82,61)	153 (17,39)	5,318	33.
naděje*	834	680 (81,53)	154 (18,47)	4,465	74.
despekt	31	25 (80,65)	6 (19,35)	2,269	627.
nepřátelství*	74	59 (79,73)	15 (20,27)	1,798	907.
respekt*	1574	1218 (77,38)	356 (22,62)	6,308	15.
opovržení	41	28 (68,29)	13 (31,71)	1,5	1138.
nechut'	35	23 (65,71)	12 (34,29)	0,629	1975.
nevraživost	35	22 (62,86)	13 (37,14)	0,732	1863.
zloba	33	20 (60,61)	13 (39,39)	0,786	1808.
předsudek	86	52 (60,47)	34 (39,53)	2,11	721.
láska*	640	374 (58,44)	266 (41,56)	4,051	116.
úmysl	98	51 (52,04)	47 (47,96)	2,212	659.
důvěra*	609	316 (51,89)	293 (48,11)	3,737	171.
podezření*	425	201 (47,29)	224 (52,71)	3,353	250.
vášeň	95	42 (44,21)	53 (55,79)	1,611	1036.
víra	133	58 (43,61)	75 (56,39)	2,423	548.
přátelství*	98	42 (42,86)	56 (57,14)	1,505	1132.
iluze	139	58 (41,73)	81 (58,27)	2,346	589.
odpor	307	122 (39,74)	185 (60,26)	3,064	306.
přesvědčení	125	47 (37,6)	78 (62,4)	2,171	690.
přízeň	93	33 (35,48)	60 (64,52)	1,563	1081.

ambice	125	44 (35,2)	81 (64,8)	1,688	986.
touha	151	50 (33,11)	101 (66,89)	1,922	830.
vztah	1314	301 (22,91)	1013 (77,09)	3,977	133.
slabost	72	14 (19,44)	58 (80,56)	0,478	2170.
myšlenka	261	44 (16,86)	217 (83,14)	2,05	747.
sen	255	36 (14,12)	219 (85,88)	2	781.
přání	166	22 (13,25)	144 (86,75)	2,069	739.
představa	526	69 (13,12)	457 (86,88)	3,414	237.
záměr	134	16 (11,94)	118 (88,06)	1,242	1358.
obava	429	49 (11,42)	380 (88,58)	1,874	858.
postoj	362	35 (9,67)	327 (90,33)	1,497	1141.
pocit	1844	174 (9,44)	1670 (90,56)	2,554	500.
obraz	208	14 (6,73)	194 (93,27)	1,188	1400.
plán	222	10 (4,5)	212 (95,5)	0,805	1784.
ohled	167	5 (2,99)	162 (97,01)	4,072	114.
názor	956	18 (1,88)	938 (98,12)	2,531	506.
potřeba	624	3 (0,48)	621 (99,52)	2,975	333.

Tabulka 12: Substantiva tvořící APKS se slovesem *chovat* seřazená podle procentuální četnosti spojení (sloupec „APKS + vztažné (%)“) a doplněná o hodnotu logDice a pořadí v seznamu této kolokační míry. Hvězdičkou jsou označena lemmata z VALLEXu a Radimského.

Možná by bylo zajímavé vyzkoušet opačný postup, tj. vyjít při hledání kolokací od predikativního jména. Krátkou sondou v korpusu SYN2020 pro substantivum *zášť*, které se slovesem *chovat* utváří APKS v 95 % případů, získáme jen poměrně omezenou paletu sloves (v závorce hodnota logDice, tučně kategoriální slovesa): *pociťovat* (6,1), ***chovat*** (5,84), ***projevovat*** (5,19), *cítit* (4,52), ***nosit*** (3,17), ***zapomenout*** (2,4), *mít* (-1,51)¹²¹ a *být* (-2,88). Domnívám se, že využití kolokačních měr v tomto směru může být pro hledání členů APKS příhodnější, ale pro potvrzení této domněnky by bylo nutné provést detailnější analýzu. Pro hlavní záměr práce, tj. zlepšení morfologické disambiguace, si nicméně vystačíme s existujícími seznamy APKS, které pokrývají prominentní část těchto spojení.

5.3.1.3.3 Struktury spojení

Díky zvolenému způsobu zpracování dat¹²² je možné za využití dalšího skriptu anotované věty abstrahovat na jednotlivé slovní druhy spojení. Ty pak lze vyhodnotit a nabytou znalost dále využít pro morfologickou disambiguaci, jak bylo nastíněno výše (viz 5.1).

¹²¹ Záporné hodnoty značí negativní asociaci – vzájemné odpuzování.

¹²² Stažení zpracovávaných konkordancí včetně morfologických tagů a lemmat.

Jako *strukturu* označuji sekvenci všech slov od kategoriálního predikátu po predikativní jméno (a naopak), struktura zde odpovídá KWIC v rámci používaných korpusových dotazů a nebere ohled na valenci spojení ani na členy stojící před KWIC nebo po něm. Minimální podoba takové struktury jsou dva členy APKS – predikativní jméno a kategoriální sloveso, které vytvářejí strukturu V-N, potažmo N-V.

K informacím o strukturách jsem se dostala skrze vlastní skript (viz 5.4.1), který mi pomohl sestavit podklady pro další analýzu. Na základě hodnoty první a druhé pozice ze morfologického tagu, tj. slovního druhu (POS) a případně i jeho bližšího určení (2. pozice v tagsetu ČNK), jsem si vygenerovala několik tabulek reprezentujících struktury.

Celkové množství 18 500 analyzovaných spojení lze namapovat na více než 8000 různých sekvencí slovních druhů, variabilita sekvencí je tedy obrovská. Veškeré APKS přitom reprezentuje pouhých 800 sekvencí (10 %), přestože APKS představují téměř polovinu z analyzovaných spojení (cca 9000 spojení). Pro zlepšení morfologické anotace nyní potřebujeme najít co nejfrekventovanější struktury, které se vyskytují výhradně v konstrukcích s APKS; cílem je získat co největší pokrytí (*recall*) a zároveň zachovat 100% úspěšnost (*precision*) budoucích pravidel.

Začnu tím, že data setřídím frekvenčně podle absolutního množství APKS v dané konstrukci. Vznikne Tabulka 13 níže, ze které vyplývá, že první dvě struktury (V-N a V-A-N) pokryjí cca 50 % všech analyzovaných spojení. Prvních 26 z cca 800 struktur pokrývá dokonce přes 80 % výskytů APKS s kategoriálním slovesem *chovat* v analyzovaných datech. V další analýze se zaměřím na těchto 26 spojení, protože ostatní struktury se vyskytují pouze v jednotkách případů, a proto pro ně není efektivní formulovat pravidla. Ve všech 26 případech platí, že kategoriální sloveso a predikativní jméno od sebe nejsou příliš vzdálené – nejdelší ze zmínovaných konstrukcí má šest členů, tj. pouze čtyři vložená slova. I z tohoto důvodu jde tedy o ideální materiál pro tvorbu pravidel s cílem zlepšit značkování.

POS	příklad	APKS (%)	APKS kumulativní %	jiné
V-N	Víš, jaké k tobě chová city .	2132 (25,65)	25,65	0
V-A-N	I já jsem choval literární ambice .	1985 (23,88)	49,53	0
N-V	Naději choval do poslední chvíle.	456 (5,49)	55,02	88
V-P-N	On dnes vůči lidem i věcem nechová žádné iluze .	273 (3,28)	58,3	0

V-R-P-A-N	Chovali k sobě srdečnou nenávist.	237 (2,85)	61,15	0
V-D-A-N	Chovám také velký obdiv k filmům Miloše Formana.	212 (2,55)	63,7	0
V-R-P-N	Nechovám k tobě nepřátelství.	205 (2,47)	66,17	5
V-R-N-A-N	Chovám k mužům hluboký obdiv.	173 (2,08)	68,25	0
V-R-N-N	Mladík chová ke komunistům odpor.	132 (1,59)	69,84	0
N-R-P-V	Přízeň k němu chovala i císařovna Marie Terezie.	115 (1,38)	71,22	0
V-P-A-N	Dnes chovají ta nejtemnější podezření celé tři pětiny Evropanů.	93 (1,12)	72,34	0
V-N-J-N	Vždycky jsem se k nim snažil chovat respekt a úctu.	86 (1,03)	73,37	0
V-R-P-N-A-N	Chová ke svému otci smíšené pocity.	64 (0,77)	74,14	0
V-A-N-J-N	Chovám nesmírný obdiv a úctu ke Karlově univerzitě.	64 (0,77)	74,91	0
N-R-N-V	Obránce Sýkora zášť k zámoří nechová.	56 (0,67)	75,58	2
V-R-A-N-A-N	Také hráči chovají k dvaapadesátiletému stratégovi velký respekt.	51 (0,61)	76,19	0
V-D-N	K Československu jsem chovala odedávna sympatie.	50 (0,6)	76,79	0
V-T-A-N	Fanoušci k němu chovají až božskou úctu.	44 (0,53)	77,32	0
V-A-A-N	Ta v srdci dodnes chová pevné levicové přesvědčení.	39 (0,47)	77,79	0
V-R-N-P-N	Nechovám vůči policii žádnou zášť.	36 (0,43)	78,22	0
V-R-P-N-N	Chovej k té vlajce respekt.	35 (0,42)	78,64	0
V-R-N-N-N	Ten přiznává, že chová pro historii vinnohradnictví slabost.	30 (0,36)	79,0	0
V-A-J-A-N	K ní chovám obdivný a uctivý vztah.	30 (0,36)	79,36	0
V-R-P-P-N	Chovám k němu veškerou úctu.	28 (0,34)	79,7	0
V-T-N	Chovají totiž opovržení vůči sociálním médiím a internetu jako takovému.	27 (0,32)	80,02	0
N-R-A-N-V	Lásku k silným mašinám chovají oba slavní bráchové.	27 (0,32)	80,34	0

Tabulka 13: Přehled struktur APKS, resp. slovních druhů (POS) jednotlivých tokenů v anotovaných KWIC (v příkladech tučně); N – substantivum, A – adjektivum, P – pronomen, V – verbum, D – adverbium, R – prepozice, J – konjunkce, T – partikule; sloupec „APKS (%)“ označuje procentuální podíl ze všech analyzovaných APKS (Tabulka 9).

V analýze se nejdříve zaměřím na struktury tvořící výhradně APKS (s nulovou hodnotou ve sloupci „jiné“). Z výše uvedeného plyne, že zhruba ve 26 % jsou tato spojení v nepříznačném slovosledu složená pouze z kategoriálního slovesa a predikativního jména. Téměř identické zastoupení (24 %) má struktura s rozvitím predikativního jména adjektivem (V-A-N), na což ostatně poukazovaly mnohé přechodí studie (viz oddíl 2.3.3). Detailní třídění adjektiv jsem systematicky neprováděla, ale ze zběžného pohledu na data vyplývá, že se ve spojeních vyskytují adjektiva kvalifikační/jakostní (*něžné, silnější, hluboké, hlubší city, velké/malé/sebemenší ambice* atp.) i vztahová/relační (*stejně/vzájemně/mateřské/otcovské city, politické/tvůrčí/umělecké ambice* atp.). Nedomnívám se, že by se jednalo o adjektiva specifická pro predikativní jméno, stejná adjektiva se s daným substantivem pojí i mimo kontext APKS. Adjektiva přivlastňovací (mající na druhé pozici morfologického tagu hodnotu *U*) se v analyzovaných datech téměř nevyskytují. Ve zmíněné struktuře se nacházejí právě jednou (80), v ostatních anotovaných datech najdeme přivlastňovací adjektivum ještě šestnáctkrát, nikdy nicméně nerozvíjí predikativní jméno, ale je součástí jeho dalšího valenčního doplnění (81).

(80) Ta pozice v sobě **chová umělcovu touhu**.

(81) Krátká noticka u úvodníku informovala, že policie ani prokuratura **nechová vůči Ylviným příbuzným sebemenší podezření**.

U frekvence výskytu dalších struktur následuje poměrně rychlý propad. Přes 5 % z celkových výskytů má strukturu s opačným pořadím N-V (ale v pětina případů se nejedná o APKS, nýbrž o predikativní užití slovesa *chovat*, např. *chovat koně*). Následují spojení s predikativním jménem upřesněným záporným nebo neurčitým zájmenem (*žádné/nějaké ambice*), dále spojení predikativního jména (ne)rozvitého adjektivem s fakultativním dativním valenčním doplněním (*chovat k němu nepřátelství, chovat ke komunistům odpor*). Adjektivum může být dále zpřesněno příslovcem (*velmi něžné city, podezřele silné ambice*), příp. je kategoriální sloveso rozvíjeno několika predikativními jmény (*chovat respekt a úctu, chovat pochybnosti a nedůvěru*); vyskytují se i další variace.

Podívejme se ještě na struktury z výše uvedené Tabulky 13, které v některých případech klasifikovat jako APKS nelze (nenulová hodnota ve sloupci „jiné“). V pravidlech pro morfologické

značkování se s nimi vypořádám případ od případu: je možné je zcela pominout, ošetřit pravidla tak, aby konkrétní strukturu nezahrnovala, anebo na ně pravidlo přesto aplikovat, neboť je substantivum v akuzativu, přestože není součástí APKS.

První z této podmnožiny struktur je těsné spojení substantiva a slovesa (N-V; třetí řádek výše uvedené Tabulky 13). Bereme-li v potaz pouze KWIC, není žádný způsob, jak od sebe u této struktury odlišit APKS (82) a jiné struktury (83) a (84). Ve velké části jde ve skutečnosti o spojení se slovesem *chovat se* (83), přičemž *se* následuje až po KWIC. Ve zbylé části případů jde o spojení slovesa *chovat* s predikativním významem (84).

(82) Já osobně k němu moc velké **city nechovám**.

(83) Lidé pak necítí **potřebu chovat se** ekologicky

(84) Ten ve Svinově pro svou **potřebu choval** dvacet slepic.

Zatímco většina nalezené struktury s předložkami se řadí mezi APKS ve všech výskytech, u jiných tomu tak není (V-R-P-N, N-R-N-V). Na druhou skupinu jsem proto nahlédla ještě optikou použitých předložek, shrnutí viz Tabulka 14 níže.

předl.	APKS (%)	příklad	jiné	protipříklad
k/ke	954 (82,1)	Chovám k nim důvěru.	0	
v/ve	88 (7,57)	Chovám v sobě naději.	1	Žilková už v dětství měla potřebu v náruči chovat dítě. (N-R-N-V)
vůči	75 (6,45)	Nechováme vůči vám nenávisť.	0	
pro	10 (0,86)	Chováte pro něco vašeň?	5	Moje rodina žije na vesnici a pár slepiček jsme také chovali pro svou potřebu . (V-R-P-N)
o	10 (0,86)	Nechová o sobě pražádné iluze.	0	
na	7 (0,6)	Naději na zlepšení chová jen málokdo.	0	
od	5 (0,43)	Velmi jsem tuto stavbu obdivoval a chovám od té doby velký obdiv ke všemu.	1	Schopnost oddělovat osobní pocity od potřeby chovat se jako manažer profesionálně. (N-R-N-V)
proti	3 (0,26)	Nechovám proti nikomu nenávisť.	0	

za	3 (0,26)	K protivníkovi je třeba chovat za všech okolností úctu.	0	
z	2 (0,17)	Předstírat nadšení, když k nim chováte z duše odpor , není v pořádku.	0	
před	2 (0,17)	Klatovské mužstvo chová před Jihočechy respekt.	0	
podle	2 (0,17)	K doktorce chová podle svých slov stejnou důvěru a vděčnost jako k doktorce Kaňkové.	0	
díky	1 (0,09)	K Čechám chovám díky nedávným angažmá silný vztah.	0	

Tabulka 14: Předložky použité ve strukturách V-R-P-N, N-R-N-V se slovesem *chovat*; seřazeno podle frekvence; procenta odpovídají podílu jednotlivých předložek v obou strukturách.

Z tabulky plyne, že v 82 % případů obě zkoumané struktury obsahují předložku *k/ke*. Ta se vždy pojí s fakultativním doplněním predikativního jména v dativu (*chovat k někomu důvěru/cit...*), které navíc sloveso s predikativním významem *chovat* nemá. Z dalších dativních předložek se zde vyskytují předložky *vůči*, *proti* a *díky*. Ostatní předložky se vážou s jinými pády a bylo by vhodnější je analyzovat vždy na základě valenčního rámce predikativního jména, což je analýza výrazně přesahující rámec jednoduchých pravidel pro disambiguaci. Zároveň zde narážíme na problém, že analýza zahrnuje i substantiva, která APKS vytvářejí jen v zanedbatelném množství případů (jak bylo předvedeno v Tabulce 10); přičemž ve strukturách, které nejsou součástí APKS, se vyskytují právě ona.

Data jsem při zpracování obohatila o druhou pozici morfologického tagu (detailní určení slovního druhu). Údaj nakonec nepřinesl žádná neočekávaná zjištění, ale využila jsem ho alespoň pro zpětnou kontrolu. Všechny spojky například byly souřadící; pokud by se vyskytla spojka podřadící, jednalo by se o chybu anotace. Co se zájmen týče, v analyzovaných strukturách z Tabulky 13 nenajdeme žádný krátký tvar osobního zájmena (*mě*, *mi*, *ti*, *mu*) ani žádné vztažné nebo tázací zájmeno. Zastoupení jednotlivých tvarů zájmen viz níže v Tabulce 15.

detailní určení	příklad	APKS	%
zájmeno <i>on</i> , <i>oni</i> ve tvarech po předložce (tj. <i>n-: něj, něho, nich</i> ...)	A co víc, chovám k nim respekt. (VB-RR-P5-NN)	394	35,37

záporné zájmeno (<i>nic, nikdo, nijaký, žádný ...</i>)	Ale ona ke mně nechovala žádné city . (Vp-PW-NN)	228	20,47
ukazovací zájmeno (<i>ten, onen ...</i>)	Litujeme toho, protože chováme k tomuto lidu přátelský vztah . (VB-RR-PD-NN-AA-NN)	134	12,03
neurčité zájmeno (<i>nějaký, některý, číkoli, cosi ...</i>)	Věty sugerují představu, že chovám jakési sympatie vůči KSČM, a tím hraničí se lží. (VB-PZ-NN)	121	10,86
reflexivní zájmeno <i>se</i> v dlouhých tvarech (<i>sebe, sobě, sebou</i>)	Lidé se zdraví a chovají k sobě úctu . (VB-RR-P6-NN)	97	8,71
osobní zájmeno	Nechovám k tobě nepřátelství . (VB-RR-PP-NN)	70	6,28
přivlastňovací zájmeno <i>svůj</i>	Lidé v Tatobitech chovají ke své lípě velikou úctu . (VB-RV-P8-NN-AA-NN)	39	3,5
přivlastňovací zájmeno <i>můj, tvůj, jeho</i> (vč. plurálu)	Jestli můj synovec chová vaše názory v úctě, vezme je v potaz, i když se s vámi možná bude donekonečna hádat. (VB-PS-NN)	19	1,71
neurčité zájmeno <i>všechen, sám</i>	Můžete chovat k veškerému sportu vyslovený odpor , ale pro vašeho psa je každý pohyb odměnou. (Vf-RR-PL-NN-AA-NN)	10	0,9
reflexivní zájmeno <i>se, si</i> pouze v těchto tvarech, a dále <i>ses, sis</i>	Chovám si naději , že by byli také zvoleni. (VB-P7-NN)	2	0,18

Tabulka 15: Druhy zájmen u struktur, které tvoří 80 % všech APKS; zkratka v závorce reprezentuje první dvě pozice z tagu pro každé slovo v KWIC; řazeno podle frekvence, jednotlivé popisky viz *Příručka ČNK* (Cvrček a Richterová 2022b).

V neposlední řadě se zaměříme na takové struktury, které ve zkoumaných datech APKS nikdy nevytvářejí. Tyto struktury představují cca 7200 z celkových 8000 struktur (přes 9000 KWIC z 18 000), ale nevykazují žádné silné tendence ani pravidla. Nejfrekventovanější struktury jsou uvedené v tabulce níže (tabulka obsahuje pouze struktury, které se ve vzorku objevily minimálně 27krát; tentýž počet výskytů mají nejméně frekventované struktury, které APKS naopak tvoří, viz Tabulka 13). Plný výčet struktur je obsažen v souboru *prvni_studie_ostatni.csv*, který nabízí i možnost filtrování.

POS	příklad	jiné
N-Z-V	Mění názory, chová se jinak na veřejnosti a jinak doma.	122
N-Z-N-P-R-P-N-V-V	Absurdní příběh o tom, jak babička naplní svůj sen, matka se ke svým dcerám přestane chovat jako k dětem [...].	106
N-J-V	Dejte jí najevo svou lásku a chovejte se k ní něžně.	104

N-Z-J-P-V	Mám pocit, že se chová davově a stádně.	99
N-V	Ten ve Svinově pro svou potřebu choval dvacet slepic.	88
N-Z-J-P-N-V	Máte pocit, že se policie chová nemravně?	45
N-Z-J-P-R-P-N-V	Máme k němu jiný vztah, nebo se ke každému vozu chováte stejně?	39
V-D-J-V-N	Navíc se vláda chovala zmateně a ztratila respekt.	36
N-Z-D-P-V	S otcem jsem měla komplikované vztahy, často se choval agresivně.	34
V-D-Z-J-V-N	Lidé se mají chovat zásadově, ale nemít předsudky, říká se.	34
N-Z-J-N-P-V	Hráči mají podezření, že hra se chová jako živá.	30
N-Z-J-P-V-V	Máte pocit, že se umíme chovat k cizincům jako k sobě rovným?	27

Tabulka 16: Přehled struktur pro KWIC s jiným než predikativním významem

Poměrně kuriózní je struktura N-Z-N-P-R-P-N-V-V na druhém řádku, která mne zaujala svou délkou a komplexností (oproti N-Z-V a N-J-V) v kombinaci s vysokou frekvencí. Pohledem na data i dotazem do korpusu se ukázalo, že se jedná o jedinou větu, která je v korpusu¹²³ přítomná právě 106krát, a proto dokáže zkruslit data. Jedná se o krátký popis k českému filmu *Výlet*, který Mladá fronta DNES v totožném znění opakovaně otiskovala v jednotlivých vydáních.

5.3.1.3.4 Slovesný způsob

K analyzovaným datům jsem zpětně doplnila ještě hodnotu verbtagu pro kategoriální sloveso *chovat*. Na jeho druhé pozici se skrývá informace o slovesném způsobu, přičemž ve spojení je jasně preferován indikativ. Tabulka 17 srovnává rozložení slovesného způsobu u APKS s rozložením v korpusu SYN2005 (tj. v reprezentativním vzorku psané češtiny) zachyceným v publikaci *Statistiky češtiny* (Bartoň et al. 2009, s. 163).

slovesný způsob	SYN2005 (%)	APKS (%)
indikativ	80,1	92,5
kondicionál	4,7	1,8
imperativ	1,4	0,39
infinitiv	13,6	5,17
transgresiv	0,14	0,08
ostatní funkce trpného přičestí	-	0,06

Tabulka 17: Srovnání slovesného způsobu u klauzí obsahujících APKS se slovesem *chovat* s rozložením v SYN2005, jak ho uvádí publikace *Statistiky češtiny* (Bartoň et al. 2009)

¹²³ Konkordance dostupná z: <https://www.korpus.cz/kontext/view?q=~bcUuqWuw0Eg>.

V souvislosti s infinitivem zmiňuje Radimský (2010, s. 16) v návaznosti na lexikogramatiku slovesné *operátory*, které přidávají matematické funkce a měly by mít velmi širokou kolokabilitu (*Válka, kterou hodlá/mohla vést Francie s Anglií.*). S infinitivem *chovat* vyskytujícím se spolu s predikativním jménem registruji nejčastěji slovesa *moci, začít a přestat*.

5.3.1.3.5 Další pozorování

Zkoušela jsem v datech zkoumat i další jevy a ověřit pozorování předchozích badatelů, nicméně výsledky buď nebyly přesvědčivé, nebo pro takový účel data nejsou vhodná.

Úmyslně jsem se např. zaměřila na fenomén porušení koreference (viz 2.1.2.2). Nepodařilo se mi nicméně najít žádný přesvědčivý příklad tohoto jevu, přestože jsem pečlivě prošla všechna spojení s genitivem v pravém kontextu KWIC. Srov. příklad, který uvádí Radimský (85) a příklady (86) a (87), kdy není koreference porušena – v (86) je vypravěčka chůvou, chová tedy svoji lásku; v (87) jsou černoši imigranty, kteří sní svůj sen.

(85) Starosta **přednesl požadavek Technických služeb** o zvýšení cen za prodej hrobového místa. (JR, ČNK)

(86) Kráčela tou svou opatrnou chůzí s dvěma malými děvčátky, **chovajíce** k nim **lásku starostlivé chůvy**, tak jak to celé dětství u své matky vídávala.

(87) Avšak černoši, kteří přišli do Ameriky z Karibiku či z Afriky, ve svých srdcích **chovají sen imigranta** o lepším životě a jsou srozuměni s tím, že takový sen je vykoupěn tvrdou prací a usilovným studiem.

Co se týká čísla predikativního jména (zmiňovaného v kombinaci s videm slovesa v pododdílu 2.2.2.3), v datech se predikativní jméno objevuje třikrát častěji v singuláru než v plurálu. Výhradně v plurálu se vyskytuje predikativní jméno *cit*, dále je plurál typický pro substantiva *předsudek, sympatie, cit, iluze, pocit, antipatie, obava, ambice* (plurál ve více než 75 % případů). Naopak vždy v singuláru jsou predikativní jména *opovržení, vášeň, slabost, skepse, respekt, přízeň, přátelství, potřeba, nepřátelství, odpor, obdiv, nevraživost, nedůvěra, nechuť, důvěra, despekt, zášť*. Podrobné výsledky viz Příloha 5.

5.3.1.4 Stručné shrnutí

Pokud jde o spojení slovesa *chovat/chovávat* s predikativním jménem, ukázalo se, že výčet predikativních jmen ve VALLEXu ani v Radimském zdaleka není vyčerpávající (viz 5.3.1.1). Mnou upravený seznam byl na jednu stranu pojatý až příliš široce, a tudíž se do analýzy dostaly i struktury, které APKS vytvářejí jen výjimečně (*chovat názor, potřebu* atp.; viz Tabulka 10). Na druhou stranu jsem doplnila predikativní jména, která se zkoumaným slovesem vstupují do

APKS ve velkém množství případů a ve VALLEXu, resp. jeho pracovní verzi, zahrnuta nejsou (*antipatie, skepse, despekt, nechut'*; tamtéž). Krátká sonda do kolokací slovesa *chovat/chovávat* ukázala, že hledat predikativní jména na základě kolokačních měr nepřináší kýžené výsledky.

Při analýze materiálu z korpusu SYNv10 se ukázalo, že téměř polovina vyhledaných spojení odpovídá definici APKS. Více než 80 % těchto APKS má nepříznačkové pořadí členů V-N a většina tvoří velice těsné struktury, tj. takové, kdy se mezi kategoriálním slovesem a predikativním jménem nevyskytuje žádné další slovo, případně jich je pouze několik (průměr 1,61 slova za všechna spojení).

V další analýze jsem se zaměřila výhradně na APKS s oběma členy v rámci jedné klauze (tj. nikoliv na konstrukce rozdělené do dvou klauzí, jako v případě spojení *cit, který k ní chovala...*; tato podkategorie vykazuje poměrně velkou variabilitu, zároveň by bylo složité ji dále využít pro rozšíření pravidly řízené disambiguace). Nejčastější strukturou je dvojice kategoriálního slovesa a predikativního jména (v tomto pořadí) a dále spojení kategoriálního slovesa, adjektiva a predikativního jména. Tyto dvě struktury představují téměř 50 % APKS spojení a spolu s dalšími 24 konstrukcemi pokrývají 80 % všech analyzovaných APKS se slovesem *chovat/chovávat*. Z této charakteristiky je možné vyjít pro další zpracování těchto konstrukcí v rámci pravidlové disambiguace systému LanGr.

5.3.2 Druhá případová studie: ověření nalezených struktur na dalších APKS s bezpředložkovým akuzativem

Pilotní studie (oddíl 5.3.1) potvrdila předpoklad, že minimálně u APKS se slovesem *chovat/chovávat* lze znalost dvojic vstupujících do těchto spojení použít k řešení morfologické homonymie predikativního jména (viz podkapitola 5.1). V této kapitole se zaměřím na dva cíle: formulaci kritérií pro vyhledávání APKS v korpusu a jejich následné ověření na dalších spojení. Pokud se oba cíle splní, bude možné formulovat pravidla pro pravidlovou disambiguaci LanGr.

Pro formulaci kritérií využiji 25 struktur, které pokrývají téměř 80 % APKS s kategoriálním slovesem *chovat*, a vyřadím naopak nejednoznačnou strukturu N-V. Předpokládám, že po splnění následujících kritérií budou výsledky dotazu (KWIC) obsahovat pouze APKS:

- (a) hledané sloveso a substantivum v bezpředložkovém akuzativu spolu utvářejí APKS
- (b) hledaný výraz odpovídá jedné z těchto struktur: N-R-A-N-V, N-R-N-V, N-R-P-V, V-A-A-N, V-A-J-A-N, V-A-N, V-A-N-J-N, V-D-A-N, V-D-N, V-N, V-N-J-N, V-P-A-N, V-P-N, V-R-A-N-A-N, V-R-A-N-N, V-R-N-A-N, V-R-N-N, V-R-N-

N-N, V-R-N-P-N, V-R-P-A-N, V-R-P-N, V-R-P-N-A-N, V-R-P-N-N, V-R-P-P-N, V-T-A-N¹²⁴ (řazeno abecedně)

- (c) je-li součástí struktury spojka (J), musí to být spojka souřadící
- (d) je-li uvnitř struktury zájmeno (P), musí mít na druhé pozici tagu hodnotu 5, 6, 7, 8, D, L, P, S, W, nebo Z; tj. jedná se o zájmeno záporné, ukazovací, neurčité, reflexivní, osobní, přivlastňovací, nebo zájmeno *on, oni* ve tvarech po předložce
- (e) pro struktury V-R-P-N a N-R-N-V se omezím pouze na struktury s dativními předložkami *k/ke/vůči/proti/díky*

Ilustrujme si kritéria na následujícím příkladu. Vyhledáme lemma slovesa *chovat* a lemma substantiva *láska* v Asg nebo Apl, mezi nimiž je 0–4 tokenů (maximální délka struktury je 6 tokenů včetně substantiva i slovesa). Na základě kritérií stanovených výše předpokládáme, že nalezené výsledky bude vždy možné klasifikovat jako APKS, pokud bude KWIC odpovídat struktuře V-N (chovat *lásku*), V-A-N (chovat *mateřskou lásku*) nebo dalším sekvencím definovaným výše v bodě (b) a zároveň bude splňovat další kritéria v bodech (c)–(e). U jiných struktur se správnou délkou naopak buď víme, že APKS spolehlivě nevytvářejí (např. věta (88) – spojení s předložkou tam, kde očekáváme APKS s bezpředložkovým akuzativem), anebo k závěru nemůžeme dojít bez další analýzy, jako u struktury N-V v příkladech (89) a (90).

(88) Koně musíte **chovat pro lásku**, ne pro peníze.

(89) Matka k ní zjevně velkou **lásku nechovala**.

(90) Netuší, co v takové situaci dělat, jen u srdce si jako památku na ztracenou **lásku chová** myšlenku, že nakonec už na ničem nezáleží.

Pro praktický test využiji seznam APKS ze studie Kettnerové a Lopatkové (2017, s. 21–23),¹²⁵ konkrétně přílohu A. *Typ koreference ACT_{subst}-ACT_{verb}*, která obsahuje celkem 517 spojení (46 sloves a jejich možné kombinace s predikativním jménem). Jedná se o APKS, u kterých sloveso sdílí stejný valenční rámec s již analyzovaným slovesem *chovat/chovávat* a predikativní jméno je v bezpředložkovém akuzativu. Seznam je dostupný v Příloze 6 této práce, anotovaná data jsou k práci přiložena elektronicky v podobě souboru *druha_studie.csv*.

¹²⁴ Zkratky slovních druhů (N – substantivum, A – adjektivum, P – pronomen, V – verbum, D – adverbium, R – prepozice, J – konjunkce, T – partikule), viz *Příručka ČNK* (Cvrček a Richterová 2022b).

¹²⁵ Pro verifikaci formulovaných kritérií bude existující seznam plně dostačující, přestože by ho bylo možné rozšířit např. o aktuální pracovní verzi VALLEXu.

5.3.2.1 Získání dat a jejich anotace

Při zadávání korpusového dotazu formulovaného na základě kritérií výše jsem bohužel narazila na počítačové limity ČNK. Dotaz pokrývající zmíněné struktury je sice možné zkonstruovat, ale KonText u většího množství lemmat selže při jejich vyhledávání. Systém buď nevrátí žádné výsledky, nebo vyhodnotí jen část dotazu, přestože jednotlivé složky dotazu (kategoriální sloveso, predikativní jméno a případná slova mezi nimi) jsou správně zadané a v korpusu existují. Řešením se ukázalo nahradit přesné specifikace tagů, které se mohou vyskytovat mezi kategoriálním slovesem a predikativním jménem, obecnou strukturou [] značící libovolný token v dané pozici a pouze omezit počet tokenů, a to na 2–3 pro slovosled N-V a 0–4 pro slovosled V-N. Filtrování vybraných struktur následně probíhá při anotaci, a to v pythonovském skriptu, který k anotaci propustí pouze věty odpovídající uvedeným kritériím. Skript samotný je součástí elektronické přílohy práce (viz 5.4.2).

Pro rychlejší stahování dat jsem také upustila od zobrazení celé věty (vyžaduje znovunastavení v menu KonTextu při každém vyhledávání) a exportovala jednotlivé výsledky pouze s 20 tokeny před KWIC i po něm, což je pro anotaci dostatečný kontext.

Následně jsem u každého kategoriálního slovesa posuzovala vždy prvních 100 výsledků odpovídajících ověřovaným kritériím, s výjimkou sloves zvrtných.¹²⁶ Jednalo se tedy celkem o 4600 vět k anotaci. Při hodnocení jsem postupovala totožně jako v první studii: hodnotila jsem, zda jde o APKS, nebo o jiné spojení.

5.3.2.2 Analýza dat

Předpoklad, že analyzované věty budou obsahovat APKS, se naplnil – téměř všechna spojení odpovídají APKS nebo se jedná o spojení s polysémními dějovými jmény (např. *uzavřít vyšetřování* – přestat vyšetřovat vs. *uzavřít konkrétní policejní případ*), která v tomto případě počítám k APKS, přestože můžeme vidět rozdíly mezi predikativností spojení u jednotlivých případů jako v příkladech (91) a (92).

(91) **Uzavřeli vyšetřování** té nehody.

(92) Kriminalisté stále **neuzavřeli vyšetřování** loňské vraždy faráře Ladislava Kubíčka (78) z Třebenic na Litoměřicku.

Pouze v malém množství analyzovaných KWIC se o APKS nejednalo (zejména u spojení se strukturou V-N). Šlo o případy, kdy (a) můžeme spojení označit jako jazykově neobratné na

¹²⁶ Vynechány jsou APKS *činit si nárok, dělat si iluze/legrace/nárok/srandu* a *utvořit/utvářet si názor/představu/úsudek/vazbu/vztah*.

místě, kde by mohlo být použito vhodnější synonymum (srov. příklad (93) s variantou *najít východ z bludiště*, případně variantou *najít východisko* bez určení *z bludiště*), (b) se jedná o čisté predikativní užití daného slovesa místo APKS (viz příklad (94)).

(93) Nechce Tomášovi nic oplácet. Chce **najít východisko** z bludiště. Ví, že se stala jeho přítěží: bere věci příliš vážně, ze všeho dělá tragédii [...].

(94) Zřejmě **měl smysly** otupělé a zakrnělé penězi a jejich vyděláváním.

Pokud bychom při morfologické disambiguaci chtěli vycházet výhradně z APKS, tyto relativně řídké případy nesprávně určených spojení by bylo nutné vyloučit při ruční kontrole, případně na základě pozorování zpřesnit korpusové výsledky použitím negativního filtru. Cílem této disertační práce je nicméně vytvořit pravidla pro morfologickou disambiguaci, v níž jde především o určení správného pádu. I v těchto „sporných“ případech se přitom jedná o akuzativ, tudíž je možné je s klidným svědomím ponechat na místě bez dalších úprav.

Distribuci struktur shrnuje Tabulka 18 níže. V tabulce jasně vidíme převahu struktur V-N a V-A-N ve srovnání s ostatními sekvencemi, která je ještě výraznější než u struktur se slovesem *chovat* (v nejčastějších 26 strukturách se slovesem *chovat* se struktura V-N vyskytuje v 32 % a V-A-N v 30 %). To může vyplývat ze skutečnosti, že seznam predikativních jmen ke kategoriálnímu slovesu *chovat* byl co nejširší, zatímco ve druhé studii vycházím ze seznamu mnohem užšího.

POS	příklad	počet APKS (podíl ze všech APKS v %)	kumula- tivní %
V-N	Stejně ale chovám naději , že skutečnost je jiná.	2430 (52,94)	52,94
V-A-N	Přesto musím přebírat plnou odpovědnost za všechny své činy [...].	1061 (23,12)	76,06
V-P-N	Ještě žes našel tu odvahu a řekl mi pravdu.	414 (9,02)	85,08
V-P-A-N	Již příliš dlouho uplatňují své přemrštěné nároky na ovlivňování našich životů.	157 (3,42)	88,5
V-D-A-N	Na levém kraji obrany podal velmi dobrý výkon .	109 (2,37)	90,87
V-A-A-N	[V]dova pojala šílený vdovský úmysl učinit pánem nad rozháraným koncertním jednatelstvím mě.	86 (1,87)	92,74
V-D-N	ČNB ovšem následně vydala opět rozhodnutí o pokutě 1,7 milionu korun [...].	59 (1,29)	94,03

V-R-N-N	Dnes by měl soud vynést nad Tichým rozsudek.	48 (1,05)	95,08
V-R-N-A-N	Má na Broadwayi domovské právo.	29 (0,63)	95,71
V-R-P-A-N	Od začátku ji ochraňoval, jevil o ni upřímný zájem a od okamžiku, kdy se na něho poprvé okouzleně podívala, stal se pro ni hrdinou [...].	25 (0,54)	96,25
N-R-P-V	Když Karin opět promluví, slyším, že už boj proti tomu vzdala.	22 (0,48)	96,73
N-R-A-N-V	V úterý totiž pokus o stejný vrchol vzdal jiný vyhlášený osmitisícovkář, Španěl Carlos Pauner.	21 (0,46)	97,19
V-A-J-A-N	Podat poctivý a kvalitní výkon podpořený dobrou koncovkou.	16 (0,35)	97,54
V-R-A-N-N	Doufám, že si z toho hráči vezmou do dalších utkání ponaučení.	14 (0,31)	97,85
V-N-J-N	Jen našli sílu a odvahu se ke svému problému přiznat.	14 (0,31)	98,16
V-R-A-N-A-N	Získal na hlavní pole minutový náskok, ale to bylo všechno.	13 (0,28)	98,44
V-T-A-N	Padang o svou novou partnerku projevuje opravdu velký zájem.	12 (0,26)	98,7
V-R-P-N-N	[P]ůvodní nájemce má na jeho vystěhování čas do konce března.	10 (0,22)	98,92
V-R-N-N-N	Vznáším proti generálu Stoneovi stížnost.	10 (0,22)	99,14
V-R-P-P-N	Když si vytíral chlebovou kůrkou omáčku, vydával ze sebe nějaké zvuky naznačující spokojenost.	8 (0,17)	99,31
V-A-N-J-N	Někdy je nutné přijmout tvrdá rozhodnutí a opatření, která nemusí každý vítat s radostí.	8 (0,17)	99,48
V-R-P-N	Riskoval tím však, že se ho později bude Junk vyptávat na jeho telepatické kouzlo a pojme vůči němu podezření.	7 (0,15)	99,63
N-R-N-V	Ostrava čeká, jaký postoj k věci zaujme vedení klubu z Uherského Hradiště.	6 (0,13)	99,76
V-R-P-N-A-N	Oni uplatňovali podle svého chápání přirozené právo majorit vůči minoritám.	6 (0,13)	99,89
V-R-N-P-N	Od začátku nám kladl na srdce, že nemáme vyvíjet na Blunta žádný nátlak, aby neemigroval.	5 (0,11)	100,0

Tabulka 18: Distribuce zkoumaných struktur v anotovaných datech

Vzorek bohužel není vhodný pro analýzu vidu kategoriálního slovesa ani čísla predikativního jména. Kvůli širší záběru se každé zde analyzované kategoriální sloveso pojí s několika predikativními jmény, z nichž každé se v analyzovaném vzorku dat (100 vět na jedno sloveso) vyskytuje pouze několikrát, vyvozené závěry by proto nebyly relevantní.

5.3.3 Tvorba pravidel pro morfologickou disambiguaci LanGr

Jak jsem nastínila v podkapitole o motivaci tohoto výzkumu (viz 5.1), jeho cílem je rozšířit pravidla pro automatickou morfologickou disambiguaci systému LanGr a díky ustálenosti APKS řešit morfologickou homonymii predikativních jmen. Jelikož víme, která predikativní jména a kategoriální slovesa do APKS společně vstupují (seznamy můžeme čerpat např. ze zdrojů uvedených podkapitole 2.6), a zároveň známe pád daného predikativního jména v APKS, můžeme vytvořit pravidlo, které provede pádovou disambiguaci.

Pravidla se uplatňují v přesně specifikovaném kontextu, kterým budou struktury nalezené v první případové studii (5.3.1) a ověřené na větším seznamu sloves se stejnou valencí ve studii druhé (5.3.2). V tomto oddíle se nejprve ještě krátce vrátím k širší morfologické homonymie a poté se zaměřím na samotné disambiguační pravidlo pro systém LanGr.

5.3.3.1 K míře morfologické homonymie predikativních jmen

Na první pohled se sice může zdát, že se morfologická homonymie týká jen zlomku predikativních jmen tvořících APKS, avšak tento problém zdaleka není marginální.

Demonstrujme to na seznamu z Přílohy 6 pocházejícím ze studie Kettnerové a Lopatkové (2017, s. 21–23), který jsem použila v druhé případové studii (5.3.2). Z něj nejprve extrahujeme všechna predikativní jména a odstraníme duplicity (vzniklé skutečností, že se totožné predikativní jméno může pojit s více slovesy), čímž nám vznikne soupis 339 substantiv. Dále vyjdeme ze zjednodušeného scénáře, kdy se každé substantivum, které je součástí APKS, může uplatnit jak v jednotném, tak v množném čísle akuzativu.¹²⁷ Vezmeme v potaz i skutečnost, že substantivum *hrátky* nemá singulár a mnohá jiná substantiva naopak v akuzativu mají více tvarů plurálu (např. *chut'* – *chutě*, *chuti*). Celý tento proces proběhl následovně:

Vytvořila jsem skript, který v rozhraní *Internetové jazykové příručky*¹²⁸ vyhledal jednotlivá substantiva v Nsg a doplnil k nim tvary Asg a Apl. Akuzativní tvary následně v příručce vyhledal znovu. První dotaz sloužil k vyhledání homonymie v rámci paradigmatu jednoho

¹²⁷ Uvědomuji si, že v úzu může docházet k situacím, kdy se u některých APKS predikativní jméno vyskytuje výhradně v singuláru, nebo výhradně v plurálu, jak ukázala analýza predikativních jmen u APKS s kategoriálním slovesem *chovat*, viz 5.3.1.3.5.

¹²⁸ Dostupné z <https://prirucka.ujc.cas.cz/>.

substantiva (tedy homonymie vlastní), druhý dotaz doplnil případnou homonymií nevlastní. Nakonec skript přidal informaci uvedenou v oddíle „Zadaný řetězec (bez ohledu na diakritiku) byl dále nalezen v následujících heslech:“. Zde byla nakonec nutná pečlivá ruční kontrola, tento totiž oddíl nezohledňuje diakritiku a např. k A_{sg} *úctu* od substantiva *úcta* nabízí i tvar „učíst 1. osoba, jednotné číslo: učtu“. Použité skripty jsou součástí elektronické přílohy práce (viz 5.4.3); vyčištěná data jsou ve zjednodušené formě dostupná v Příloze 7, v plném znění jsou k nalezení v elektronické příloze práce jako soubor *priloha_7.csv*.

K 600 z 681 tvarů existuje alespoň jeden homonymní tvar, ať už jde o homonymií vlastní, nebo nevlastní. Nevlastní homonymní tvary se v mnoha případech vyskytují u sloves (ty pravidlo pro LanGr pokrývat nebude), najdeme je ale i u podstatných jmen – jmenujme např. tvar *politiku* pro A_{sg} substantiva *politika*, homonymní se substantivy *politik* (V_{sg}) a *politikum* (D_{sg}, L_{sg}). Vezmeme-li v potaz pouze vlastní morfologickou homonymii (tj. morfologii v rámci paradigmatu jednoho substantiva) s akuzativem v singuláru nebo plurálu, získáme přehled shrnutý v Tabulce 19:

počet homonymních pádů	pády homonymní s A _{sg} /A _{pl}	celkem tvarů	příklad v N _{sg}
2	N _{sg} , A _{sg}	167	boj, chuť, cit, cvik
3	D _{sg} , A _{sg} , L _{sg}	28	akce, demise, dokumentace
	N _{sg} , A _{sg} , V _{sg}	7	gesto, jméno, právo, riziko
	N _{pl} , A _{pl} , V _{pl}	1	hrátky
4	G _{sg} , N _{pl} , A _{pl} , V _{pl}	103	důvěra, fotka, idea, iniciativa
	N _{pl} , A _{pl} , V _{pl} , I _{pl}	125	klid, kompromis, koncert, let
6	N _{sg} , G _{sg} , V _{sg} , N _{pl} , A _{pl} , V _{pl}	28	energie, invaze, kopie, praxe
7	G _{sg} , D _{sg} , V _{sg} , L _{sg} , N _{pl} , A _{pl} , V _{pl}	34	lítost, možnost, nelibost, nenávisť, samostatnost
10	N _{sg} , G _{sg} , D _{sg} , A _{sg} , V _{sg} , L _{sg} , N _{pl} , G _{pl} , A _{pl} , V _{pl}	102	opatření, označení, oznámení, podezření, pochopení

Tabulka 19: Přehled pádů s morfologickou homonymií s A_{sg} nebo A_{pl}

Část homonymních tvarů je již samozřejmě pokrytá stávajícími pravidly, nicméně nemalá část vyřešená není, pravidlová disambiguace konkrétního homonymního substantiva neproběhne a rozhodnutí zůstává až na statistickém taggeru. Přidáme-li pravidlo založené na APKS, které homonymii úspěšně vyřeší, zpřesní se znalost o dané větě a před předáním rozpracované, tj. částečně disambigované věty statistickému taggeru se mohou uplatnit další pravidla.

5.3.3.2 Zápisy pravidel

Aby bylo možné kritéria z druhé případové studie (5.3.2) transformovat na pravidla pro morfologickou disambiguaci, je nutné provést dva kroky – sestavit seznam APKS, který budou pravidla využívat, a pravidla samotná.

Začnu sestavením seznamu (tabulky) bigramů, přičemž bigram zde označuje dvojici lemmat kategoriálního slovesa a predikativního jména oddělených čárkou (řádky 10–12). Ukázkou zápisu zachycuje Zdrojový kód 3 níže.

```
1. charset "unix";
2.
3. /* Seznam bigramů lemmat kategoriální sloveso - predikativní jméno v akuzativu
   */
4.
5. parse PositionType;
6.
7. shared StringTupleList VerbNounLemmaList;
8.
9. VerbNounLemmaList = [
10. {"chovat", "zášť"},
11. {"provést", "výpočet"},
12. {"činit", "opatření"}
13. ];
```

Zdrojový kód 3: Ukáзка zápisu seznamu (tabulky) APKS v programovacím jazyce LanGr

Pokud jde o samotná pravidla, jejich implementaci představím na zjednodušeném příkladu pokrývajícím nejprostší APKS struktury, u nichž mezi kategoriální sloveso a predikativní jméno může být vloženo ještě syntaktické adjektivum (tj. adjektivum, zájmeno nebo číslovka se syntaktickou funkcí), adverbium, příp. částice:

(95) Soudce **vynesl rozsudek**. (vI)

(96) Jana **nesla velkou odpovědnost** za správný chod podniku. (VP)

(97) Nechtěli **brát vůbec ohledy** na staré lidi. (VP)

Mějme tedy v klauzi možný slovesný tvar (který není pasivním přičestím), po kterém může (ale nemusí) následovat posloupnost syntaktických adjektiv, adverbíí nebo částic, a strukturu uzavírá možné substantivum v akuzativu.

Pokud je tato konfigurace splněna, můžeme přikročit k samotnému vykonání disambiguační akce. Najde-li se (*HOLDS ON*) v seznamu bigramů dvojice, jejíž první člen je lemmatem slovesného tvaru nalezeného v textu a druhý člen je lemmatem substantivního tvaru nalezeného v textu v akuzativu, pak se první možný (*Possible*) slovesný tvar určí jako jistý (*IsSafe*)

slovesný tvar a zároveň se druhý možný (*Possible*) substantivní tvar v akuzativu určí jako jistý (*IsSafe*) substantivní tvar v akuzativu.

Toto pravidlo zachycuje Zdrojový kód 4 níže:

```
14. Rule VerbNom {
15.
16. Verb1 = ITEM (Possible Verb) and (MustNotBe PassiveParticiple);
17. SEQUENCE OF IsSafe AdjectiveSynt or AdverbSynt or Particle;
18. Noun2 = ITEM Possible Noun and Accusative;
19.
20. IF (((Possible (lemma == VN.string1)) HOLDS ON Verb1) and
21. ((Possible (lemma == VN.string2)) HOLDS ON Noun2))
22. WHERE VN IN VerbNounLemmaList THEN {
23. LEAVE ONLY Verb IN Verb1;
24. LEAVE ONLY Noun and Accusative IN Noun2;
25. };
26.
27. };
```

Zdrojový kód 4: Zápis pravidla pro základní strukturu APKS v programovacím jazyce LanGr

Výše popsané pravidlo pokrývá prominentní část struktur APKS s bezpředložkovým akuzativem, ale bohužel systematicky neošetřuje hraniční případy jako (98), (99) a (100) níže. Ty charakterizuje omezený kontext, ve kterém mohou APKS s daným slovesem utvářet obě přítomná substantiva.

(98) **Soud** (Nsg|Asg) vynesl **rozsudek** (Nsg|Asg). (vl)

(99) **Soud** (Nsg|Asg) vynesl **verdikt** (Nsg|Asg). (vl)

(100) **Dozor** (Nsg|Asg) vykonal **úkon** (Nsg|Asg). (vl)

V aktuální implementaci pravidla záleží na pořadí, v jakém jsou dvojice slovesa a substantiva utvářející APKS uvedeny v seznamu. Akuzativ se určí na základě prvního spojení v pořadí a ve čtení seznamu se nepokračuje. Máme-li dvě konkurenční spojení se stejným kategoriálním slovesem (*vynést rozsudek* a *vynést soud*), musíme je seřadit tak, aby pravidlo jako první našlo spojení *vynést rozsudek*, který je správnou variantou pro disambiguaci věty (98).

Takové řešení nicméně bude fungovat jen u vět, v nichž predikativní jméno v akuzativu následuje po kategoriálním slovesu; pravidlo se totiž aplikuje výhradně na sekvence začínající slovesem, zatímco sekvence opačné neošetřuje. Větu z příkladu (101) by tedy disambigovalo nesprávně a akuzativ by přiřadilo substantivu *soud*.

(101) **Rozsudek** (Nsg|Asg) vynesl Ústavní **soud** (Nsg|Asg). (vl)

Do budoucna se nabízí několik řešení: na hraniční případy pravidla s bigramy neaplikovat, nebo naopak zpracovat trigramy hraničních případů do pravidel, která se bigramům předradí.

Bohužel nelze přesně změřit účinnost jednotlivých pravidel vycházejících ze seznamů APKS, jelikož je nelze otestovat izolovaně, nicméně pravidla jsou funkční a víceznačnost úspěšně řeší. Používaný seznam bigramů bude postupně rozšířen o APKS ze seznamů pořizovaných při zpracování teoretických poznatků o APKS. Po APKS s bezpředložkovým akuzativem a se zvratným slovesem se pozornost přesune na další, násobně méně frekventované bezpředložkové pády, vhodná se zdají být spojení s genitivem (*dojít naplnění, nechat práce*) a dativem (*podlehnout pocitu, propadnout iluzím, uniknout pozornosti*).

5.4 Používané skripty

Velkou část úsilí věnovaného této disertační práci je možné vytušit jen mezi řádky, z toho důvodu se ještě krátce zaměřím na používané skripty. Ty jsou součástí elektronické přílohy této práce jako *skripty.pdf*, zároveň jsem je publikovala ve veřejném repozitáři GitHub,¹²⁹ v němž bude případně k dispozici i jejich aktualizovaná verze. U veřejně dostupných skriptů existuje větší šance, že nezůstanou jen artefaktem vykonané práce, ale pomohou i dalším badatelům a budou se dále vyvíjet. Ráda bych jejich publikací také podpořila otevřenost lingvistického výzkumu.¹³⁰

Skripty využívané pro analýzu a vyhodnocování jazykových dat jsou napsané v programovacím jazyce Python, ve kterém ale běžně neprogramuji. Přestože výborně posloužily mému záměru a jsou plně funkční, je možné, že jejich struktura a komentáře neodpovídají zažitým zvyklostem. Je tedy záhodno, aby s touto skutečností případný uživatel počítal. K práci přikládám pouze skripty komplexnější, případně skripty, které nejsou zcela jednoúčelové (tj. neřeší čištění dat, třídění atp.).

Pokud ve skriptech pracuji s konkordancí staženou z ČNK, je třeba, aby odpovídala těmto formálním požadavkům:

- exportované CSV soubory obsahují následující sloupce: *left* (levý kontext), *kwic* (KWIC), *right* (pravý kontext); další sloupce jsou volitelné
- KWIC obsahuje lemma a morfologický tag, volitelně i verbtage

¹²⁹ Zdrojové soubory jsou dostupné na adrese <https://github.com/vyskoczilova/phd-public>.

¹³⁰ Otevřenosti lingvistického výzkumu se např. věnovalo celé číslo 104 časopisu *Naše řeč*, kde Chromý a Cvrček (2021) zdůrazňují význam tohoto tématu.

- levý a pravý kontext může obsahovat stejné údaje jako KWIC výše
- pracuji s původním pojmenováním CSV souboru (*kontext-conc_download* + datum stažení), před který jako předponu ukládám typ dat, např: *N-V_kontext-conc_download-2022-04-11.csv*

Skripty navzájem sdílejí některé funkce a datové struktury, které jsou součástí složky *scripts*; není potřeba je měnit, ale jsou nutné pro běh jednotlivých skriptů.

5.4.1 Skripty použité v první studii

Skripty pracují s následující strukturou dat: zpracovávaná data jsou uložena ve složce pojmenované podle kategoriálního slovesa, v ní jsou dále roztržena do složek pojmenovaných podle jednotlivých predikativních jmen.

prvni_studie_annotace_predikativnich_jmen.py – Skript pracuje se staženou konkordancí a frekvenčním seznamem z ní vytvořeným. Umožňuje procházet jednotlivá substantiva z frekvenčního seznamu a zobrazuje k nim příklady z konkordance. Pomocí dialogu umožňuje přiřadit danému predikativnímu jménu anotaci a připsat vzorovou větu. Při spuštění skriptu je třeba zadat do parametru adresář s daty. Přímou ve zdrojovém kódu je lze upravit další parametry – seznam „předschválených“ predikativních jmen a počet zobrazených příkladů.

prvni_studie_annotace_dat.py – Při spuštění skript očekává parametr s názvem složky kategoriálního slovesa, názvem složky predikativního jména, a případně nepovinný parametr *--type* s prefixem aktuálně anotované konkordance (např. „N-V“). Při procházení jednotlivých vět se zobrazuje aktuálně anotovaný řádek se zvýrazněným KWIC a odstraněnými dodatečnými informacemi (jako je tag, lemma apod). Skript očekává zadání vstupu s výslednou anotací do příkazového řádku, viz Obrázek 6 níže:

[1/2]

Vrhl se jako smyslů zbavený s napřaženým mečem proti muži, který právě zranil ženu, kterou ... ke které **choval jisté city** .

Je to příklad APKS?

0 = ne, 1 = ano, 2 = možná, 3 = vztažné, x = skončit: █

Obrázek 6: Ukázka anotace jazykových dat

prvni_studie_data.py – Skript sloučí všechna anotovaná data do jednoho XLSX souboru pro jednodušší zpracování. Jeden list obsahuje APKS a druhý ostatní spojení; oba typy spojení jsou

doplněné údajem o struktuře (slovní druh a druhá pozice) a délce KWIC; v elektronické příloze tomuto výstupu odpovídají soubory *prvni_studie_apks_vztazne.csv* a *prvni_studie_ostatni.csv*.

prvni_studie_statistika.py – Skript připraví základní statistické přehledy pro další analýzu (konkrétně Grafu 1, Tabulky 8, Tabulky 10, Přílohy 2, Přílohy 3, Přílohy 4). Jako parametr očekává název složky s daty.

prvni_studie_statistika_pos.py – Skript z dat vytvoří několik pomocných přehledů a zejména citované tabulky (konkrétně Tabulky 13, Tabulky 14, Tabulky 1229, Tabulky 16) s informacemi z morfologického tagu (slovní druh, druhá pozice tagu) a lemmatu. Při spuštění skriptu je jako parametr potřeba zadat název složky se všemi anotovanými soubory.

prvni_studie_pos_ke_kolokacim.py – Ke staženému souboru s kolokacemi skript doplní slovní druhy (na základě konkordancí s anotovanými daty používanými v předchozích skriptech), případně u substantiv ze seznamu existujících APKS doplní i zdroj, který ho jako predikativní jméno označil.

5.4.2 Skripty použité v druhé studii

Skripty pracují s následující strukturou: CSV konkordance k analýze jsou všechny uloženy v jedné společné složce, výsledky analýzy se ukládají do podsložky s názvem *_data*.

druha_studie_anotace_dat.py – Skript slouží k anotaci jednotlivých konkordancí stažených z korpusu a analyzovaných v oddílu 5.3.2. Projde jednotlivé řádky zdrojového souboru a propustí k anotaci jen KWIC odpovídající hledaným strukturám. Anotace probíhá přes příkazový řádek, stejně jako u skriptu používaného v první případové studii. Při spuštění skriptu je třeba zadat dva povinné parametry: název složky, ve které jsou všechna data k anotaci uložená, a název předpony souboru (v mém případě sloveso). Nepovinně je možné přidat i parametr *--output_column*, který umožní ukládat anotaci do jiného sloupce (např. pro druhou kontrolní anotaci).

druha_studie_statistika.py – Navazuje na předchozí skript a z anotovaných souborů zpracovává XLSX soubor s výslednými daty (v elektronické příloze jako *druha_pripadova_studie.csv*). Sestaví základní statistiku, příklady (které je třeba ručně ve zdrojovém kódu aktualizovat) a procentuální zastoupení jednotlivých struktur – v práci jako Tabulka 18. Při spuštění skriptu je jako parametr potřeba zadat název složky obsahující anotované soubory.

5.4.3 Skripty pro analýzu morfologické homonymie

Malý set skriptů pro zpracování morfologické homonymie, parametry jsou nastavovány přímo ve zdrojovém kódu.

homonyma_v_priruce.py – Skript, který získává informace o vlastní morfologické homonymii z *Internetové jazykové příručky* a experimentálně doplňuje informaci o dalších tvarech, které příručka uvádí.¹³¹ V kódu je třeba upravit seznam hledaných substantiv v nominativu a pád, pro který mají být hledány další tvary. Skript si otevře okno prohlížeče a bude vyhledávat jednotlivá slova stejně, jako kdybychom je vyhledávali v příručce ručně, ovšem násobně rychleji. Zároveň zjistí, ve kterých pádech k homonymii tvarů dochází. Výstup se zobrazí v příkazovém řádku. Z výsledků skriptu vychází Příloha 7.

homonyma_v_priruce_statistika.py – Data ze skriptu výše jsou statisticky zpracována a výsledkem je přehled různých typů morfologické homonymie a jejich počtu v analyzovaných datech, viz Tabulka 19. Uvnitř skriptu je třeba upravit název zdrojového souboru a případně oddělovač jednotlivých sloupců v CSV. Výstup se zobrazí v příkazovém řádku.

5.4.4 Skript pro konverzi dat z Excelu do CSV

Při analýze dat jsem pracovala s tabulkami v editoru Excel (formát XLSX), tato data jsem kvůli povinnému formátu elektronické přílohy musela převést do CSV.

xlsx_do_csv.py – Jednoduchý skript, který při spuštění očekává parametr s názvem souboru ke konverzi a příp. volitelný parametr *--tabs*, ve kterém lze specifikovat záložky, které se mají vyexportovat. Nejsou-li tímto parametrem záložky vybrány, vyexportují se všechny záložky do jednotlivých CSV souborů.

¹³¹ Bohužel ale příručka řeší shodu s hledaným tvarem bez rozlišování diakritiky, takže je nutné tyto informace ručně překontrolovat.

6 Závěr

Předkládaná disertační práce stojí na pomezí teoretické a aplikované lingvistiky. Ilustruje, jak v praxi zúročit teoretické poznání analytických predikátů s kategoriálním slovesem (APKS), vytěžit ho při morfologické disambiguaci, a tím potenciálně ovlivnit i syntaktickou analýzu.

Práce si kladla dva hlavní cíle. Teoretickým cílem bylo APKS charakterizovat, popsat současný stav jejich poznání a usouvztažnit příspěvky k tématu, které vyšly v průběhu posledních patnácti let. Praktickým cílem bylo zjistit, zda jsou daná spojení v jazykových datech přítomná ve formě charakteristických jednotek, které by bylo možné využít k formulování pravidel pro morfologickou disambiguaci LanGr, případně taková pravidla zkonstruovat a zlepšit tak morfologickou anotaci korpusů ČNK. Oba záměry se podařilo naplnit.

Teoretická část práce se věnovala třem základním okruhům vytvářejícím východiska pro část empirickou: APKS samotným, značkování korpusů a podrobné charakteristice systému LanGr.

Kapitola 2 věnovaná analytickému predikátu s kategoriálním slovesem shrnula předchozí poznatky uceleně po jednotlivých oblastech výzkumu, nikoliv po jednotlivých autorech, jak bývá často zvykem. Nejprve sumarizuje kritéria a testy pro určování APKS, a to jak na povrchové a sémantické rovině, tak na rovině hloubkové, kde představuje i relativně nové pojetí testu koreference. K APKS přistupuje z pohledu lingvistických disciplín a zaměřuje se zejména na výzkum jevu v posledních patnácti letech, kdy po vydání Radimského (2010) monografie k problematice vyšlo poměrně velké množství článků věnujících se APKS z pohledu PDT a jejich implementaci do slovníku VALLEX. Podstatnou část kapitoly tvoří charakteristika APKS jako analytického predikátu v oblastech spojených s morfologií a syntaxí: valence jednotlivých členů i celého spojení, typologie kategoriálních sloves, problematika desémantizace kategoriálního slovesa a další. V neposlední řadě kapitola odkazuje na existující seznamy českých APKS, které lze dále využít při morfologické disambiguaci.

Značkování korpusů (zejména korpusů ČNK) se věnovala kapitola 3, která čtenáře provedla celým procesem – od tokenizace, větné segmentace a morfologické analýzy (včetně přidání verbtagů) přes proces morfologické disambiguace až po finalizaci a dodatečné úpravy. Nejpodrobněji kapitola popisuje morfologickou disambiguaci (výběr správné značky a lemmatu z celé množiny potenciálně možných údajů) a jednotlivé metody, kterými lze disambiguaci provádět, tj. disambiguaci statistickou, disambiguaci na základě pravidel a tzv. hybridní metodu, která obě zmíněné spojuje. Kapitola zachycuje i poslední změny v oblasti značkování

korpusů, jako je využití neuronových sítí při disambiguaci a zavedení tzv. verbtagů při značkování sloves.

Kapitola 4 zevrubně popisuje systém pravidly řízené redukční automatické morfologické disambiguace LanGr, který je vyvíjen na ÚTKL a využíván při značkování korpusů ČNK. Představuje jeho fungování na teoretické úrovni a ilustruje lingvistické uvažování při tvorbě jednotlivých pravidel. Zaměřuje se také na techničtější stránku tvorby pravidel, konkrétně na jejich zápis v programovacím jazyce LanGr, který demonstruje na příkladu dvou pravidel. Během zpracovávání kapitoly vznikl též soupis příkazů a operátorů, který tvoří Přílohu 1.

Praktická část (kapitola 5) disertační práce propojuje jednotlivé oblasti představené v části teoretické. Úspěšnosti morfologické disambiguace stojí v cestě rozsáhlá morfologická homonymie substantiv, což má dopad i na případnou syntaktickou analýzu jednotlivých vět. Nejpalčivějším problémem je v tomto ohledu homonymie nominativu a akuzativu, kde nesprávné rozlišení pádu ústí v potenciálně špatné určení podmětu a předmětu věty. K řešení tohoto problému práce přispívá analýzou a následným popisem pravidelností struktur s APKS ve dvou případových studiích, jelikož APKS nejsou stávajícími pravidly dostatečně pokryty.

První případová studie na rozsáhlém jazykovém materiálu (přes 18 000 vět) zkoumá APKS se slovesem *chovat/chovávat* ve spojení s maximálně rozšířeným seznamem predikativních jmen. Analýza těchto spojení nalezených v korpusu poukázala na existenci predikativních jmen, která tvoří APKS téměř ve všech souvýskytech se slovesem *chovat/chovávat* (např. *zášť, sympatie, nedůvěra, obdiv* aj.). Hlavní pozornost jsem věnovala složení výsledného KWIC odpovídajícího hledanému APKS, kde data potvrdila, že sekvence slovních druhů tvořících KWIC (APKS) vykazují pravidelnost, kterou jiná spojení těchto substantiv a daného slovesa neoplývají. Pro APKS nalezené v korpusu je nejtypičtější sekvence slovesa a substantiva (*chovat úctu*) společně se sekvencí slovesa, adjektiva a substantiva (*chovat mateřský cit*). Kromě nich bylo v analyzovaném materiálu identifikováno ještě dalších 24 sekvencí, které jsou pro APKS charakteristické.

Na základě uvedených pozorování jsou sestavena kritéria pro vyhledávání, po jejichž zadání by měla výsledná data odpovídat APKS. Tento předpoklad ověřuje druhá případová studie na APKS se stejným valenčním rámcem, kde analýza vzorku 100 vět ke každému z 46 hledaných kategoriálních sloves a jeho predikativních jmen validitu kritérií potvrzuje.

Kritéria byla následně přetavena v pravidla pro morfologickou disambiguaci LanGr pokrývající APKS v bezpředložkovém akuzativu. Existující seznamy APKS (resp. jejich část) byly nejprve převedeny na tabulku dvojic lemmatu predikativního jména a kategoriálního

slovesa. Ta posloužila jako podklad pro základní pravidlo ošetřující nejčastější struktury zahrnující sloveso, sekvenci syntaktických adjektiv, adverbii a příp. částic zakončené substantivem v akuzativu. Pravidlo pokrývá prominentní část struktur APKS, bude možné ho dále rozšířit o další vytipované unikátní struktury a po menší úpravě také o APKS s jinými bezpředložkovými pády.

Hlavním přínosem disertační práce je zlepšení úspěšnosti pravidlové morfologické disambiguace LanGr za využití znalostí APKS a jejich struktur. Kromě toho práce sumarizuje aktuální poznatky o APKS a shromážděné seznamy bude např. možné zahrnout do databáze víceslovných jednotek LEMUR. Praktická část ilustruje způsob myšlení stojící za tvorbou ručních pravidel pro morfologickou disambiguaci LanGr ve všech jednotlivých krocích. Nadto podává informace o stavu systému LanGr téměř patnáct let od jeho prvního použití a dodává chybějící část dokumentace. V rámci zpracovávání jazykového materiálu také vzniklo několik skriptů v jazyce Python zjednodušujících proces anotace korpusových dat a jejich vyhodnocování, skripty jsou jednoduše dostupné k využití dalšími výzkumníky. V neposlední řadě práce poukazuje na možná vylepšení korpusového prohlížeče Kontext (zejm. uvedení slovního druhu při zobrazování kolokací).

Zdroje

Použité nástroje a korpusy

Morfologický analyzátor ajka. Brno: Masarykova univerzita, Fakulta informatiky, NLP Centrum. [cit. 2018-04-27]. Dostupné z: <https://nlp.fi.muni.cz/projekty/wwwajka/WwwAjkaSkripty/morph.cgi>.

Internetová jazyková příručka (2008–2022). Praha: Ústav pro jazyk český AV ČR, v. v. i. Dostupné z: <https://prirucka.ujc.cas.cz/>.

BARANČIKOVÁ, P. a V. KETTNEROVÁ (2018). *ParaDi 2.0 (2018-01-24)*. LINDAT/CLARIAH-CZ digital library at the Institute of Formal and Applied Linguistics (ÚFAL), Faculty of Mathematics and Physics, Charles University. [cit. 2019-09-28]. Dostupné z: <http://hdl.handle.net/11234/1-2605>.

HAIČ, J., et al. (2018). *Prague Dependency Treebank 3.5*. LINDAT/CLARIAH-CZ digital library at the Institute of Formal and Applied Linguistics (ÚFAL), Faculty of Mathematics and Physics, Charles University. Dostupné z: <http://hdl.handle.net/11234/1-2621>.

HAIČ, J., J. HLAVÁČOVÁ, M. MIKULOVÁ, M. STRAKA a B. ŠTĚPÁNKOVÁ (2020). *MorfFlex CZ 2.0*. LINDAT/CLARIN digital library at the Institute of Formal and Applied Linguistics (ÚFAL), Faculty of Mathematics and Physics, Charles University. Dostupné z: <http://hdl.handle.net/11234/1-3186>.

KŘEN, M., V. CVRČEK, J. HENYŠ, M. HNÁTKOVÁ, T. JELÍNEK, J. KOCEK, D. KOVÁŘIKOVÁ, J. KŘIVAN, J. MILIČKA, V. PETKEVIČ, P. PROCHÁZKA, H. SKOUMALOVÁ, J. ŠINDLEROVÁ a M. ŠKRABAL (2020). *SYN2020: reprezentativní korpus psané češtiny*. Praha: Ústav Českého národního korpusu FF UK. Dostupné z: <https://www.korpus.cz>.

KŘEN, M., V. CVRČEK, M. HNÁTKOVÁ, M. JELÍNEK, J. KOCEK, D. KOVÁŘIKOVÁ, J. KŘIVAN, J. MILIČKA, V. PETKEVIČ, P. PROCHÁZKA, H. SKOUMALOVÁ, J. ŠINDLEROVÁ a M. ŠKRABAL (2022). *Korpus SYN, verze 10 z 22. 2. 2022*. Praha: Ústav Českého národního korpusu FF UK. Dostupné z: <https://www.korpus.cz>.

LOPATKOVÁ, M., V. KETTNEROVÁ, A. VERNEROVÁ, E. BEJČEK a Z. ŽABOKRTSKÝ (2020). *VALLEX 4.0*. LINDAT/CLARIAH-CZ Digital Library at the Institute of Formal and Applied Linguistics, Faculty of Mathematics and Physics, Charles University. Dostupné z: <http://hdl.handle.net/11234/1-3524>.

MACHÁLEK, T. (2014). *KonText – aplikace pro práci s jazykovými korpusy*. Praha: FF UK. Dostupné z: <https://kontext.korpus.cz>.

STRAKA, M. a J. STRAKOVÁ (2014). *MorphoDiTa*. Praha: ÚFAL MFF UK. Dostupné z: <http://lindat.mff.cuni.cz/services/morphodita/>.

UREŠOVÁ, Z., A. BÉMOVÁ, E. FUČIKOVÁ, J. HAIČ, V. KOLÁŘOVÁ, M. MIKULOVÁ, P. PAJAS, J. PANEVOVÁ a J. ŠTĚPÁNEK (2021). *PDT-Vallex: Czech Valency lexicon linked to treebanks 4.0*: LINDAT/CLARIAH-CZ digital library at Institute of Formal and Applied Linguistics, Charles University in Prague. Dostupné z: <http://hdl.handle.net/11234/1-3499>.

Literatura

- BARANČIKOVÁ, P. a V. KETTNEROVÁ (2017): ParaDi: Dictionary of Paraphrases of Czech Complex Predicates with Light Verbs *Proceedings of the 13th Workshop on Multiword Expressions (MWE 2017)*: Association for Computational Linguistics, s. 1–10.
- BARTOŇ, T., V. CVRČEK, F. ČERMÁK, T. JELÍNEK a V. PETKEVIČ (2009): *Statistiky češtiny*. Praha: Nakladatelství Lidové noviny.
- BEČKA, J. (1933): Kapitoly o českém slovese. *Naše řeč*, 17, č. 8, s. 225–233.
- BEČKA, J. (1946): Vyjadřování slovesné a jmenné. In A. OPRAVIL ed. *První hovory o českém jazyce*. Praha: Kruh přátel českého jazyka, s. 64–79.
- BEČKA, J. V. (1992): *Česká stylistika*. Praha: Academia.
- BEJČEK, E., J. PANEVOVÁ, J. POPELKA, P. STRAŇÁK, M. ŠEVČÍKOVÁ, J. ŠTĚPÁNEK a Z. ŽABOKRTSKÝ (2012): Prague Dependency Treebank 2.5 – a Revisited Version of PDT 2.0. In M. KAY a C. BOITET eds. *COLING 2012, 24th International Conference on Computational Linguistics, Proceedings of the Conference: Technical Papers, 8-15 December 2012, Mumbai, India*: Indian Institute of Technology Bombay, s. 231–246.
- BOJANOWSKI, P., E. GRAVE, A. JOULIN a T. MIKOLOV (2017): Enriching Word Vectors with Subword Information. *Transactions of the Association for Computational Linguistics*, 5, s. 135–146.
- BRILL, E. (1992): A simple rule-based part of speech tagger *Proceedings of the third conference on Applied natural language processing (ANLC '92)*, Stroudsburg, PA, USA: Association for Computational Linguistics, s. 152–155.
- BRINTON, L. a M. AKIMOTO (1999): *Collocational and Idiomatic Aspects of Composite Predicates in the History of English*. Amsterdam – Philadelphia: John Benjamins.
- BUTT, M. J. (2003): The Light Verb Jungle. *Harvard Working Papers in Linguistics*, s. 1–49.
- BUTT, M. J. (2010): The light verb jungle: still hacking away. *Complex predicates: cross-linguistic perspectives on event structure*, s. 48–78.
- CATTELL, R. (1984): *Composite Predicates in English*. Sydney – New York – London: Academic Press.
- CINKOVÁ, S. (2009): *Words that matter: Towards a Swedish-Czech Colligational Lexicon of Basic Verbs*. Disertační práce. Ústav formální a aplikované lingvistiky MFF UK.
- CVRČEK, V. a O. RICHTEROVÁ (2022a): cnk:syn2020:lemmatizace. In *Příručka ČNK*. [cit. 2022-06-13]. Dostupné z: <http://wiki.korpus.cz/doku.php?id=cnk:syn2020:lemmatizace&rev=1608771216>.
- CVRČEK, V. a O. RICHTEROVÁ (2022b): seznamy:tagy. In *Příručka ČNK*. [cit. 2022-06-11]. Dostupné z: <http://wiki.korpus.cz/doku.php?id=seznamy:tagy&rev=1650299545>.
- ČERMÁK, F. (1974): Víceslovná pojmenování typu verbum-substantivum v češtině. *Slovo a slovesnost*, 35, č. 4, s. 287–306.

- ČERMÁK, F. (2017a): Frazém a idiom. In P. KARLÍK, M. NEKULA A J. PLESKALOVÁ eds. *CzechEncy – Nový encyklopedický slovník češtiny*. [cit. 2022-05-23]. Dostupné z: <https://www.czechency.org/slovník/FRAZ%C3%89M%20A%20IDIOM>.
- ČERMÁK, F. (2017b): *Korpus a korpusová lingvistika*. Praha: Univerzita Karlova, nakladatelství Karolinum.
- ČERMÁK, F. (2017c): Kvazifrazém. In P. KARLÍK, M. NEKULA A J. PLESKALOVÁ eds. *CzechEncy – Nový encyklopedický slovník češtiny*. [cit. 2022-05-23]. Dostupné z: <https://www.czechency.org/slovník/KVAZIFRAZ%C3%89M>.
- ČERMÁK, F. (2017d): Verbo-nominální frazém. In P. KARLÍK, M. NEKULA A J. PLESKALOVÁ eds. *CzechEncy – Nový encyklopedický slovník češtiny*. [cit. 2022-05-23]. Dostupné z: <https://www.czechency.org/slovník/VERBO-NOMIN%C3%81LN%C3%8D%20FRAZ%C3%89M>.
- ČERMÁK, F. a J. HRONEK (2009): *Slovník české frazeologie a idiomatiky 3, Výrazy slovesné*. Praha: Leda.
- DANEŠ, F., M. GREPL a Z. HLAVSA (1987): *Mluvnice češtiny 3, Skladba*. Praha: Academia.
- GOLÁŇOVÁ, H. (2011): Novočeský lexikální archiv a excerptce v průběhu let 1911–2011. *Slovo a slovesnost*, 72, č. 4, s. 287–300.
- GREGOR, J. (2008): *Verbonominální spojení v ruštině (v částečné konfrontaci s češtinou)*. České Budějovice: Vysoká škola evropských a regionálních studií.
- GRIMSHAW, J. B. (1990): *Argument structure*. Cambridge: The MIT Press.
- GROSS, G. (1989): *Les constructions converses du français*. Geneve: Librairie Droz.
- GROSS, M. (1981): Les bases empiriques de la notion de prédicat sémantique. *Langages*, 15, č. 63, s. 7–52.
- HAIJČ, J. (2004): *Disambiguation of Rich Inflection*. Praha: Karolinum.
- HELBIG, G. (1984): Probleme der Beschreibung von Funktionsverbgefügen im Deutschen. In *Studien zur deutschen Syntax*. Leipzig: VEB Verlag Enzyklopädie, s. 163–188.
- HLADKÁ, Z. (2017): Desémantizace. In P. KARLÍK, M. NEKULA A J. PLESKALOVÁ eds. *CzechEncy - Nový encyklopedický slovník češtiny*. [cit. 2022-07-24]. Dostupné z: <https://www.czechency.org/slovník/DES%C3%89MANTIZACE>.
- HNÁTKOVÁ, M. (2002): Značkování frazémů a idiomů v Českém národním korpusu s pomocí Slovníku české frazeologie a idiomatiky. *Slovo a slovesnost*, 63, s. 117–126.
- HNÁTKOVÁ, M., T. JELÍNEK, V. PETKEVIČ, A. ROSEN, H. SKOUMALOVA, M. KOPŘIVOVÁ, P. VONDŘIČKA, M. KOPŘIVOVÁ a P. VONDŘIČKA (2018): Lepší vrabec v hrsti nežli holub na střeše. Víceslovné lexikální jednotky v češtině: typologie a slovník. *Korpus – gramatika – axiologie*, č. 17, s. 3–22.
- HNÁTKOVÁ, M., M. KŘEN, P. PROCHÁZKA a H. SKOUMALOVÁ (2014): The SYN-series corpora of written Czech *Proceedings of LREC 2014*, Reykjavík: ELRA, s. 160–164.

- HNÁTKOVÁ, M. a V. PETKEVIČ (2017): Morphological Disambiguation of Multiword Expressions and Its Impact on the Disambiguation of Their Environment in a Sentence. *Jazykovedný časopis*, 68, s. 145–155.
- CHROMÝ, J. a V. CVRČEK (2021): Lingvistika jako otevřená a transparentní disciplína. *Naše řeč*, 104, č. 1, s. 5–14.
- IBRAHIM, A. H. (2007): Intruduction. In Z. HARRIS ed. *La langue et l'information*. Paris: Cellule de recherche en linguistique.
- JAKUBÍČEK, M., V. KOVÁŘ a P. ŠMERK (2011): Czech Morphological Tagset Revisited. In HORÁK a RYCHLÝ eds. *Proceedings of Recent Advances in Slavonic Natural Language Processing 2011*: Tribun EU, s. 29–42.
- JEDLIČKA, A. (1969): Univerbizace a multiverbizace v pojmenovacích strukturách. *Acta Universitatis Carolinae, Slavica Pragensia*, 11, s. 93–101.
- JELÍNEK, M. (1967): Jména dějová. In M. DOKULIL, J. KUCHAR A F. DANEŠ eds. *Tvoření slov v češtině 2, Odvozování podstatných jmen*. Praha: Academia, s. 562–653.
- JELÍNEK, M. (2002): Verbonominální spojení v českých obrozenských textech. In *Sborník prací Filozoficko-přírodovědecké fakulty Slezské univerzity v Opavě*. Opava: D2, s. 72–89.
- JELÍNEK, M. (2003): O verbonominálních spojeních ve spisovné češtině *Přednášky a besedy z 36. běhu Letní školy slovanských studií*: Masarykova univerzita, s. 37–51.
- JELÍNEK, M. (2007): Verbonominální synonyma ke slovesům *Přednášky a besedy z 40. běhu Letní školy slovanských studií*: Masarykova univerzita, s. 40–49.
- JELÍNEK, T. (2008): Nové značkování v Českém národním korpusu. *Naše řeč*, 91, č. 1, s. 13–20.
- JELÍNEK, T., J. KŘIVAN, V. PETKEVIČ, H. SKOUMALOVÁ a J. ŠINDLEROVÁ (2021): SYN2020: A New Corpus of Czech with an Innovated Annotation *Text, Speech, and Dialogue: 24th International Conference, TSD 2021, Olomouc, Czech Republic, September 6–9, 2021, Proceedings*: Springer-Verlag, s. 48–59.
- JELÍNEK, T. a V. PETKEVIČ (2011): Systém jazykového značkování korpusů současné psané češtiny. In V. PETKEVIČ a A. ROSEN eds. *Korpusová lingvistika Praha 2011, sv. 3: Gramatika a značkování korpusů*. Praha: Nakladatelství Lidové noviny, s. 154–170.
- JELÍNEK, T., V. PETKEVIČ a H. SKOUMALOVÁ (2022): Hodnoty slovesných morfologických kategorií v korpusu SYN2020 – atribut verbtage. *Časopis pro moderní filologii*, 104, č. 1, s. 89–109.
- JESPERSEN, O. (1942): *A Modern English Grammar on Historical Principles, VI. Morphology*. Copenhagen: Ejnar Munksgaard.
- KARLSSON, F., A. VOUTILAINEN, J. HEIKKILÄ a A. ANTILA (1995): *Constraint Grammar: a language-independent system for parsing unrestricted text*. Berlin a New York: Mouton de Gruyter.
- KETTNEROVÁ, V. (2017): Syntaktická struktura komplexních predikátů v češtině. *Slovo a*

slovesnost, 78, č. 1, s. 3–24.

KETTNEROVÁ, V. a E. BEJČEK (2016): Distribution of Valency Complements in Czech Complex Predicates: Between Verb and Noun: European Language Resources Association (ELRA), s. 515–521.

KETTNEROVÁ, V., V. KOLÁŘOVÁ a A. VERNEROVÁ (2017): Deverbal Nouns in Czech Light Verb Constructions: Springer International Publishing, s. 205–219.

KETTNEROVÁ, V. a M. LOPATKOVÁ (2013): The Representation of Czech Light Verb Constructions in a Valency Lexicon. In E. HAJIČOVÁ, K. GERDES a L. WANNER eds. *Proceedings of DepLing 2013*: Matfyzpress, s. 147–156.

KETTNEROVÁ, V. a M. LOPATKOVÁ (2015): At the Lexicon-Grammar Interface: The Case of Complex Predicates in the Functional Generative Description *Proceedings of the Third International Conference on Dependency Linguistics (Depling 2015)*: Uppsala University, s. 191–200.

KETTNEROVÁ, V. a M. LOPATKOVÁ (2017): Ke koreferenci u komplexních predikátů s kategoriálním slovesem. *Korpus – gramatika – axiologie*, 16, s. 3–26.

KETTNEROVÁ, V. a M. LOPATKOVÁ (2020): Reciprocity in Czech Light Verb Constructions: The Dependency Perspective. *Jazykovedný časopis*, 71, č. 1, s. 41–68.

KETTNEROVÁ, V., M. LOPATKOVÁ a P. BARANČIKOVÁ (2016): Lexicographic Description of Czech Complex Predicates: Between Lexicon and Grammar. In G. M. TINATIN MARGALITADZE ed. *Proceedings of the 17th EURALEX International Congress*, Tbilisi, Georgia: Ivane Javakhishvili Tbilisi University Press, s. 881–892.

KETTNEROVÁ, V., M. LOPATKOVÁ a E. BEJČEK (2012): The Syntax-Semantics Interface of Czech Verbs in the Valency Lexicon *Euralex International Congress 2012*: University of Oslo.

KETTNEROVÁ, V., M. LOPATKOVA, E. BEJČEK a P. BARANČIKOVÁ (2018): Enriching VALLEX with Light Verbs: From Theory to Data and Back Again. *The Prague Bulletin of Mathematical Linguistics*, 111, s. 29–56.

KETTNEROVÁ, V., M. LOPATKOVÁ, E. BEJČEK, A. VERNEROVÁ a M. PODOBOVÁ (2013): Corpus Based Identification of Czech Light Verbs *Proceedings of the Seventh International Conference Slovo 2013; Natural Language Processing, Corpus Linguistics, E-learning*: RAM-Verlag, s. 118–128.

KOLÁŘOVÁ, V. (2010): *Valence deverbativních substantiv v češtině (na materiálu substantiv s dativní valenci)*. Praha: Karolinum Press.

KOLÁŘOVÁ, V. (2014): Nominalizované struktury se dvěma aktanty ve formě bezpředložkového genitivu. *Naše řeč*, 97, č. 4–5, s. 286–299.

KRALČÁK, Ľ. (2005): *Analytické verbo-nominálne spojenia v slovenčine. Synchronno-diachronný pohľad*. Nitra: Univerzita Konštantína Filozofa.

KŘIVAN, J. a J. ŠINDLEROVÁ (2022): Změny v morfologické anotaci korpusů řady SYN: nové možnosti zkoumání české gramatiky a lexikonu. *Slovo a slovesnost*, 83, č. 2, s. 122–145.

KVĚTOŇ, P. (2006a): *Rule-based Morphological Disambiguation*. Disertační práce. Univerzita Karlova, Matematicko-fyzikální fakulta, Ústav formální a aplikované lingvistiky.

KVĚTOŇ, P. (2006b): Rule-Based Morphological Disambiguation On Computational Complexity of the LanGR formalism. *The Prague Bulletin of Mathematical Linguistics*, 85, s. 57–72.

KVĚTOŇ, P. (2008): *Rule language*. Dostupné z: <https://grimm.ruk.cuni.cz/langr/public/doc/rl.pdf>.

LOPATKOVÁ, M., V. KETTNEROVÁ, E. BEJČEK, A. VERNEROVÁ a Z. ŽABOKRTSKÝ (2016): *Valenční slovník českých sloves – VALLEX*. Praha: Univerzita Karlova, nakladatelství Karolinum.

LOPATKOVÁ, M., V. KETTNEROVÁ, A. VERNEROVÁ, E. VEJČEK a Z. ŽABOKRTSKÝ (2021): *Valenční slovník českých sloves VALLEX*. Praha: ÚFAL MFF UK. TR-2021-68.

MA, X., Z. HU, J. LIU, N. PENG, G. NEUBIG a E. HOVY (2018): Stack-Pointer Networks for Dependency Parsing *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics: ACL*, s. 1403–1414.

MACHÁČKOVÁ, E. (1979): *Analytická spojení typu sloveso + abstraktní substantivum (analytické vyjadřování predikátů)*. Kandidátská práce. Ústav pro jazyk český ČSAV.

MACHÁČKOVÁ, E. (1983): Analytické predikáty. *Jazykové aktuality*, 20, č. 3–4, s. 122–176.

MACHÁČKOVÁ, E. (1994): Constructions with verbs and abstract nouns in Czech (analytical predicates). In S. ČMEJRKOVÁ a F. ŠTÍCHA eds. *The syntax of sentence and text a festschrift for František Daneš*. Amsterdam, Philadelphia: J. Benjamins Pub. Co., vol. 42, s. 365–374.

MACHÁLEK, T. (2020): KonText: Advanced and Flexible Corpus Query Interface *Proceedings of LREC 2020*, s. 7005–7010.

MARTÍNEK, F. (2008a): Analytická verbonominální spojení v dnešní češtině a v češtině doby střední: příprava ke zkoumání textu Hájkovy Kroniky české *Jazyková interakce a jazykové rozhraní a strategie „cutting-edge“*. *Sborník příspěvků z 8. mezinárodní konference Setkání mladých lingvistů, konané na Filozofické fakultě Univerzity Palackého v Olomouci ve dnech 14.–16. května 2007*, s. 185–195.

MARTÍNEK, F. (2008b): *Analytická verbonominální spojení v Hájkově Kronice české (1541)*. Diplomová práce. Univerzita Karlova, Filozofická fakulta, Ústav českého jazyka a teorie komunikace.

MARTÍNEK, F. (2009): K analytickým verbonominálním spojení v Hájkově Kronice české (1541). *Varia*, XVII, s. 340–351.

MARTÍNEK, F. (2010): Analytická verbonominální spojení – důkaz ovlivnění starší češtiny latinou a němčinou? *Kulturní translace – Kulturelle Translationen – Translacje kulturowe: příspěvky z mezinárodní studentské konference interFaces VII v červnu 2009 v Praze: Akropolis*, s. 37–47.

MARTÍNEK, F. (2011): K valenci u analytických verbonominálních spojení (s příklady ze starší i dnešní češtiny). *Bohemica Olomucensia*, 2, s. 124–132.

- MARTÍNEK, F. (2013): K slovo tvorným typům českých deverbativních dějových (diachronní pohled) *Gramatika a korpus 2012. Sborník příspěvků ze stejnojmenné konference*, 28. – 30. 11. 2012: Gaudeamus.
- MARTÍNEK, F. (2016): *Funktionsverbgefüge im humanistischen Tschechisch*. Disertační práce. Univerzita Karlova v Praze, Filozofická fakulta, Ústav českého jazyka a teorie komunikace; Eberhard Karls Universität Tübingen, Philosophische Fakultät, Slavisches Seminar.
- MATHESIUS, V. (1966): *Řeč a sloh*. Praha: Československý spisovatel.
- MIKULOVÁ, M., et al. (2005): *Anotace na tektogramatické rovině Pražského závislostního korpusu (Anotátorská příručka)*. Praha: Ú. M. UK. TR-2005-28.
- MIKULOVÁ, M., et al. (2006): *Anotace na tektogramatické rovině Pražského závislostního korpusu (Referenční příručka)*. Praha: Ú. M. UK. TR-2006-31.
- NIVRE, J., M.-C. DE MARNEFFE, F. GINTER, J. HAJIČ, C. D. MANNING, S. PYYSALO, S. SCHUSTER, F. TYERS a D. ZEMAN (2020): Universal Dependencies v2: An Evergrowing Multilingual Treebank Collection. In N. CALZOLARI, F. BÉCHET, P. BLACHE, K. CHOUKRI, C. CIERI, T. DECLERCK, S. GOGGI, H. ISAHARA, B. MAEGAARD, J. MARIANI, H. MAZO, A. MORENO, J. ODIJK a S. PIPERIDIS eds. *Proceedings of the 12th Language Resources and Evaluation Conference*, Marseille, France: European Language Resources Association, s. 4034–4043.
- OLIVA, K., M. HNÁTKOVÁ, P. KVĚTOŇ a V. PETKEVIČ (2000): The Linguistic Basis of a Rule-Based Tagger of Czech. In P. SOJKA, I. KOPEČEK a K. PALA eds., Berlin, Heidelberg: Springer-Verlag, s. 3–8.
- OLIVA, K. a P. KVĚTOŇ (2002): (Semi-)Automatic Detection of Errors in PoS- Tagged Corpora. In S. C. TSENG ed. *Proceedings of the 19th International Conference on Computational Linguistics*: Morgan Kaufmann Publishers, s. 509–515.
- OSOLSOBĚ, K., J. HLAVÁČOVÁ, V. PETKEVIČ, M. SVÁŠEK a J. ŠIMANDL (2017): Nová automatická morfolo gická analýza češtiny. *Naše řeč*, 100, č. 4, s. 225–234.
- PANEVOVÁ, J. (2017): Funkční generativní popis. In P. KARLÍK, M. NEKULA a J. PLESKALOVÁ eds. *CzechEncy - Nový encyklopedický slovník češtiny*.
- PETKEVIČ, V. (2004): Využití pravidel pro negaci v automatickém značkování českých korpusů. In Z. HLADKÁ a P. KARLÍK eds. *Čeština – univerzália a specifika 5: Sborník 5. mezinárodního setkání bohemistů v Brně*, Praha: Nakladatelství Lidové noviny, s. 143–150.
- PETKEVIČ, V. (2006): Reliable Morphological Disambiguation of Czech: Rule-Based Approach is Necessary. In *Insight into the Slovak and Czech Corpus Linguistics*. Bratislava: Veda, s. 26–44.
- PETKEVIČ, V. (2014a): Ambiguity, language structures and corpora. *La Linguistique*, 50, č. 2, Le Cercle linguistique de Prague – II: D’hier à aujourd’hui, s. 63–82.
- PETKEVIČ, V. (2014b): Kontrola české gramatiky (český grammar checker). *Studie z aplikované lingvistiky*, 5, č. 2, s. 48–86.
- PETKEVIČ, V. (2014c): *Morfologická homonymie v současné češtině*. Praha: Nakladatelství Lidové noviny.

- PETKEVIČ, V. (2014d): Problémy automatické morfologické disambiguace češtiny. *Naše řeč*, 97, č. 4–5, s. 194–207.
- PETKEVIČ, V. (2015): Slovnědruhová a morfologická homonymie, homofonie a homografie v současné češtině. *Časopis pro moderní filologii*, 97, č. 2, s. 127–135.
- PETKEVIČ, V., J. HLAVÁČOVÁ, K. OSOLSOBĚ, M. SVÁŠEK a J. ŠIMANDL (2019): Parts of Speech in NovaMorf, A New Morphological Annotation of Czech. *Jazykovedný časopis*, 70, č. 2, s. 358–369.
- PETKEVIČ, V. a M. HNÁTKOVÁ (2002): Automatická morfologická disambiguace předložkových skupin v Českém národním korpusu. *ČEŠTINA – univerzália a specifika* 4, s. 243–252.
- PETKEVIČ, V., M. KOPŘIVOVÁ, M. HNÁTKOVÁ, T. JELÍNEK, P. KOPŘIVA, A. ROSEN, H. SKOUMALOVÁ a P. VONDŘIČKA (2020): Typologie víceslovných jednotek v češtině a frekvenční zastoupení jejich hlavních vlastností v žánrově vyváženém korpusu. *Studie z aplikované lingvistiky*, 11, č. 2, s. 37–62.
- POLENZ, P. v. (1963): *Funktionsverben im heutigen Deutsch. Sprache in der rationalisierten Welt*. Düsseldorf: Beiheft 5 zur Zeitschrift „Wirkendes Wort“.
- POŘÍZKA, P. (2014): *Tvorba korpusů a vytěžování jazykových dat: metody, modely, nástroje*. Olomouc: Univerzita Palackého v Olomouci.
- POVOLNÁ, M. (2012): *Verbonominální predikát s kategoriálním slovesem v evropských právních textech*. Diplomová práce. Jihočeská univerzita v Českých Budějovicích, Filozofická fakulta, Ústav romanistiky.
- RADIMSKÝ, J. (2010): *Verbo-nominální predikát s kategoriálním slovesem*. České Budějovice: Jihočeská univerzita v ČB.
- RADIMSKÝ, J. (2012): Vers une description des prédicats analytiques. *La Linguistique*, 48, s. 51–81.
- RADIMSKÝ, J. (2017a): Analytický predikát s kategoriálním slovesem. In P. KARLÍK, M. NEKULA a J. PLESKALOVÁ eds. *CzechEncy – Nový encyklopedický slovník češtiny*. [cit. 2022-05-23]. Dostupné z: <https://www.czechency.org/slovník/ANALYTICK%C3%9D%20PREDIK%C3%81T%20S%20KATEGORI%C3%81LN%C3%8DM%20SLOVESEM>.
- RADIMSKÝ, J. (2017b): Kategoriální sloveso. In P. KARLÍK, M. NEKULA a J. PLESKALOVÁ eds. *CzechEncy – Nový encyklopedický slovník češtiny*. [cit. 2022-05-23]. Dostupné z: <https://www.czechency.org/slovník/KATEGORI%C3%81LN%C3%8D%20SLOVESO>.
- SAG, I. A., T. BALDWIN, F. BOND, A. COPESTAKE a D. FLICKINGER (2002): *Multiword expressions: A pain in the neck for NLP*: Springer, s. 1–15.
- SEDLÁČEK, R. (1999): *Morfologický analyzátor češtiny*. Diplomová práce. Masarykova univerzita, Fakulta informatiky.
- SGALL, P. (1967): *Generativní popis jazyka a česká deklinace*. Praha: Academia, nakladatelství Československé akademie věd.

- SKOUMALOVÁ, H. (2011): Porovnání úspěšnosti tagování korpusu. In V. PETKEVIČ A A. ROSEN eds. *Korpusová lingvistika Praha 2011 – 3 Gramatika a značkování korpusů*. Praha: Nakladatelství lidové noviny, s. 199–207.
- SPOUSTOVÁ, D., J. HAJIČ, J. VOTRUBEC, P. KRBEC a P. KVĚTOŇ (2007): The Best of Two Worlds: Cooperation of Statistical and Rule-Based Taggers for Czech *Proceedings of the Workshop on Balto-Slavonic Natural Language Processing*, s. 67–74.
- SPOUSTOVÁ, D. j., J. HAJIČ, J. RAAB a M. SPOUSTA (2009): Semi-supervised Training for the Averaged Perceptron POS Tagger *Proceedings of the 12th Conference of the European Chapter of the ACL, EACL '09: Association for Computational Linguistics.*, s. 763–771.
- STEHLÍKOVÁ, L. (2010): *Morfosyntaktické vlastnosti deverbálních jmen na -ní/-tí*. Disertační práce. FF MU.
- STRAKA, M., J. STRAKOVÁ a J. HAJIČ (2019): Czech Text Processing with Contextual Embeddings: POS Tagging, Lemmatization, Parsing and NER. In. Cham: Springer International Publishing, s. 137–150.
- STRAKOVÁ, J., M. STRAKA a J. HAJIČ (2014): Open-Source Tools for Morphology, Lemmatization, POS Tagging and Named Entity Recognition *Proceedings of 52nd Annual Meeting of the Association for Computational Linguistics: System Demonstrations: Association for Computational Linguistics.*, s. 13–18.
- ŠMERK, P. (2009): Fast Morphological Analysis of Czech. In P. SOJKA A A. HORÁK eds. *Proceedings of the Raslan Workshop 2009*: Masarykova univerzita, s. 6–9.
- TĚSITELOVÁ, M. (1966): *O morfologické homonymii v češtině*. Praha: Academia.
- VLKOVÁ, V. (1978): K problematice tzv. multiverbizačních spojení, zvláště v odborném stylu. *Slovo a slovesnost*, 39, č. 2, s. 107–115.
- VLKOVÁ, V. (1990): Příspěvek k analýze multiverbálních spojení typu „provádět rekonstrukci“. *Slovo a slovesnost*, 51, č. 1, s. 1–15.
- VOTRUBEC, J. (2005): *Volba vhodné sady rysů pro morfologické značkování češtiny*. Diplomová práce. Univerzita Karlova, Matematicko-fyzikální fakulta, Ústav formální a aplikované lingvistiky.

Seznam zkratek

acc	akuzativ
ACT	Konatel, název aktantu v FGP
ADDR	Adresát, název aktantu v FGP
Adj	adjektivum
APKS	analytický predikát s kategoriálním slovesem
Apl	akuzativ plurálu
Asg	akuzativ singuláru
CPHR	compound phraseme, (část) složeného predikátu – hodnota funktoru v PDT
ČNK	Český národní korpus
dat	dativ
Dpl	dativ plurálu
Dsg	dativ singuláru
DVS	dvourovinná valenční syntax
FF UK	Filozofická fakulta Univerzity Karlovy
FGP	funkčně generativní popis
FI MU	Fakulta informatiky Masarykovy univerzity
gen	genitiv
Gpl	genitiv plurálu
Gsg	genitiv singuláru
inst	instrumentál
Ipl	instrumentál plurálu
Isg	instrumentál singuláru
KWIC	key word in context (klíčové slovo v kontextu) – označení pro hledaný výraz v korpusové konkordanci
LG	Lexique-grammaire (Lexikogramatická škola)
loc	lokál
Lpl	lokál plurálu
Lsg	lokál singuláru
MFF UK	Matematicko-fyzikální fakulta Univerzity Karlovy
MUNI	Masarykova univerzita
nom	nominativ

Npl	nominativ plurálu
Nsg	nominativ singuláru
N-V	pořadí v rámci věty – nejprve substantivum a až po něm sloveso
os.	osoba
PAT	Patiens, název aktantu v FGP
PDT	Prague Dependency Treebank (Pražský závislostní korpus)
pl.	plurál
POS	part of speech – slovní druh, resp. první pozice v morfologickém tagu v korpusech ČNK
Prep	prepozice
Pron	pronomén
sg.	singulár
Subst	substantivum
ÚČNK	Ústav Českého národního korpusu FF UK
ÚFAL	Ústav formální a aplikované lingvistiky MFF UK
UK	Univerzita Karlova
ÚTKL	Ústav teoretické a počítačové lingvistiky FF UK
VALLEX	Valenční slovník českých sloves
Verb	verbum (sloveso)
V-N	pořadí v rámci věty – nejprve sloveso a až po něm substantivum
voc	vokativ
Vpl	vokativ plurálu
Vsg	vokativ singuláru

Zdroje a autoři příkladů

Konkrétní publikace je zřejmá z kontextu; názvy korpusů jsou uvedeny celé.

(JR)	Jan Radimský
(PK)	Pavel Květoň
(VaK et al.)	Václava Ketnerová a kolektiv dalších lingvistů z kruhu PDT
(VaK)	Václava Kettnerová
(VeK)	Veronika Kolářová
(vl)	vlastní příklad, introspekce autorky disertační práce
(VP)	Vladimír Petkevič

Seznam tabulek

Tabulka 1: Ukázka redukčního testu u spojení s kategoriálním a predikativním slovesem Zdroj: Radimský (2017a)	9
Tabulka 2: Derivační škála deverbativ na ose proces–designace podle Radimského Zdroj: Radimský (2010, s. 59–60, 62, 64)	38
Tabulka 3: Srovnání typologií kategoriálního slovesa navržených Radimským a Kettnerovou	46
Tabulka 4: Škála skalární desémantizace kategoriálního slovesa podle Radimského Zdroj: Radimský (2010, s. 121)	48
Tabulka 5: Reprezentace slovního tvaru „je“ v pozičním tagsetu taggeru MorphoDiTa Zdroj: Tagger MorphoDiTa, dostupné z http://lindat.mff.cuni.cz/services/morphodita/	64
Tabulka 6: Reprezentace slovního tvaru „je“ v atributivním tagsetu taggeru ajka Zdroj: Tagger ajka, dostupné z https://nlp.fi.muni.cz/projekty/wwwajka/WwwAjkaSkripty/morph.cgi	64
Tabulka 7: Hodnoty přiřazené v rámci anotace jednotlivých vět při analýze potenciálních APKS se slovesem <i>chovat/chovávat</i>	86
Tabulka 8: Shrnutí analyzovaných APKS a jiných spojení se slovesem <i>chovat</i> , procenta se vztahují k celkovému počtu analyzovaných spojení	87
Tabulka 9: Shrnutí analyzovaných APKS se slovesem <i>chovat</i>	87
Tabulka 10: Predikativní jména tvořící APKS se slovesem <i>chovat</i> . Seřazeno podle procentuálního výskytu APKS + vztažné. Hvězdičkou jsou označena spojení uvedená ve VALLEXu a Radimském.	91
Tabulka 11: 10 nejčastějších substantivních kolokátů slovesa <i>chovat</i> podle míry logDice....	93
Tabulka 12: Substantiva tvořící APKS se slovesem <i>chovat</i> seřazená podle procentuální četnosti spojení (sloupec „APKS + vztažné (%)“) a doplněná o hodnotu logDice a pořadí v seznamu této kolokační míry. Hvězdičkou jsou označena lemmata z VALLEXu a Radimského.....	95
Tabulka 13: Přehled struktur APKS, resp. slovních druhů (POS) jednotlivých tokenů v anotovaných KWIC (v příkladech tučně); N – substantivum, A – adjektivum, P – pronomen, V – verbum, D – adverbium, R – prepozice, J – konjunkce, T – partikule; sloupec „APKS (%)“ označuje procentuální podíl ze všech analyzovaných APKS (Tabulka 9).....	98
Tabulka 14: Předložky použité ve strukturách V-R-P-N, N-R-N-V se slovesem <i>chovat</i> ; seřazeno podle frekvence; procenta odpovídají podílu jednotlivých předložek v obou strukturách.	100
Tabulka 15: Druhy zájmen u struktur, které tvoří 80 % všech APKS; zkratka v závorce reprezentuje první dvě pozice z tagu pro každé slovo v KWIC; řazeno podle frekvence, jednotlivé popisky viz <i>Příručka ČNK</i> (Cvrček a Richterová 2022b).....	101

Tabulka 16: Přehled struktur pro KWIC s jiným než predikativním významem.....	102
Tabulka 17: Srovnání slovesného způsobu u klauzí obsahujících APKS se slovesem <i>chovat</i> s rozložením v SYN2005, jak ho uvádí publikace <i>Statistiky češtiny</i> (Bartoň et al. 2009)	102
Tabulka 18: Distribuce zkoumaných struktur v anotovaných datech	108
Tabulka 19: Přehled pádů s morfologickou homonymií s <i>Asg</i> nebo <i>Apl</i>	110

Seznam grafů, obrázků a zdrojových kódů

Seznam grafů

Graf 1: Počet slov mezi kategoriálním slovesem a predikativním jménem, procentuální srovnání mezi N-V a V-N89

Seznam obrázků

Obrázek 1: Schéma dvojí analýzy Zdroj: Radimský (2010, s. 17)13

Obrázek 2: Zobrazení APKS a koreferenčních vztahů mezi jeho členy v PDT – bílé čtverce značí uzly, které nejsou povrchově vyjádřené, oranžové šipky znázorňují koreferenci Zdroj: PDT 3.529

Obrázek 3: Propojení slovesných valenčních doplnění se jmennými doplněními (černé šipky) a jejich saturace nominálními sémantickými participanty v APKS *podat hlášení* Zdroj: Kettnerová a Lopatková (2013, s. 150)31

Obrázek 4: Propojení slovesných valenčních doplnění se jmennými doplněními (černé šipky) a jejich saturace nominálními sémantickými participanty v APKS *podat výkon*, viz zejm. nepropojený a odstraněný Adresát Zdroj: Kettnerová a Lopatková (2013, s. 150)31

Obrázek 5: Zjednodušená tektogramatická struktura věty z příkladu (25) Zdroj: Kettnerová a Bejček (2016, s. 519).....32

Obrázek 6: Ukázka anotace jazykových dat114

Seznam zdrojových kódů

Zdrojový kód 1: Zápis negativního bigramu po sobě následujících finitních sloves v LanGru Zdroj: Květoň (2006a, s. 56); aktualizováno dle současně používané verze LanGru.....76

Zdrojový kód 2: Zápis v LanGru pro tvar *jehož* využívající sjednocené sekvence s omezením Zdroj: Květoň (2006a, s. 68); aktualizováno dle současně používané verze LanGru.....77

Zdrojový kód 3: Ukázka zápisu seznamu (tabulky) APKS v programovacím jazyce LanGr111

Zdrojový kód 4: Zápis pravidla pro základní strukturu APKS v programovacím jazyce LanGr112

Příloha 1

Soupis příkazů, operátorů a komentářů používaných v programovacím jazyce LanGr

Níže následuje ruční soupis příkazů, operátorů a komentářů používaných v programovacím jazyce LanGr, který byl pořízen při studiu vzorových pravidel a komentářů k nim (Květoň 2006a, s. 56–75) a slouží pro ilustraci toho, jaké operace je možné při disambiguaci provádět. S rezervou o něm mohu říci, že se jedná o první ucelenější soupis, protože Květoň sám publikoval pouze odborné články, nikoliv plnou dokumentaci. Soupis také reflektuje aktuální stav, tedy zachycuje změny a úpravy provedené později.

LanGr obsahuje funkce, datové typy a proměnné z RL jádra (tj. programovacího jazyka, ve kterém byl LanGr napsán, viz úvod podkapitoly 4.3) a výkonná klíčová slova (viz seznam níže), které dále doplňují lingvisty definované globální identifikátory (např. *PronounRelativePoss*), skupiny (např. *Adjektiva*) a pravidla samotná (např. *Jehoz*).

Klíčová slova

charset <*charset*> – definice použitého kódování (nutno definovat, pokud pravidlo obsahuje jiné znaky než ASCII)

rule – začátek pravidla

rulevariant – každé pravidlo musí obsahovat alespoň jednu variantu uvnitř *rule*

ITEM <*condition*> – token, který odpovídá po něm následující podmínce (argumentem je podmínka), která musí být splněna

ITEM SentenceStart – začátek věty (nelze uložit do proměnné)

ITEM SentenceEnd – konec věty (nelze uložit do proměnné)

PRE-SENTENCE ITEM <*condition*> – odpovídá začátku věty nebo reálné pozici odpovídající argumentu příkazu *ITEM*

POST-SENTENCE ITEM <*condition*> – odpovídá konci věty nebo reálné pozici odpovídající argumentu příkazu *ITEM*

FACULTATIVE ITEM <*condition*> – na rozdíl od *ITEM* nemusí být přítomný

SEQUENCE OF <*condition*> – sekvence tokenů odpovídající uvedené podmínce; může být prázdná

NonEmpty SEQUENCE OF <condition> – sekvence tokenů odpovídající uvedené podmínce; nesmí být prázdná

REPORT <args> – vytvoří textovou proměnnou, která je zapamatována a použita ve výkonné části, pokaždé když je odstraněn nějaký tag

emphasize (<string>) – funkce, která zvýrazní daný text

SELECT <condition> FROM <sequence> – umožní vybrat subsekvenci ze sekvence (vyjádřené příkazem SEQUENCE)

IF <condition> THEN {<action>} ELSE {<action>} – pokud je splněna podmínka, vykoná akci, pokud není, vykoná jinou (druhá část podmínky nemusí být přítomna)

<condition> HOLDS ON <variable> – vrací hodnotu *true*, pokud je podmínka splněná pro danou proměnnou; používá se v podmínce

UNIFIABLE <variable> WITH <variable> IN <array> – vrací hodnotu *true*, pokud je možné provést sjednocení; používá se v podmínce

Používání tabulek seznamu bigramů a trigramů (nově přidáno)

Nově je možné definovat novou datovou strukturu – seznam bigramů nebo trigramů (viz Zdrojový kód 3). K jeho využití v pravidlech pomocí konstrukce WHERE a HOLDS ON viz Zdrojový kód 4.

Operátory

IsSafe – všechny značky daného tokenu musí splňovat podmínku; slouží k výběru jednoznačného kontextu, ve kterém bude pravidlo operovat

Possible – alespoň jedna ze značek daného tokenu musí splňovat podmínku; slouží k výběru tokenu, který bude dále upravován vytvářeným pravidlem

MustNotBe – žádná ze značek daného tokenu nesmí splňovat podmínku

EMPTY <variable> – daná proměnná je prázdná

member of <array> – požadovaná vlastnost je ve sdíleném seznamu

=, ==, !=. and, not, or

lower form

Výkonné příkazy

DELETE <tags> FROM <position> – odstraní vybranou značku/značky z dané proměnné

LEAVE ONLY <tags> IN <position> – ponechá pouze vybranou značku/značky z dané proměnné (opak příkazu *DELETE*)

UNIFY <sequence> IN <category> – sjednotí vybranou sekvenci pozic (min. dvě) v dané kategorii; *<sequence>* může být také reprezentována konstrukcí *<position> WITH <position>*, přičemž *<position>* může být vybrána ze sekvence pomocí *<position> FROM <position>*

UNIFY CONDITIONALLY <sequence> IN <category> – podmíněné sjednocení, podobně jako *UNIFY* nabízí sjednocení, ale předtím ověří, zda dané pozice mohou být sjednoceny; pokud není sjednocení možné, není příkaz proveden; v případě *UNIFY* by v takovém případě došlo ke smazání všech značek u všech pozic v dané sekvenci

UNIFY UNILATERALLY <sequence> IN <category> – jednostranné sjednocení, sjednotí pouze vybrané kategorie u první sjednocené pozice z dané sekvence a druhá zůstane nedotčená

INSERT <variable> INTO <sequence> – vloží proměnnou do sekvence (a pak na ní provede např. sjednocení)

REMOVE <variable> FROM <sequence> – odebere proměnnou ze sekvence

Komentáře a jejich formátování

// – řádkový komentář

/* */ – blokový komentář

\it{<arg>} – argument je vytisknut kurzívou

\bf{<arg>} – argument je vytisknut tučně

\par – znak nového odstavce

\nl – znak nového řádku

\appliedto{<arg>} – pravidlo se aplikuje na větu v argumentu; vzhledem k tomu, že je používáno při ověřování, zda je pravidlo platné, nesmí obsahovat žádné další formátování

\doesntapplyto{<arg>} – opak výše uvedeného *appliedto*, pravidlo se neaplikuje na větu v argumentu

\itemize – začátek nečíslovaného seznamu

\enditemize – konec nečíslovaného seznamu

\item – položka seznamu

Příloha 2

Počet slov (vzdálenost) mezi kategoriálním slovesem a predikativním jménem v APKS s *chovat/chovávat*

vzdálenost	celkem V-N	celkem V-N (%)	celkem N-V	celkem N-V (%)
0	2132	18,11	453	27,34
1	2406	20,44	94	5,67
2	972	8,26	211	12,73
3	872	7,41	111	6,7
4	468	3,98	61	3,68
5	199	1,69	30	1,81
6	101	0,86	17	1,03
7	55	0,47	12	0,72
8	32	0,27	6	0,36
9	20	0,17	3	0,18
10	18	0,15	2	0,12
11	10	0,08	1	0,06
12	5	0,04	0	0
13	5	0,04	0	0
14	0	0	4	0,24
15	1	0,01	3	0,18
19	2	0,02	0	0
21	2	0,02	1	0,06

Příloha 3

Predikativní jména tvořící APKS se slovesem *chovat* v pořadí V-N

Seřazeno podle procentuálního výskytu APKS + vztažné. Hvězdičkou jsou označena spojení uvedená ve VALLEXu a Radimském.

lemma	celkem	APKS (%)	modus vzdálenosti APKS	ø vzdálenost APKS	jiné (%)
zášť*	299	291 (97,32)	0	1,55	8 (2,68)
nedůvěra*	231	219 (94,81)	0	1,85	12 (5,19)
sympatie*	571	534 (93,52)	1	1,32	37 (6,48)
antipatie	51	47 (92,16)	1	2	4 (7,84)
obdiv*	373	339 (90,88)	1	1,71	34 (9,12)
úcta*	1332	1208 (90,69)	0	1,72	124 (9,31)
náklonnost*	238	211 (88,66)	1	2,2	27 (11,34)
naděje*	678	600 (88,5)	0	0,73	78 (11,5)
skepse	34	30 (88,24)	1	2,7	4 (11,76)
nenávisť*	305	265 (86,89)	0	1,98	40 (13,11)
respekt*	1313	1132 (86,21)	1	1,57	181 (13,79)
cit*	613	522 (85,15)	1	1,91	91 (14,85)
nepřátelství*	65	55 (84,62)	0	1,84	10 (15,38)
despekt	25	21 (84,0)	0	1,71	4 (16,0)
nechuť	23	17 (73,91)	0	1,82	6 (26,09)
nevraživost	19	14 (73,68)	1	2,14	5 (26,32)
opovržení	34	24 (70,59)	1	1,62	10 (29,41)
předsudek	43	30 (69,77)	0	1,27	13 (30,23)
úmysl	62	39 (62,9)	1	1,72	23 (37,1)
podezření*	303	186 (61,39)	0	0,87	117 (38,61)
důvěra*	430	263 (61,16)	1	1,66	167 (38,84)
zloba	23	14 (60,87)	0	1,29	9 (39,13)
iluze	84	51 (60,71)	1	1,08	33 (39,29)
přesvědčení	73	44 (60,27)	0	0,8	29 (39,73)
láska*	386	225 (58,29)	1	1,89	161 (41,71)
víra	90	48 (53,33)	1	1,15	42 (46,67)
vášeň	59	30 (50,85)	3	2,2	29 (49,15)

odpor	213	108 (50,7)	0	1,5	105 (49,3)
přátelství*	62	30 (48,39)	3	2,23	32 (51,61)
ambice	73	35 (47,95)	1	1,54	38 (52,05)
sen	56	26 (46,43)	1	2,5	30 (53,57)
touha	91	36 (39,56)	1	2,19	55 (60,44)
představa	168	60 (35,71)	1	1,63	108 (64,29)
vztah	723	232 (32,09)	1	2,26	491 (67,91)
přízeň	57	18 (31,58)	1	1,22	39 (68,42)
myšlenka	132	36 (27,27)	1	1,58	96 (72,73)
slabost	49	13 (26,53)	0	1	36 (73,47)
obava	161	38 (23,6)	1	0,95	123 (76,4)
přání	87	18 (20,69)	1	1,33	69 (79,31)
pocit	639	120 (18,78)	1	1,97	519 (81,22)
záměr	76	11 (14,47)	1	2,64	65 (85,53)
postoj	212	25 (11,79)	1	1,96	187 (88,21)
obraz	91	9 (9,89)	1	1,78	82 (90,11)
plán	116	8 (6,9)	3	1,88	108 (93,1)
ohled	100	5 (5,0)	1	2	95 (95,0)
názor	459	11 (2,4)	3	1,82	448 (97,6)
potřeba	376	2 (0,53)	0	1,5	374 (99,47)

Příloha 4

Predikativní jména tvořící APKS se slovesem *chovat* v pořadí N-V

Seřazeno podle procentuálního výskytu APKS + vztažné. Hvězdičkou jsou označena spojení uvedená ve VALLEXu a Radimském.

lemma	celkem	APKS + vztažné %	APKS (%)	modus vzdálenosti APKS	ø vzdálenost APKS	vztažné (%)	jiné (%)
zášť*	52	43 (82,69)	41 (78,85)	0	1,76	2 (3,85)	9 (17,31)
sympatie*	148	114 (77,03)	84 (56,76)	0	1,36	30 (20,27)	34 (22,97)
cít*	267	205 (76,78)	137 (51,31)	0	1,35	68 (25,47)	62 (23,22)
náklonnost*	60	46 (76,67)	30 (50,0)	0	1,33	16 (26,67)	14 (23,33)
obdiv*	98	73 (74,49)	52 (53,06)	0	1,42	21 (21,43)	25 (25,51)
nedůvěra*	64	44 (68,75)	20 (31,25)	0	2	24 (37,5)	20 (31,25)
despekt	6	4 (66,67)	4 (66,67)	4	2,75	0 (0,0)	2 (33,33)
nenávisť*	71	46 (64,79)	32 (45,07)	0	1,22	14 (19,72)	25 (35,21)
antipatie	11	7 (63,64)	6 (54,55)	0	0,83	1 (9,09)	4 (36,36)
úcta*	311	193 (62,06)	111 (35,69)	0	1,68	82 (26,37)	118 (37,94)
zloba	10	6 (60,0)	5 (50,0)	0	1,4	1 (10,0)	4 (40,0)
láska*	254	149 (58,66)	90 (35,43)	0	2,14	59 (23,23)	105 (41,34)
skepse	7	4 (57,14)	2 (28,57)	6	4	2 (28,57)	3 (42,86)
opovržení	7	4 (57,14)	0 (0,0)			4 (57,14)	3 (42,86)
předsudek	43	22 (51,16)	7 (16,28)	0	1	15 (34,88)	21 (48,84)
naděje*	156	79 (50,64)	67 (42,95)	0	2,94	12 (7,69)	77 (49,36)
nechuť	12	6 (50,0)	3 (25,0)	1	1,33	3 (25,0)	6 (50,0)
nevraživost	16	8 (50,0)	6 (37,5)	2	1,67	2 (12,5)	8 (50,0)
nepřátelství*	9	4 (44,44)	2 (22,22)	0	0	2 (22,22)	5 (55,56)
přízeň	36	15 (41,67)	5 (13,89)	3	1,8	10 (27,78)	21 (58,33)
přátelství*	36	12 (33,33)	4 (11,11)	1	0,75	8 (22,22)	24 (66,67)
vášeň	36	12 (33,33)	5 (13,89)	0	1,2	7 (19,44)	24 (66,67)
respekt*	261	86 (32,95)	57 (21,84)	0	1,35	29 (11,11)	175 (67,05)
úmysl	36	11 (30,56)	7 (19,44)	0	2	4 (11,11)	25 (69,44)
důvěra*	179	53 (29,61)	32 (17,88)	0	1,5	21 (11,73)	126 (70,39)
touha	60	14 (23,33)	14 (23,33)	0	0,93	0 (0,0)	46 (76,67)

víra	43	10 (23,26)	9 (20,93)	0	3,89	1 (2,33)	33 (76,74)
ambice	52	9 (17,31)	9 (17,31)	1	1	0 (0,0)	43 (82,69)
odpor	94	14 (14,89)	9 (9,57)	3	2,11	5 (5,32)	80 (85,11)
iluze	55	7 (12,73)	4 (7,27)	0	1	3 (5,45)	48 (87,27)
podezření*	122	15 (12,3)	13 (10,66)	0	0,77	2 (1,64)	107 (87,7)
vztah	591	69 (11,68)	55 (9,31)	0	1,75	14 (2,37)	522 (88,32)
záměr	58	5 (8,62)	3 (5,17)	3	4	2 (3,45)	53 (91,38)
postoj	150	10 (6,67)	7 (4,67)	2	2,43	3 (2,0)	140 (93,33)
myšlenka	129	8 (6,2)	7 (5,43)	0	1,57	1 (0,78)	121 (93,8)
přesvědčení	52	3 (5,77)	3 (5,77)	0	0	0 (0,0)	49 (94,23)
prání	79	4 (5,06)	2 (2,53)	0	2	2 (2,53)	75 (94,94)
sen	199	10 (5,03)	9 (4,52)	0	1,33	1 (0,5)	189 (94,97)
pocit	1205	54 (4,48)	30 (2,49)	0	1,33	24 (1,99)	1151 (95,52)
slabost	23	1 (4,35)	0 (0,0)			1 (4,35)	22 (95,65)
obraz	117	5 (4,27)	3 (2,56)	2	1	2 (1,71)	112 (95,73)
obava	268	11 (4,1)	9 (3,36)	0	0,22	2 (0,75)	257 (95,9)
představa	358	9 (2,51)	6 (1,68)	3	1,5	3 (0,84)	349 (97,49)
plán	106	2 (1,89)	1 (0,94)	5	5	1 (0,94)	104 (98,11)
názor	497	7 (1,41)	6 (1,21)	0	1	1 (0,2)	490 (98,59)
potřeba	248	1 (0,4)	1 (0,4)	0	0	0 (0,0)	247 (99,6)
ohled	67	0 (0,0)	0 (0,0)			0 (0,0)	67 (100,0)

Příloha 5

Číslo predikativních jmen (APKS s kategoriálním slovesem *chovat*)

lemma	singulár (%)	plurál (%)	celkem
ohled	0 (0,0)	5 (100,0)	5
předsudek	1 (1,92)	51 (98,08)	52
sympatie	19 (2,93)	630 (97,07)	649
cít	95 (13,07)	632 (86,93)	727
iluze	10 (17,24)	48 (82,76)	58
pocit	32 (18,39)	142 (81,61)	174
antipatie	11 (20,37)	43 (79,63)	54
obava	10 (20,41)	39 (79,59)	49
ambice	9 (20,45)	35 (79,55)	44
úmysl	16 (31,37)	35 (68,63)	51
názor	6 (33,33)	12 (66,67)	18
myšlenka	21 (47,73)	23 (52,27)	44
záměr	8 (50,0)	8 (50,0)	16
představa	36 (52,17)	33 (47,83)	69
postoj	23 (65,71)	12 (34,29)	35
plán	7 (70,0)	3 (30,0)	10
sen	28 (77,78)	8 (22,22)	36
obraz	11 (78,57)	3 (21,43)	14
přání	18 (81,82)	4 (18,18)	22
vztah	252 (83,72)	49 (16,28)	301
naděje	579 (85,15)	101 (14,85)	680
touha	43 (86,0)	7 (14,0)	50
zloba	18 (90,0)	2 (10,0)	20
přesvědčení	43 (91,49)	4 (8,51)	47
podezření	185 (92,04)	16 (7,96)	201
víra	57 (98,28)	1 (1,72)	58
láska	369 (98,66)	5 (1,34)	374
nenávisť	308 (99,04)	3 (0,96)	311
náklonnost	256 (99,61)	1 (0,39)	257
úcta	1396 (99,64)	5 (0,36)	1401
opovržení	28 (100,0)	0 (0,0)	28
vášeň	42 (100,0)	0 (0,0)	42
slabost	14 (100,0)	0 (0,0)	14

skepse	34 (100,0)	0 (0,0)	34
respekt	1218 (100,0)	0 (0,0)	1218
přízeň	33 (100,0)	0 (0,0)	33
přátelství	42 (100,0)	0 (0,0)	42
potřeba	3 (100,0)	0 (0,0)	3
nepřátelství	59 (100,0)	0 (0,0)	59
odpor	122 (100,0)	0 (0,0)	122
obdiv	412 (100,0)	0 (0,0)	412
nevraživost	22 (100,0)	0 (0,0)	22
nedůvěra	263 (100,0)	0 (0,0)	263
nechuť	23 (100,0)	0 (0,0)	23
důvěra	316 (100,0)	0 (0,0)	316
despekt	25 (100,0)	0 (0,0)	25
zášť	334 (100,0)	0 (0,0)	334

Příloha 6

APKS s typem koreference ACT_{subst} - ACT_{verb}

Tato příloha je převzatá ze studie Kettnerové a Lopatkové (2017, s. 21–23) s názvem *Ke koreferenci u komplexních predikátů s kategoriálním slovesem*, konkrétně z přílohy A. Seznam APKS byl využit v druhé případové studii (viz 5.3.2); ~~přeskrtnuté~~ APKS sice původní seznam obsahuje, ale do případové studie zahrnuté nebyly.

činit *krok, narážku, opatření, pokání, pokrok, pokus, rozhodnutí, rozdíl, skutek, úkon, závěr;*

činit si *nárok;*

dělat *chybu, kompromis, kontrolu, muziku, pokrok, poznámky, práci, rámus, reklamu, rozdíl, závěr;*

~~**dělat si** *iluze, legraci, nárok, srandu;*~~

dostat, dostávat *chuť, nápad, smyk, strach, záchvat;*

chovat *cit, naděje, náklonnost, nedůvěru, nenávisť, obdiv, podezření, respekt, sympatie, úctu, zášť;*

jevit *chuť, nadšení, neklid, nezájem, ochotu, podobnost, sklon, snahu, tendenci, úmysl, zájem;*

konat *pouť, přednášku, skutek;*

mít *čas, chuť, kontrolu, možnost, nárok, obavu, období, pocit, potíž, pravdu, právo, problém, radost, smůlu, smysl, strach, šanci, štěstí, vliv, vztah, zájem, zkušenost;*

najít, nacházet *klid, kompromis, náhradu, odvahu, odpověď, rovnováhu, řešení, uplatnění, východisko, vysvětlení;*

nalézt/naleznout, nalézat *důkaz, klid, kompromis, možnost, odpověď, odvahu, rovnováhu, řešení, shodu, uplatnění, vysvětlení;*

nasbírat *silu, zkušenost;*

nést *jméno, název, odpovědnost/zodpovědnost, označení, riziko, vinu;*

podat, podávat *demisi, dovolání, odvolání, protest, přihlášku, svědectví, výklad, výkon, žalobu;*

podniknout, podnikat *akci, cestu, invazi, kroky, nájezd, obhlídku, opatření, pátrání, pokus, pouť, procházky, protiútok, přípravu, šetření, tah, tažení, útok, vycházky, výlet, výpad, výpravu, výstup;*

pojmout, pojímat *ideu, nápad, nedůvěru, nenávisť, plán, podezření, přání, sympatie, úmysl, záměr;*

pořídít, pořizovat dokumentaci, fotku, fotografii, kopii, náčrt, opis, překlad, snímek, soupis, výbavu, zápis, záznam;

projevit, projevovat cit, horlivost, lhostejnost, lítost, nadání, nadšení, nelibost, nesouhlas, nespokojenost, obavu, odvahu, ochotu, optimismus, pochopení, přání, sklon, snahu, souhlas, tendenci, údiv, vstřícnost, vůli, zájem, zvědavost;

provést, provádět analýzu, cvičení, cvik, dokazování, experiment, kontrolu, korekci, korekturu, manévr, masáž, měření, ohledání, operaci, opravu, pitvu, pokus, prohlídku, průzkum, převrat, razii, rekonstrukci, revizi, rozbor, řez, selekci, šetření, test, transakci, údržbu, úklid, úkon, úpravu, vyhodnocení, výpočet, vyšetření, výzkum, zákrok;

provozovat činnost, dopravu, hratky, obchod, praxi, rituál, sex, sport;

přijmout, přijímat odpovědnost/zodpovědnost, opatření, rozhodnutí, usnesení, závazek;

převzít, přebrat, přebírat iniciativu, odpovědnost/zodpovědnost, riziko, ručení, splácení, úkol, vedení, velení, vládu, záštitu, závazek;

sebrat, sbírat kuráž, odvahu, sílu, zkušenost;

učinit doznání, krok, krůček, objev, opatření, oznámení, pokrok, pokus, posuněk, prohlášení, rozhodnutí, skutek, slib, udání, úkon, úsudek, výjimku, závěr;

udělat gesto, chybu, kariéru, krok, krůček, pokrok, pořádek, výjimku, výlet, zkoušku;

~~**udělat si** obrázek, představu;~~

uplatňovat, uplatnit nárok, právo;

uskutečnit, uskutečňovat akci, cvičení, činnost, hovor, koncert, koupi, krok, let, nákup, operaci, plán, pokus, politiku, projekt, průzkum, přenos, reformu, spojení, studium, studii, test, transakci, výpravu, výzkum, záměr;

~~**utvořit si, utvářet si** názor, představu, úsudek, vazbu, vztah;~~

uvést, uvádět důkaz, důvod, prohlášení, přehled, výčet;

uzavřít, uzavírat debatu, vyšetřování;

vést boj, evidenci, kampaň, válku, vyšetřování, záznamy, život;

vydat, vydávat prohlášení, rozhodnutí, svědectví, usnesení, vůni, vyhlášku, výkřik, zápach, zvuk;

vyhlásit, vyhlášovat boj, mír, nezávislost, pátrání, poplach, příměří, rozsudek, řízení, samostatnost, sbírku, stávku, útok, válku;

vyjádřit, vyjadřovat domněnku, lítost, mínění, naději, názor, nedůvěru, nespokojenost, obavu, obdiv, odhodlání, ochotu, podiv, pohrdání, pochybnost, politování, postoj, přání, přesvědčení, skepsi, solidaritu, spokojenost, stanovisko, sympatie, úctu, údiv, znechucení, znepokojení;

vykonat, vykonávat čin, činnost, dohled, dozor, kontrolu, modlitbu, nátlak, návštěvu, obhajobu, obřad, pomstu, popravu, poradenství, potřebu, pouť, povinnost, práci, pravomoc, praxi, prohlídku, rituál, skutek, službu, soulož, správu, úkol, úkon, zkoušku;

vynešt, vynášet rozsudek, soud, verdikt;

vypovídat, vypovědět dohodu, příměří, smlouvu, spolupráci, úmluvu;

vyslovit, vyslovovat dohad, domněnku, kletbu, lítost, myšlenku, námitku, názor, nesouhlas, obavu, obdiv, podezření, pochybnost, pochyby, politování, přání, předpoklad, přesvědčení, připomínku, údiv, věštbu, výhradu;

vytvořit, vytvářet náskok, převahu, tlak;

vyvinout, vyvíjet aktivitu, činnost, energii, iniciativu, nátlak, rychlost, sílu, snahu, tlak, úsilí;

vzdát, vzdávat boj, kariéru, naději, pokus, snahu, účast, zápas;

~~**vzít (si), brát (si)** povinnost, odpovědnost/zodpovědnost;~~

~~**vzít si, brát si** dovolenou, ponaučení, poučení, volno;~~

vznést, vznášet dotaz, námitku, nárok, obvinění, obžalobu, otázku, požadavek, protest, připomínku, stížnost, výhrady, žádost, žalobu;

zaujmout, zaujímat názor, odstup, postoj, postavení, stanovisko;

zastávat myšlenku, názor, postoj, přesvědčení, stanovisko, zásadu;

zastat, zastávat pochůzky, povinnost, práci, sekání;

zaznamenat, zaznamenávat nárůst, návštěvnost, pokles, pokrok, posun, propad, přírůstek, rekord, rozmach, rozvoj, růst, úbytek, úpadek, úspěch, ústup, vzestup, ztrátu;

získat, získávat důvěru, kontrolu, náskok, popularitu, pověst, poznatek, přehled, převahu, sebe-důvěru, vzdělání, zkušenost;

ztratit, ztrácet důvěru, důvěryhodnost, chuť, kontakt, kontrolu, naději, náskok, odvahu, orientaci, přehled, rovnováhu, sebedůvěru, sebekontrolu, schopnost, soudnost, trpělivost, víru, zájem.

Příloha 7

Přehled homonymních tvarů (Asg a Apl) predikativních jmen z Přílohy 6

Tato tabulka je také součástí elektronické přílohy *přiloha_7.csv*, obsahuje navíc údaje o tom, jestli byl záznam v tabulce zpracován automaticky nebo skriptem, a dále případné informace o dalších výskytech v rámci *Internetové jazykové příručky*, kterou vyhledávání poskytlo, a také sloupec se souhrnným údajem o morfologické homonymii obecně (vlastní i nevlastní).

nominativ	akuzativ	číslo	počet tvarů	homonymní tvary
akce	akci	sg	3	Dsg, Asg, Lsg
akce	akce	pl	6	Nsg, Gsg, Vsg, Npl, Apl, Vpl
aktivita	aktivitu	sg	1	Asg
aktivita	aktivitu	pl	4	Gsg, Npl, Apl, Vpl
analýza	analýzu	sg	1	Asg
analýza	analýzy	pl	4	Gsg, Npl, Apl, Vpl
boj	boj	sg	2	Nsg, Asg
boj	boje	pl	4	Gsg, Npl, Apl, Vpl
cesta	cestu	sg	1	Asg
cesta	cesty	pl	4	Gsg, Npl, Apl, Vpl
cit	cit	sg	2	Nsg, Asg
cit	city	pl	4	Npl, Apl, Vpl, Ipl
cvičení	cvičení	sg	10	Nsg, Gsg, Dsg, Asg, Vsg, Lsg, Npl, Gpl, Apl, Vpl
cvičení	cvičení	pl	10	Nsg, Gsg, Dsg, Asg, Vsg, Lsg, Npl, Gpl, Apl, Vpl
cvik	cvik	sg	2	Nsg, Asg
cvik	cviky	pl	4	Npl, Apl, Vpl, Ipl
čas	čas	sg	2	Nsg, Asg
čas	časy	pl	4	Npl, Apl, Vpl, Ipl
čin	čin	sg	2	Nsg, Asg
čin	činy	pl	4	Npl, Apl, Vpl, Ipl
činnost	činnost	sg	2	Nsg, Asg
činnost	činnosti	pl	7	Gsg, Dsg, Vsg, Lsg, Npl, Apl, Vpl
debata	debatu	sg	1	Asg
debata	debaty	pl	4	Gsg, Npl, Apl, Vpl

demise	demisi	sg	3	Dsg, Asg, Lsg
demise	demise	pl	6	Nsg, Gsg, Vsg, Npl, Apl, Vpl
dohad	dohad	sg	2	Nsg, Asg
dohad	dohady	pl	4	Npl, Apl, Vpl, Ipl
dohled	dohled	sg	2	Nsg, Asg
dohled	dohledy	pl	4	Npl, Apl, Vpl, Ipl
dohoda	dohodu	sg	1	Asg
dohoda	dohody	pl	4	Gsg, Npl, Apl, Vpl
dokazování	dokazování	sg	10	Nsg, Gsg, Dsg, Asg, Vsg, Lsg, Npl, Gpl, Apl, Vpl
dokazování	dokazování	pl	10	Nsg, Gsg, Dsg, Asg, Vsg, Lsg, Npl, Gpl, Apl, Vpl
dokumentace	dokumentaci	sg	3	Dsg, Asg, Lsg
dokumentace	dokumentace	pl	6	Nsg, Gsg, Vsg, Npl, Apl, Vpl
domněnka	domněnku	sg	1	Asg
domněnka	domněnky	pl	4	Gsg, Npl, Apl, Vpl
doprava	dopravu	sg	1	Asg
dotaz	dotaz	sg	2	Nsg, Asg
dotaz	dotazy	pl	4	Npl, Apl, Vpl, Ipl
dovolání	dovolání	sg	10	Nsg, Gsg, Dsg, Asg, Vsg, Lsg, Npl, Gpl, Apl, Vpl
dovolání	dovolání	pl	10	Nsg, Gsg, Dsg, Asg, Vsg, Lsg, Npl, Gpl, Apl, Vpl
doznání	doznání	sg	10	Nsg, Gsg, Dsg, Asg, Vsg, Lsg, Npl, Gpl, Apl, Vpl
doznání	doznání	pl	10	Nsg, Gsg, Dsg, Asg, Vsg, Lsg, Npl, Gpl, Apl, Vpl
dozor	dozor	sg	2	Nsg, Asg
dozor	dozory	pl	4	Npl, Apl, Vpl, Ipl
důkaz	důkaz	sg	2	Nsg, Asg
důkaz	důkazy	pl	4	Npl, Apl, Vpl, Ipl
důvěra	důvěru	sg	1	Asg
důvěra	důvěry	pl	4	Gsg, Npl, Apl, Vpl
důvěryhod- nost	důvěryhodnost	sg	2	Nsg, Asg
důvěryhod- nost	důvěryhod- nosti	pl	7	Gsg, Dsg, Vsg, Lsg, Npl, Apl, Vpl
důvod	důvod	sg	2	Nsg, Asg
důvod	důvody	pl	4	Npl, Apl, Vpl, Ipl
energie	energii	sg	3	Dsg, Asg, Lsg
energie	energie	pl	6	Nsg, Gsg, Vsg, Npl, Apl, Vpl
evidence	evidenci	sg	3	Dsg, Asg, Lsg
evidence	evidence	pl	6	Nsg, Gsg, Vsg, Npl, Apl, Vpl

experiment	experiment	sg	2	Nsg, Asg
experiment	experimenty	pl	4	Npl, Apl, Vpl, Ipl
fotka	fotku	sg	1	Asg
fotka	fotky	pl	4	Gsg, Npl, Apl, Vpl
fotografie	fotografii	sg	3	Dsg, Asg, Lsg
fotografie	fotografie	pl	6	Nsg, Gsg, Vsg, Npl, Apl, Vpl
gesto	gesto	sg	3	Nsg, Asg, Vsg
gesto	gesta	pl	4	Gsg, Npl, Apl, Vpl
horlivost	horlivost	sg	2	Nsg, Asg
horlivost	horlivosti	pl	7	Gsg, Dsg, Vsg, Lsg, Npl, Apl, Vpl
hovor	hovor	sg	2	Nsg, Asg
hovor	hovory	pl	4	Npl, Apl, Vpl, Ipl
hrátky	hrátky	pl	3	Npl, Apl, Vpl
chuť	chuť	sg	2	Nsg, Asg
chuť	chutě	pl	4	Gsg, Npl, Apl, Vpl
chuť	chuti	pl	7	Gsg, Dsg, Vsg, Lsg, Npl, Apl, Vpl
chyba	chybu	sg	1	Asg
chyba	chyby	pl	4	Gsg, Npl, Apl, Vpl
idea	ideu	sg	1	Asg
idea	idey	pl	4	Gsg, Npl, Apl, Vpl
idea	ideje	pl	4	Gsg, Npl, Apl, Vpl
iniciativa	iniciativu	sg	1	Asg
iniciativa	iniciativy	pl	4	Gsg, Npl, Apl, Vpl
invaze	invazi	sg	3	Dsg, Asg, Lsg
invaze	invaze	pl	6	Nsg, Gsg, Vsg, Npl, Apl, Vpl
jméno	jméno	sg	3	Nsg, Asg, Vsg
jméno	jména	pl	4	Gsg, Npl, Apl, Vpl
kampaň	kampaň	sg	2	Nsg, Asg
kampaň	kampaně	pl	4	Gsg, Npl, Apl, Vpl
kariéra	kariéru	sg	1	Asg
kariéra	kariéry	pl	4	Gsg, Npl, Apl, Vpl
kletba	kletbu	sg	1	Asg
kletba	kletby	pl	4	Gsg, Npl, Apl, Vpl
klid	klid	sg	2	Nsg, Asg
klid	klidy	pl	4	Npl, Apl, Vpl, Ipl
kompromis	kompromis	sg	2	Nsg, Asg
kompromis	kompromisy	pl	4	Npl, Apl, Vpl, Ipl
koncert	koncert	sg	2	Nsg, Asg
koncert	koncerty	pl	4	Npl, Apl, Vpl, Ipl
kontakt	kontakt	sg	2	Nsg, Asg
kontakt	kontakty	pl	4	Npl, Apl, Vpl, Ipl
kontrola	kontrolu	sg	1	Asg
kontrola	kontroly	pl	4	Gsg, Npl, Apl, Vpl

kopie	kopii	sg	3	Dsg, Asg, Lsg
kopie	kopie	pl	6	Nsg, Gsg, Vsg, Npl, Apl, Vpl
korekce	korekci	sg	3	Dsg, Asg, Lsg
korekce	korekce	pl	6	Nsg, Gsg, Vsg, Npl, Apl, Vpl
korektura	korekturu	sg	1	Asg
korektura	korektury	pl	4	Gsg, Npl, Apl, Vpl
koupě	koupi	sg	3	Dsg, Asg, Lsg
koupě	koupě	pl	6	Nsg, Gsg, Vsg, Npl, Apl, Vpl
krok	krok	sg	2	Nsg, Asg
krok	kroky	pl	4	Npl, Apl, Vpl, Ipl
krůček	krůček	sg	2	Nsg, Asg
krůček	krůčky	pl	4	Npl, Apl, Vpl, Ipl
kuráž	kuráž	sg	2	Nsg, Asg
kuráž	kuráže	pl	4	Gsg, Npl, Apl, Vpl
let	let	sg	2	Nsg, Asg
let	lety	pl	4	Npl, Apl, Vpl, Ipl
lhostejnost	lhostejnost	sg	2	Nsg, Asg
lhostejnost	lhostejnosti	pl	7	Gsg, Dsg, Vsg, Lsg, Npl, Apl, Vpl
lítost	lítost	sg	2	Nsg, Asg
lítost	lítosti	pl	7	Gsg, Dsg, Vsg, Lsg, Npl, Apl, Vpl
manévr	manévr	sg	2	Nsg, Asg
manévr	manévry	pl	4	Npl, Apl, Vpl, Ipl
masáž	masáž	sg	2	Nsg, Asg
masáž	masáže	pl	4	Gsg, Npl, Apl, Vpl
měření	měření	sg	10	Nsg, Gsg, Dsg, Asg, Vsg, Lsg, Npl, Gpl, Apl, Vpl
měření	měření	pl	10	Nsg, Gsg, Dsg, Asg, Vsg, Lsg, Npl, Gpl, Apl, Vpl
mínění	mínění	sg	10	Nsg, Gsg, Dsg, Asg, Vsg, Lsg, Npl, Gpl, Apl, Vpl
mínění	mínění	pl	10	Nsg, Gsg, Dsg, Asg, Vsg, Lsg, Npl, Gpl, Apl, Vpl
mír	mír	sg	2	Nsg, Asg
mír	míry	pl	4	Npl, Apl, Vpl, Ipl
modlitba	modlitbu	sg	1	Asg
modlitba	modlitby	pl	4	Gsg, Npl, Apl, Vpl
možnost	možnost	sg	2	Nsg, Asg
možnost	možnosti	pl	7	Gsg, Dsg, Vsg, Lsg, Npl, Apl, Vpl
muzika	muziku	sg	1	Asg
muzika	muziky	pl	4	Gsg, Npl, Apl, Vpl
myšlenka	myšlenku	sg	1	Asg
myšlenka	myšlenky	pl	4	Gsg, Npl, Apl, Vpl
náčrt	náčrt	sg	2	Nsg, Asg

náčrt	náčrty	pl	4	Npl, Apl, Vpl, Ipl
nadání	nadání	sg	10	Nsg, Gsg, Dsg, Asg, Vsg, Lsg, Npl, Gpl, Apl, Vpl
nadání	nadání	pl	10	Nsg, Gsg, Dsg, Asg, Vsg, Lsg, Npl, Gpl, Apl, Vpl
naděje	naději	sg	3	Dsg, Asg, Lsg
naděje	naděje	pl	6	Nsg, Gsg, Vsg, Npl, Apl, Vpl
nadšení	nadšení	sg	10	Nsg, Gsg, Dsg, Asg, Vsg, Lsg, Npl, Gpl, Apl, Vpl
nadšení	nadšení	pl	10	Nsg, Gsg, Dsg, Asg, Vsg, Lsg, Npl, Gpl, Apl, Vpl
náhrada	náhradu	sg	1	Asg
náhrada	náhrady	pl	4	Gsg, Npl, Apl, Vpl
nájezd	nájezd	sg	2	Nsg, Asg
nájezd	nájezdy	pl	4	Npl, Apl, Vpl, Ipl
nákup	nákup	sg	2	Nsg, Asg
nákup	nákupy	pl	4	Npl, Apl, Vpl, Ipl
námitka	námitku	sg	1	Asg
námitka	námitky	pl	4	Gsg, Npl, Apl, Vpl
nápad	nápad	sg	2	Nsg, Asg
nápad	nápady	pl	4	Npl, Apl, Vpl, Ipl
narážka	narážku	sg	1	Asg
narážka	narážky	pl	4	Gsg, Npl, Apl, Vpl
nárok	nárok	sg	2	Nsg, Asg
nárok	nároky	pl	4	Npl, Apl, Vpl, Ipl
nárůst	nárůst	sg	2	Nsg, Asg
nárůst	nárůsty	pl	4	Npl, Apl, Vpl, Ipl
náskok	náskok	sg	2	Nsg, Asg
náskok	náskoky	pl	4	Npl, Apl, Vpl, Ipl
nátlak	nátlak	sg	2	Nsg, Asg
nátlak	nátlaky	pl	4	Npl, Apl, Vpl, Ipl
návštěva	návštěvu	sg	1	Asg
návštěva	návštěvy	pl	4	Gsg, Npl, Apl, Vpl
návštěvnost	návštěvnost	sg	2	Nsg, Asg
návštěvnost	návštěvnosti	pl	7	Gsg, Dsg, Vsg, Lsg, Npl, Apl, Vpl
název	název	sg	2	Nsg, Asg
název	názvy	pl	4	Npl, Apl, Vpl, Ipl
názor	názor	sg	2	Nsg, Asg
názor	názory	pl	4	Npl, Apl, Vpl, Ipl
nedůvěra	nedůvěru	sg	1	Asg
nedůvěra	nedůvěry	pl	4	Gsg, Npl, Apl, Vpl
neklid	neklid	sg	2	Nsg, Asg
neklid	neklidy	pl	4	Npl, Apl, Vpl, Ipl

nelibost	nelibost	sg	2	Nsg, Asg
nelibost	nelibosti	pl	7	Gsg, Dsg, Vsg, Lsg, Npl, Apl, Vpl
nenávist	nenávist	sg	2	Nsg, Asg
nenávist	nenávisti	pl	7	Gsg, Dsg, Vsg, Lsg, Npl, Apl, Vpl
nesouhlas	nesouhlas	sg	2	Nsg, Asg
nesouhlas	nesouhlasy	pl	4	Npl, Apl, Vpl, Ipl
nespokojenost	nespokojenost	sg	2	Nsg, Asg
nespokojenost	nespokojenosti	pl	7	Gsg, Dsg, Vsg, Lsg, Npl, Apl, Vpl
nezájem	nezájem	sg	2	Nsg, Asg
nezájem	nezájmy	pl	4	Npl, Apl, Vpl, Ipl
nezávislost	nezávislost	sg	2	Nsg, Asg
nezávislost	nezávislosti	pl	7	Gsg, Dsg, Vsg, Lsg, Npl, Apl, Vpl
obava	obavu	sg	1	Asg
obava	obavy	pl	4	Gsg, Npl, Apl, Vpl
obdiv	obdiv	sg	2	Nsg, Asg
obdiv	obdivy	pl	4	Npl, Apl, Vpl, Ipl
obdoba	obdobu	sg	1	Asg
obdoba	obdoby	pl	4	Gsg, Npl, Apl, Vpl
obhajoba	obhajobu	sg	1	Asg
obhajoba	obhajoby	pl	4	Gsg, Npl, Apl, Vpl
obhlídka	obhlídku	sg	1	Asg
obhlídka	obhlídky	pl	4	Gsg, Npl, Apl, Vpl
obchod	obchod	sg	2	Nsg, Asg
obchod	obchody	pl	4	Npl, Apl, Vpl, Ipl
objev	objev	sg	2	Nsg, Asg
objev	objevy	pl	4	Npl, Apl, Vpl, Ipl
obřad	obřad	sg	2	Nsg, Asg
obřad	obřady	pl	4	Npl, Apl, Vpl, Ipl
obvinění	obvinění	sg	10	Nsg, Gsg, Dsg, Asg, Vsg, Lsg, Npl, Gpl, Apl, Vpl
obvinění	obvinění	pl	10	Nsg, Gsg, Dsg, Asg, Vsg, Lsg, Npl, Gpl, Apl, Vpl
obžaloba	obžalobu	sg	1	Asg
obžaloba	obžaloby	pl	4	Gsg, Npl, Apl, Vpl
odhodlání	odhodlání	sg	10	Nsg, Gsg, Dsg, Asg, Vsg, Lsg, Npl, Gpl, Apl, Vpl
odhodlání	odhodlání	pl	10	Nsg, Gsg, Dsg, Asg, Vsg, Lsg, Npl, Gpl, Apl, Vpl
odpověď	odpověď	sg	2	Nsg, Asg
odpověď	odpovědi	pl	7	Gsg, Dsg, Vsg, Lsg, Npl, Apl, Vpl
odpovědnost	zodpovědnost	sg	2	Nsg, Asg

odpovědnost	zodpovědnosti	pl	7	Gsg, Dsg, Vsg, Lsg, Npl, Apl, Vpl
odstup	odstup	sg	2	Nsg, Asg
odstup	odstupy	pl	4	Npl, Apl, Vpl, Ipl
odvaha	odvahu	sg	1	Asg
odvaha	odvahy	pl	4	Gsg, Npl, Apl, Vpl
odvolání	odvolání	sg	10	Nsg, Gsg, Dsg, Asg, Vsg, Lsg, Npl, Gpl, Apl, Vpl
odvolání	odvolání	pl	10	Nsg, Gsg, Dsg, Asg, Vsg, Lsg, Npl, Gpl, Apl, Vpl
ohledání	ohledání	sg	10	Nsg, Gsg, Dsg, Asg, Vsg, Lsg, Npl, Gpl, Apl, Vpl
ohledání	ohledání	pl	10	Nsg, Gsg, Dsg, Asg, Vsg, Lsg, Npl, Gpl, Apl, Vpl
ochota	ochotu	sg	1	Asg
ochota	ochoty	pl	4	Gsg, Npl, Apl, Vpl
opatření	opatření	sg	10	Nsg, Gsg, Dsg, Asg, Vsg, Lsg, Npl, Gpl, Apl, Vpl
opatření	opatření	pl	10	Nsg, Gsg, Dsg, Asg, Vsg, Lsg, Npl, Gpl, Apl, Vpl
operace	operaci	sg	3	Dsg, Asg, Lsg
operace	operace	pl	6	Nsg, Gsg, Vsg, Npl, Apl, Vpl
opis	opis	sg	2	Nsg, Asg
opis	opisy	pl	4	Npl, Apl, Vpl, Ipl
oprava	opravu	sg	1	Asg
oprava	opravy	pl	4	Gsg, Npl, Apl, Vpl
optimismus	optimismus	sg	2	Nsg, Asg
optimismus	optimismy	pl	4	Npl, Apl, Vpl, Ipl
orientace	orientaci	sg	3	Dsg, Asg, Lsg
orientace	orientace	pl	6	Nsg, Gsg, Vsg, Npl, Apl, Vpl
otázka	otázku	sg	1	Asg
otázka	otázky	pl	4	Gsg, Npl, Apl, Vpl
označení	označení	sg	10	Nsg, Gsg, Dsg, Asg, Vsg, Lsg, Npl, Gpl, Apl, Vpl
označení	označení	pl	10	Nsg, Gsg, Dsg, Asg, Vsg, Lsg, Npl, Gpl, Apl, Vpl
oznámení	oznámení	sg	10	Nsg, Gsg, Dsg, Asg, Vsg, Lsg, Npl, Gpl, Apl, Vpl
oznámení	oznámení	pl	10	Nsg, Gsg, Dsg, Asg, Vsg, Lsg, Npl, Gpl, Apl, Vpl
pátrání	pátrání	sg	10	Nsg, Gsg, Dsg, Asg, Vsg, Lsg, Npl, Gpl, Apl, Vpl
pátrání	pátrání	pl	10	Nsg, Gsg, Dsg, Asg, Vsg, Lsg, Npl, Gpl, Apl, Vpl

pitva	pitvu	sg	1	Asg
pitva	pitvy	pl	4	Gsg, Npl, Apl, Vpl
plán	plán	sg	2	Nsg, Asg
plán	plány	pl	4	Npl, Apl, Vpl, Ipl
pocit	pocit	sg	2	Nsg, Asg
pocit	pocity	pl	4	Npl, Apl, Vpl, Ipl
podezření	podezření	sg	10	Nsg, Gsg, Dsg, Asg, Vsg, Lsg, Npl, Gpl, Apl, Vpl
podezření	podezření	pl	10	Nsg, Gsg, Dsg, Asg, Vsg, Lsg, Npl, Gpl, Apl, Vpl
podiv	podiv	sg	2	Nsg, Asg
podiv	podivy	pl	4	Npl, Apl, Vpl, Ipl
podobnost	podobnost	sg	2	Nsg, Asg
podobnost	podobnosti	pl	7	Gsg, Dsg, Vsg, Lsg, Npl, Apl, Vpl
pohrdání	pohrdání	sg	10	Nsg, Gsg, Dsg, Asg, Vsg, Lsg, Npl, Gpl, Apl, Vpl
pohrdání	pohrdání	pl	10	Nsg, Gsg, Dsg, Asg, Vsg, Lsg, Npl, Gpl, Apl, Vpl
pochopení	pochopení	sg	10	Nsg, Gsg, Dsg, Asg, Vsg, Lsg, Npl, Gpl, Apl, Vpl
pochopení	pochopení	pl	10	Nsg, Gsg, Dsg, Asg, Vsg, Lsg, Npl, Gpl, Apl, Vpl
pochůzka	pochůzku	sg	1	Asg
pochůzka	pochůzky	pl	4	Gsg, Npl, Apl, Vpl
pochybnost	pochybnost	sg	2	Nsg, Asg
pochybnost	pochybnosti	pl	7	Gsg, Dsg, Vsg, Lsg, Npl, Apl, Vpl
pochyby	pochybu	sg	1	Asg
pochyby	pochyby	pl	4	Gsg, Npl, Apl, Vpl
pokání	pokání	sg	10	Nsg, Gsg, Dsg, Asg, Vsg, Lsg, Npl, Gpl, Apl, Vpl
pokání	pokání	pl	10	Nsg, Gsg, Dsg, Asg, Vsg, Lsg, Npl, Gpl, Apl, Vpl
pokles	pokles	sg	2	Nsg, Asg
pokles	poklesy	pl	4	Npl, Apl, Vpl, Ipl
pokrok	pokrok	sg	2	Nsg, Asg
pokrok	pokroky	pl	4	Npl, Apl, Vpl, Ipl
pokus	pokus	sg	2	Nsg, Asg
pokus	pokusy	pl	4	Npl, Apl, Vpl, Ipl
politika	politiku	sg	1	Asg
politika	politiky	pl	4	Gsg, Npl, Apl, Vpl
politování	politování	sg	10	Nsg, Gsg, Dsg, Asg, Vsg, Lsg, Npl, Gpl, Apl, Vpl

politování	politování	pl	10	Nsg, Gsg, Dsg, Asg, Vsg, Lsg, Npl, Gpl, Apl, Vpl
pomsta	pomstu	sg	1	Asg
pomsta	pomsty	pl	4	Gsg, Npl, Apl, Vpl
poplach	poplach	sg	2	Nsg, Asg
poplach	poplachy	pl	4	Npl, Apl, Vpl, Ipl
poprava	popravu	sg	1	Asg
poprava	popravy	pl	4	Gsg, Npl, Apl, Vpl
popularita	popularitu	sg	1	Asg
popularita	popularity	pl	4	Gsg, Npl, Apl, Vpl
poradenství	poradenství	sg	10	Nsg, Gsg, Dsg, Asg, Vsg, Lsg, Npl, Gpl, Apl, Vpl
poradenství	poradenství	pl	10	Nsg, Gsg, Dsg, Asg, Vsg, Lsg, Npl, Gpl, Apl, Vpl
pořádek	pořádek	sg	2	Nsg, Asg
pořádek	pořádky	pl	4	Npl, Apl, Vpl, Ipl
postavení	postavení	sg	10	Nsg, Gsg, Dsg, Asg, Vsg, Lsg, Npl, Gpl, Apl, Vpl
postavení	postavení	pl	10	Nsg, Gsg, Dsg, Asg, Vsg, Lsg, Npl, Gpl, Apl, Vpl
postoj	postoj	sg	2	Nsg, Asg
postoj	postoje	pl	4	Gsg, Npl, Apl, Vpl
posun	posun	sg	2	Nsg, Asg
posun	posuny	pl	4	Npl, Apl, Vpl, Ipl
posuněk	posuněk	sg	2	Nsg, Asg
posuněk	posuněky	pl	4	Npl, Apl, Vpl, Ipl
potíž	potíž	sg	2	Nsg, Asg
potíž	potíže	pl	4	Gsg, Npl, Apl, Vpl
potřeba	potřebu	sg	1	Asg
potřeba	potřeby	pl	4	Gsg, Npl, Apl, Vpl
pout'	pout'	sg	2	Nsg, Asg
pout'	poutě	pl	4	Gsg, Npl, Apl, Vpl
pout'	pouti	pl	7	Gsg, Dsg, Vsg, Lsg, Npl, Apl, Vpl
pověst	pověst	sg	2	Nsg, Asg
pověst	pověsti	pl	7	Gsg, Dsg, Vsg, Lsg, Npl, Apl, Vpl
povinnost	povinnost	sg	2	Nsg, Asg
povinnost	povinnosti	pl	7	Gsg, Dsg, Vsg, Lsg, Npl, Apl, Vpl
poznámka	poznámku	sg	1	Asg
poznámka	poznámky	pl	4	Gsg, Npl, Apl, Vpl
poznatek	poznatek	sg	2	Nsg, Asg
poznatek	poznatky	pl	4	Npl, Apl, Vpl, Ipl
požadavek	požadavek	sg	2	Nsg, Asg
požadavek	požadavky	pl	4	Npl, Apl, Vpl, Ipl

práce	práci	sg	3	Dsg, Asg, Lsg
práce	práce	pl	6	Nsg, Gsg, Vsg, Npl, Apl, Vpl
pravda	pravdu	sg	1	Asg
pravda	pravdy	pl	4	Gsg, Npl, Apl, Vpl
právo	právo	sg	3	Nsg, Asg, Vsg
právo	práva	pl	4	Gsg, Npl, Apl, Vpl
pravomoc	pravomoc	sg	2	Nsg, Asg
pravomoc	pravomoce	pl	4	Gsg, Npl, Apl, Vpl
pravomoc	pravomoci	pl	7	Gsg, Dsg, Vsg, Lsg, Npl, Apl, Vpl
praxe	praxi	sg	3	Dsg, Asg, Lsg
praxe	praxe	pl	6	Nsg, Gsg, Vsg, Npl, Apl, Vpl
problém	problém	sg	2	Nsg, Asg
problém	problémy	pl	4	Npl, Apl, Vpl, Ipl
prohlášení	prohlášení	sg	10	Nsg, Gsg, Dsg, Asg, Vsg, Lsg, Npl, Gpl, Apl, Vpl
prohlášení	prohlášení	pl	10	Nsg, Gsg, Dsg, Asg, Vsg, Lsg, Npl, Gpl, Apl, Vpl
prohlídka	prohlídku	sg	1	Asg
prohlídka	prohlídky	pl	4	Gsg, Npl, Apl, Vpl
procházka	procházku	sg	1	Asg
procházka	procházky	pl	4	Gsg, Npl, Apl, Vpl
projekt	projekt	sg	2	Nsg, Asg
projekt	projekty	pl	4	Npl, Apl, Vpl, Ipl
propad	propad	sg	2	Nsg, Asg
propad	propady	pl	4	Npl, Apl, Vpl, Ipl
protest	protest	sg	2	Nsg, Asg
protest	protesty	pl	4	Npl, Apl, Vpl, Ipl
protiútok	protiútok	sg	2	Nsg, Asg
protiútok	protiútoky	pl	4	Npl, Apl, Vpl, Ipl
průzkum	průzkum	sg	2	Nsg, Asg
průzkum	průzkumy	pl	4	Npl, Apl, Vpl, Ipl
přání	přání	sg	10	Nsg, Gsg, Dsg, Asg, Vsg, Lsg, Npl, Gpl, Apl, Vpl
přání	přání	pl	10	Nsg, Gsg, Dsg, Asg, Vsg, Lsg, Npl, Gpl, Apl, Vpl
přednáška	přednášku	sg	1	Asg
přednáška	přednášky	pl	4	Gsg, Npl, Apl, Vpl
předpoklad	předpoklad	sg	2	Nsg, Asg
předpoklad	předpoklady	pl	4	Npl, Apl, Vpl, Ipl
přehled	přehled	sg	2	Nsg, Asg
přehled	přehledy	pl	4	Npl, Apl, Vpl, Ipl
překlad	překlad	sg	2	Nsg, Asg
překlad	překlady	pl	4	Npl, Apl, Vpl, Ipl

přenos	přenos	sg	2	Nsg, Asg
přenos	přenosy	pl	4	Npl, Apl, Vpl, Ipl
přesvědčení	přesvědčení	sg	10	Nsg, Gsg, Dsg, Asg, Vsg, Lsg, Npl, Gpl, Apl, Vpl
přesvědčení	přesvědčení	pl	10	Nsg, Gsg, Dsg, Asg, Vsg, Lsg, Npl, Gpl, Apl, Vpl
převaha	převahu	sg	1	Asg
převaha	převahy	pl	4	Gsg, Npl, Apl, Vpl
převrat	převrat	sg	2	Nsg, Asg
převrat	převraty	pl	4	Npl, Apl, Vpl, Ipl
příhláška	příhlášku	sg	1	Asg
příhláška	příhlášky	pl	4	Gsg, Npl, Apl, Vpl
příměří	příměří	sg	10	Nsg, Gsg, Dsg, Asg, Vsg, Lsg, Npl, Gpl, Apl, Vpl
příměří	příměří	pl	10	Nsg, Gsg, Dsg, Asg, Vsg, Lsg, Npl, Gpl, Apl, Vpl
připomínka	připomínku	sg	1	Asg
připomínka	připomínky	pl	4	Gsg, Npl, Apl, Vpl
příprava	přípravu	sg	1	Asg
příprava	přípravy	pl	4	Gsg, Npl, Apl, Vpl
přírůstek	přírůstek	sg	2	Nsg, Asg
přírůstek	přírůstky	pl	4	Npl, Apl, Vpl, Ipl
radost	radost	sg	2	Nsg, Asg
radost	radosti	pl	7	Gsg, Dsg, Vsg, Lsg, Npl, Apl, Vpl
rámus	rámus	sg	2	Nsg, Asg
rámus	rámusy	pl	4	Npl, Apl, Vpl, Ipl
razie	razii	sg	3	Dsg, Asg, Lsg
razie	razie	pl	6	Nsg, Gsg, Vsg, Npl, Apl, Vpl
reforma	reformu	sg	1	Asg
reforma	reformy	pl	4	Gsg, Npl, Apl, Vpl
reklama	reklamu	sg	1	Asg
reklama	reklamy	pl	4	Gsg, Npl, Apl, Vpl
rekonstrukce	rekonstrukci	sg	3	Dsg, Asg, Lsg
rekonstrukce	rekonstrukce	pl	6	Nsg, Gsg, Vsg, Npl, Apl, Vpl
rekord	rekord	sg	2	Nsg, Asg
rekord	rekordy	pl	4	Npl, Apl, Vpl, Ipl
revize	revizi	sg	3	Dsg, Asg, Lsg
revize	revize	pl	6	Nsg, Gsg, Vsg, Npl, Apl, Vpl
rituál	rituál	sg	2	Nsg, Asg
rituál	rituály	pl	4	Npl, Apl, Vpl, Ipl
riziko	riziko	sg	3	Nsg, Asg, Vsg
riziko	rizika	pl	4	Gsg, Npl, Apl, Vpl
rovnováha	rovnováhu	sg	1	Asg

rovnováha	rovnováhy	pl	4	Gsg, Npl, Apl, Vpl
rozbor	rozbor	sg	2	Nsg, Asg
rozbor	rozbory	pl	4	Npl, Apl, Vpl, Ipl
rozdíl	rozdíl	sg	2	Nsg, Asg
rozdíl	rozdíly	pl	4	Npl, Apl, Vpl, Ipl
rozhodnutí	rozhodnutí	sg	10	Nsg, Gsg, Dsg, Asg, Vsg, Lsg, Npl, Gpl, Apl, Vpl
rozhodnutí	rozhodnutí	pl	10	Nsg, Gsg, Dsg, Asg, Vsg, Lsg, Npl, Gpl, Apl, Vpl
rozmach	rozmach	sg	2	Nsg, Asg
rozmach	rozmachy	pl	4	Npl, Apl, Vpl, Ipl
rozsudek	rozsudek	sg	2	Nsg, Asg
rozsudek	rozsudky	pl	4	Npl, Apl, Vpl, Ipl
rozvoj	rozvoj	sg	2	Nsg, Asg
rozvoj	rozvoje	pl	4	Gsg, Npl, Apl, Vpl
ručení	ručení	sg	10	Nsg, Gsg, Dsg, Asg, Vsg, Lsg, Npl, Gpl, Apl, Vpl
ručení	ručení	pl	10	Nsg, Gsg, Dsg, Asg, Vsg, Lsg, Npl, Gpl, Apl, Vpl
růst	růst	sg	2	Nsg, Asg
růst	růsty	pl	4	Npl, Apl, Vpl, Ipl
rychlost	rychlost	sg	2	Nsg, Asg
rychlost	rychlosti	pl	7	Gsg, Dsg, Vsg, Lsg, Npl, Apl, Vpl
řešení	řešení	sg	10	Nsg, Gsg, Dsg, Asg, Vsg, Lsg, Npl, Gpl, Apl, Vpl
řešení	řešení	pl	10	Nsg, Gsg, Dsg, Asg, Vsg, Lsg, Npl, Gpl, Apl, Vpl
řez	řez	sg	2	Nsg, Asg
řez	řezy	pl	4	Npl, Apl, Vpl, Ipl
řízení	řízení	sg	10	Nsg, Gsg, Dsg, Asg, Vsg, Lsg, Npl, Gpl, Apl, Vpl
řízení	řízení	pl	10	Nsg, Gsg, Dsg, Asg, Vsg, Lsg, Npl, Gpl, Apl, Vpl
samostatnost	samostatnost	sg	2	Nsg, Asg
samostatnost	samostatnosti	pl	7	Gsg, Dsg, Vsg, Lsg, Npl, Apl, Vpl
sbírka	sbírku	sg	1	Asg
sbírka	sbírky	pl	4	Gsg, Npl, Apl, Vpl
sebedůvěra	sebedůvěru	sg	1	Asg
sebedůvěra	sebedůvěry	pl	4	Gsg, Npl, Apl, Vpl
sebekontrola	sebekontrolu	sg	1	Asg
sebekontrola	sebekontroly	pl	4	Gsg, Npl, Apl, Vpl
sekání	sekání	sg	10	Nsg, Gsg, Dsg, Asg, Vsg, Lsg, Npl, Gpl, Apl, Vpl

sekání	sekání	pl	10	Nsg, Gsg, Dsg, Asg, Vsg, Lsg, Npl, Gpl, Apl, Vpl
selekce	selekci	sg	3	Dsg, Asg, Lsg
selekce	selekce	pl	6	Nsg, Gsg, Vsg, Npl, Apl, Vpl
sex	sex	sg	2	Nsg, Asg
sex	sexy	pl	4	Npl, Apl, Vpl, Ipl
shoda	shodu	sg	1	Asg
shoda	shody	pl	4	Gsg, Npl, Apl, Vpl
schopnost	schopnost	sg	2	Nsg, Asg
schopnost	schopnosti	pl	7	Gsg, Dsg, Vsg, Lsg, Npl, Apl, Vpl
síla	sílu	sg	1	Asg
síla	síly	pl	4	Gsg, Npl, Apl, Vpl
skepsi	skepsi	sg	3	Dsg, Asg, Lsg
skepsi	skepse	pl	6	Nsg, Gsg, Vsg, Npl, Apl, Vpl
sklon	sklon	sg	2	Nsg, Asg
sklon	sklony	pl	4	Npl, Apl, Vpl, Ipl
skutek	skutek	sg	2	Nsg, Asg
skutek	skutky	pl	4	Npl, Apl, Vpl, Ipl
slib	slib	sg	2	Nsg, Asg
slib	sliby	pl	4	Npl, Apl, Vpl, Ipl
služba	službu	sg	1	Asg
služba	služby	pl	4	Gsg, Npl, Apl, Vpl
smlouva	smlouvu	sg	1	Asg
smlouva	smlouvy	pl	4	Gsg, Npl, Apl, Vpl
smůla	smůlu	sg	1	Asg
smůla	smůly	pl	4	Gsg, Npl, Apl, Vpl
smyk	smyk	sg	2	Nsg, Asg
smyk	smyky	pl	4	Npl, Apl, Vpl, Ipl
smysl	smysl	sg	2	Nsg, Asg
smysl	smysly	pl	4	Npl, Apl, Vpl, Ipl
snaha	snahu	sg	1	Asg
snaha	snahy	pl	4	Gsg, Npl, Apl, Vpl
snímek	snímek	sg	2	Nsg, Asg
snímek	snímky	pl	4	Npl, Apl, Vpl, Ipl
solidarita	solidaritu	sg	1	Asg
solidarita	solidarity	pl	4	Gsg, Npl, Apl, Vpl
soud	soud	sg	2	Nsg, Asg
soud	soudy	pl	4	Npl, Apl, Vpl, Ipl
soudnost	soudnost	sg	2	Nsg, Asg
soudnost	soudnosti	pl	7	Gsg, Dsg, Vsg, Lsg, Npl, Apl, Vpl
souhlas	souhlas	sg	2	Nsg, Asg
souhlas	souhlasy	pl	4	Npl, Apl, Vpl, Ipl
soulož	soulož	sg	2	Nsg, Asg

soulož	soulože	pl	4	Gsg, Npl, Apl, Vpl
soupis	soupis	sg	2	Nsg, Asg
soupis	soupisy	pl	4	Npl, Apl, Vpl, Ipl
spláčení	spláčení	sg	10	Nsg, Gsg, Dsg, Asg, Vsg, Lsg, Npl, Gpl, Apl, Vpl
spláčení	spláčení	pl	10	Nsg, Gsg, Dsg, Asg, Vsg, Lsg, Npl, Gpl, Apl, Vpl
spojení	spojení	sg	10	Nsg, Gsg, Dsg, Asg, Vsg, Lsg, Npl, Gpl, Apl, Vpl
spojení	spojení	pl	10	Nsg, Gsg, Dsg, Asg, Vsg, Lsg, Npl, Gpl, Apl, Vpl
spokojenost	spokojenost	sg	2	Nsg, Asg
spokojenost	spokojenosti	pl	7	Gsg, Dsg, Vsg, Lsg, Npl, Apl, Vpl
spolupráce	spolupráci	sg	3	Dsg, Asg, Lsg
spolupráce	spolupráce	pl	6	Nsg, Gsg, Vsg, Npl, Apl, Vpl
sport	sport	sg	2	Nsg, Asg
sport	sporty	pl	4	Npl, Apl, Vpl, Ipl
správa	správu	sg	1	Asg
správa	správy	pl	4	Gsg, Npl, Apl, Vpl
stanovisko	stanovisko	sg	3	Nsg, Asg, Vsg
stanovisko	stanoviska	pl	4	Gsg, Npl, Apl, Vpl
stávka	stávku	sg	1	Asg
stávka	stávky	pl	4	Gsg, Npl, Apl, Vpl
stížnost	stížnost	sg	2	Nsg, Asg
stížnost	stížnosti	pl	7	Gsg, Dsg, Vsg, Lsg, Npl, Apl, Vpl
strach	strach	sg	2	Nsg, Asg
strach	strachy	pl	4	Npl, Apl, Vpl, Ipl
studie	studii	sg	3	Dsg, Asg, Lsg
studie	studie	pl	6	Nsg, Gsg, Vsg, Npl, Apl, Vpl
studium	studium	sg	3	Nsg, Asg, Vsg
studium	studia	pl	4	Gsg, Npl, Apl, Vpl
svědectví	svědectví	sg	10	Nsg, Gsg, Dsg, Asg, Vsg, Lsg, Npl, Gpl, Apl, Vpl
svědectví	svědectví	pl	10	Nsg, Gsg, Dsg, Asg, Vsg, Lsg, Npl, Gpl, Apl, Vpl
sympatie	sympatii	sg	3	Dsg, Asg, Lsg
sympatie	sympatie	pl	6	Nsg, Gsg, Vsg, Npl, Apl, Vpl
šance	šanci	sg	3	Dsg, Asg, Lsg
šance	šance	pl	6	Nsg, Gsg, Vsg, Npl, Apl, Vpl
šetření	šetření	sg	10	Nsg, Gsg, Dsg, Asg, Vsg, Lsg, Npl, Gpl, Apl, Vpl
šetření	šetření	pl	10	Nsg, Gsg, Dsg, Asg, Vsg, Lsg, Npl, Gpl, Apl, Vpl

štěstí	štěstí	sg	10	Nsg, Gsg, Dsg, Asg, Vsg, Lsg, Npl, Gpl, Apl, Vpl
štěstí	štěstí	pl	10	Nsg, Gsg, Dsg, Asg, Vsg, Lsg, Npl, Gpl, Apl, Vpl
tah	tah	sg	2	Nsg, Asg
tah	tahy	pl	4	Npl, Apl, Vpl, Ipl
tažení	tažení	sg	10	Nsg, Gsg, Dsg, Asg, Vsg, Lsg, Npl, Gpl, Apl, Vpl
tažení	tažení	pl	10	Nsg, Gsg, Dsg, Asg, Vsg, Lsg, Npl, Gpl, Apl, Vpl
tendence	tendenci	sg	3	Dsg, Asg, Lsg
tendence	tendence	pl	6	Nsg, Gsg, Vsg, Npl, Apl, Vpl
test	test	sg	2	Nsg, Asg
test	testy	pl	4	Npl, Apl, Vpl, Ipl
tlak	tlak	sg	2	Nsg, Asg
tlak	tlaky	pl	4	Npl, Apl, Vpl, Ipl
transakce	transakci	sg	3	Dsg, Asg, Lsg
transakce	transakce	pl	6	Nsg, Gsg, Vsg, Npl, Apl, Vpl
trpělivost	trpělivost	sg	2	Nsg, Asg
trpělivost	trpělivosti	pl	7	Gsg, Dsg, Vsg, Lsg, Npl, Apl, Vpl
úbytek	úbytek	sg	2	Nsg, Asg
úbytek	úbytky	pl	4	Npl, Apl, Vpl, Ipl
úcta	úctu	sg	1	Asg
úcta	úcty	pl	4	Gsg, Npl, Apl, Vpl
účast	účast	sg	2	Nsg, Asg
účast	účasti	pl	7	Gsg, Dsg, Vsg, Lsg, Npl, Apl, Vpl
udání	udání	sg	10	Nsg, Gsg, Dsg, Asg, Vsg, Lsg, Npl, Gpl, Apl, Vpl
udání	udání	pl	10	Nsg, Gsg, Dsg, Asg, Vsg, Lsg, Npl, Gpl, Apl, Vpl
údiv	údiv	sg	2	Nsg, Asg
údiv	údivy	pl	4	Npl, Apl, Vpl, Ipl
údržba	údržbu	sg	1	Asg
údržba	údržby	pl	4	Gsg, Npl, Apl, Vpl
úklid	úklid	sg	2	Nsg, Asg
úklid	úklidy	pl	4	Npl, Apl, Vpl, Ipl
úkol	úkol	sg	2	Nsg, Asg
úkol	úkoly	pl	4	Npl, Apl, Vpl, Ipl
úkon	úkon	sg	2	Nsg, Asg
úkon	úkony	pl	4	Npl, Apl, Vpl, Ipl
úmluva	úmluvu	sg	1	Asg
úmluva	úmluvy	pl	4	Gsg, Npl, Apl, Vpl
úmysl	úmysl	sg	2	Nsg, Asg

úmysl	úmysly	pl	4	Npl, Apl, Vpl, Ipl
úpadek	úpadek	sg	2	Nsg, Asg
úpadek	úpadky	pl	4	Npl, Apl, Vpl, Ipl
uplatnění	uplatnění	sg	10	Nsg, Gsg, Dsg, Asg, Vsg, Lsg, Npl, Gpl, Apl, Vpl
uplatnění	uplatnění	pl	10	Nsg, Gsg, Dsg, Asg, Vsg, Lsg, Npl, Gpl, Apl, Vpl
úprava	úpravu	sg	1	Asg
úprava	úpravy	pl	4	Gsg, Npl, Apl, Vpl
úsilí	úsilí	sg	10	Nsg, Gsg, Dsg, Asg, Vsg, Lsg, Npl, Gpl, Apl, Vpl
úsilí	úsilí	pl	10	Nsg, Gsg, Dsg, Asg, Vsg, Lsg, Npl, Gpl, Apl, Vpl
usnesení	usnesení	sg	10	Nsg, Gsg, Dsg, Asg, Vsg, Lsg, Npl, Gpl, Apl, Vpl
usnesení	usnesení	pl	10	Nsg, Gsg, Dsg, Asg, Vsg, Lsg, Npl, Gpl, Apl, Vpl
úspěch	úspěch	sg	2	Nsg, Asg
úspěch	úspěchy	pl	4	Npl, Apl, Vpl, Ipl
ústup	ústup	sg	2	Nsg, Asg
ústup	ústupy	pl	4	Npl, Apl, Vpl, Ipl
úsudek	úsudek	sg	2	Nsg, Asg
úsudek	úsudky	pl	4	Npl, Apl, Vpl, Ipl
útok	útok	sg	2	Nsg, Asg
útok	útoky	pl	4	Npl, Apl, Vpl, Ipl
válka	válku	sg	1	Asg
válka	války	pl	4	Gsg, Npl, Apl, Vpl
vedení	vedení	sg	10	Nsg, Gsg, Dsg, Asg, Vsg, Lsg, Npl, Gpl, Apl, Vpl
vedení	vedení	pl	10	Nsg, Gsg, Dsg, Asg, Vsg, Lsg, Npl, Gpl, Apl, Vpl
velení	velení	sg	10	Nsg, Gsg, Dsg, Asg, Vsg, Lsg, Npl, Gpl, Apl, Vpl
velení	velení	pl	10	Nsg, Gsg, Dsg, Asg, Vsg, Lsg, Npl, Gpl, Apl, Vpl
verdikt	verdikt	sg	2	Nsg, Asg
verdikt	verdikty	pl	4	Npl, Apl, Vpl, Ipl
věštba	věštbu	sg	1	Asg
věštba	věštby	pl	4	Gsg, Npl, Apl, Vpl
vína	vinu	sg	1	Asg
vína	viny	pl	4	Gsg, Npl, Apl, Vpl
víra	víru	sg	1	Asg
víra	víry	pl	4	Gsg, Npl, Apl, Vpl

vláda	vládu	sg	1	Asg
vláda	vlády	pl	4	Gsg, Npl, Apl, Vpl
vliv	vliv	sg	2	Nsg, Asg
vliv	vlivy	pl	4	Npl, Apl, Vpl, Ipl
vstřícnost	vstřícnost	sg	2	Nsg, Asg
vstřícnost	vstřícnosti	pl	7	Gsg, Dsg, Vsg, Lsg, Npl, Apl, Vpl
vůle	vůli	sg	3	Dsg, Asg, Lsg
vůle	vůle	pl	6	Nsg, Gsg, Vsg, Npl, Apl, Vpl
vůně	vůni	sg	3	Dsg, Asg, Lsg
vůně	vůně	pl	6	Nsg, Gsg, Vsg, Npl, Apl, Vpl
výbava	výbavu	sg	1	Asg
výbava	výbavy	pl	4	Gsg, Npl, Apl, Vpl
výčet	výčet	sg	2	Nsg, Asg
výčet	výčty	pl	4	Npl, Apl, Vpl, Ipl
vyhláška	vyhlášku	sg	1	Asg
vyhláška	vyhlášky	pl	4	Gsg, Npl, Apl, Vpl
vyhodnocení	vyhodnocení	sg	10	Nsg, Gsg, Dsg, Asg, Vsg, Lsg, Npl, Gpl, Apl, Vpl
vyhodnocení	vyhodnocení	pl	10	Nsg, Gsg, Dsg, Asg, Vsg, Lsg, Npl, Gpl, Apl, Vpl
výhrada	výhradu	sg	1	Asg
výhrada	výhrady	pl	4	Gsg, Npl, Apl, Vpl
vycházka	vycházku	sg	1	Asg
vycházka	vycházky	pl	4	Gsg, Npl, Apl, Vpl
východisko	východisko	sg	3	Nsg, Asg, Vsg
východisko	východiska	pl	4	Gsg, Npl, Apl, Vpl
výjimka	výjimku	sg	1	Asg
výjimka	výjimky	pl	4	Gsg, Npl, Apl, Vpl
výklad	výklad	sg	2	Nsg, Asg
výklad	výklady	pl	4	Npl, Apl, Vpl, Ipl
výkon	výkon	sg	2	Nsg, Asg
výkon	výkony	pl	4	Npl, Apl, Vpl, Ipl
výkřik	výkřik	sg	2	Nsg, Asg
výkřik	výkřiky	pl	4	Npl, Apl, Vpl, Ipl
výlet	výlet	sg	2	Nsg, Asg
výlet	výlety	pl	4	Npl, Apl, Vpl, Ipl
výpad	výpad	sg	2	Nsg, Asg
výpad	výpady	pl	4	Npl, Apl, Vpl, Ipl
výpočet	výpočet	sg	2	Nsg, Asg
výpočet	výpočty	pl	4	Npl, Apl, Vpl, Ipl
výprava	výpravu	sg	1	Asg
výprava	výpravy	pl	4	Gsg, Npl, Apl, Vpl
výstup	výstup	sg	2	Nsg, Asg

výstup	výstupy	pl	4	Npl, Apl, Vpl, Ipl
vysvětlení	vysvětlení	sg	10	Nsg, Gsg, Dsg, Asg, Vsg, Lsg, Npl, Gpl, Apl, Vpl
vysvětlení	vysvětlení	pl	10	Nsg, Gsg, Dsg, Asg, Vsg, Lsg, Npl, Gpl, Apl, Vpl
vyšetření	vyšetření	sg	10	Nsg, Gsg, Dsg, Asg, Vsg, Lsg, Npl, Gpl, Apl, Vpl
vyšetření	vyšetření	pl	10	Nsg, Gsg, Dsg, Asg, Vsg, Lsg, Npl, Gpl, Apl, Vpl
vyšetřování	vyšetřování	sg	10	Nsg, Gsg, Dsg, Asg, Vsg, Lsg, Npl, Gpl, Apl, Vpl
vyšetřování	vyšetřování	pl	10	Nsg, Gsg, Dsg, Asg, Vsg, Lsg, Npl, Gpl, Apl, Vpl
výzkum	výzkum	sg	2	Nsg, Asg
výzkum	výzkumy	pl	4	Npl, Apl, Vpl, Ipl
vzdělání	vzdělání	sg	10	Nsg, Gsg, Dsg, Asg, Vsg, Lsg, Npl, Gpl, Apl, Vpl
vzdělání	vzdělání	pl	10	Nsg, Gsg, Dsg, Asg, Vsg, Lsg, Npl, Gpl, Apl, Vpl
vzestup	vzestup	sg	2	Nsg, Asg
vzestup	vzestupy	pl	4	Npl, Apl, Vpl, Ipl
vztah	vztah	sg	2	Nsg, Asg
vztah	vztahy	pl	4	Npl, Apl, Vpl, Ipl
záchvat	záchvat	sg	2	Nsg, Asg
záchvat	záchvaty	pl	4	Npl, Apl, Vpl, Ipl
zájem	zájem	sg	2	Nsg, Asg
zájem	zájmy	pl	4	Npl, Apl, Vpl, Ipl
zákrok	zákrok	sg	2	Nsg, Asg
zákrok	zákroky	pl	4	Npl, Apl, Vpl, Ipl
záměr	záměr	sg	2	Nsg, Asg
záměr	záměry	pl	4	Npl, Apl, Vpl, Ipl
zápach	zápach	sg	2	Nsg, Asg
zápach	zápachy	pl	4	Npl, Apl, Vpl, Ipl
zápas	zápas	sg	2	Nsg, Asg
zápas	zápasy	pl	4	Npl, Apl, Vpl, Ipl
zápis	zápis	sg	2	Nsg, Asg
zápis	zápisy	pl	4	Npl, Apl, Vpl, Ipl
zásada	zásadu	sg	1	Asg
zásada	zásady	pl	4	Gsg, Npl, Apl, Vpl
záštita	záštitu	sg	1	Asg
záštita	záštity	pl	4	Gsg, Npl, Apl, Vpl
závazek	závazek	sg	2	Nsg, Asg
závazek	závazky	pl	4	Npl, Apl, Vpl, Ipl

závěr	závěr	sg	2	Nsg, Asg
závěr	závěry	pl	4	Npl, Apl, Vpl, Ipl
záznam	záznam	sg	2	Nsg, Asg
záznam	záznamy	pl	4	Npl, Apl, Vpl, Ipl
zkouška	zkoušku	sg	1	Asg
zkouška	zkoušky	pl	4	Gsg, Npl, Apl, Vpl
zkušenost	zkušenost	sg	2	Nsg, Asg
zkušenost	zkušenosti	pl	7	Gsg, Dsg, Vsg, Lsg, Npl, Apl, Vpl
znechucení	znechucení	sg	10	Nsg, Gsg, Dsg, Asg, Vsg, Lsg, Npl, Gpl, Apl, Vpl
znechucení	znechucení	pl	10	Nsg, Gsg, Dsg, Asg, Vsg, Lsg, Npl, Gpl, Apl, Vpl
znepokojení	znepokojení	sg	10	Nsg, Gsg, Dsg, Asg, Vsg, Lsg, Npl, Gpl, Apl, Vpl
znepokojení	znepokojení	pl	10	Nsg, Gsg, Dsg, Asg, Vsg, Lsg, Npl, Gpl, Apl, Vpl
zodpovědnost	zodpovědnost	sg	2	Nsg, Asg
zodpovědnost	zodpovědnosti	pl	7	Gsg, Dsg, Vsg, Lsg, Npl, Apl, Vpl
ztráta	ztrátu	sg	1	Asg
ztráta	ztráty	pl	4	Gsg, Npl, Apl, Vpl
zvědavost	zvědavost	sg	2	Nsg, Asg
zvědavost	zvědavosti	pl	7	Gsg, Dsg, Vsg, Lsg, Npl, Apl, Vpl
zvuk	zvuk	sg	2	Nsg, Asg
zvuk	zvuky	pl	4	Npl, Apl, Vpl, Ipl
žádost	žádost	sg	2	Nsg, Asg
žádost	žádosti	pl	7	Gsg, Dsg, Vsg, Lsg, Npl, Apl, Vpl
žaloba	žalobu	sg	1	Asg
žaloba	žaloby	pl	4	Gsg, Npl, Apl, Vpl
život	život	sg	2	Nsg, Asg
život	životy	pl	4	Npl, Apl, Vpl, Ipl