

Matematicko-fyzikální fakulta
Univerzity Karlovy

**HOMONYMIE PŘEDLOŽKOVÝCH SKUPIN A
MOŽNOST JEJICH AUTOMATICKÉHO
ZPRACOVÁNÍ**

Markéta Straňáková

Disertační práce

Ústav formální a aplikované lingvistiky

Praha, 2001

Disertační práce byla vypracována v rámci doktorandského studia, které uchazeč absolvoval v Ústavu formální a aplikované lingvistiky Univerzity Karlovy v Praze v letech 1993-2001.

Uchazeč: Markéta Straňáková

Školitel: Prof. PhDr. J. Panevová, DrSc.
Ústav formální a aplikované lingvistiky, MFF UK

Školící pracoviště: Ústav formální a aplikované lingvistiky, MFF UK
Malostranské náměstí 25, 118 00 Praha 1

Oponenti: Prof. PhDr. P. Sgall, DrSc.
Ústav formální a aplikované lingvistiky, MFF UK

Doc. RNDr. V. Petkevič, CSc.
Ústav teoretické a počítačové lingvistiky, FF UK

Děkuji Jarmile Panevové za odborné vedení i lidskou podporu a povzbuzení. Děkuji za hodiny a hodiny konzultací, které podstatnou měrou přispěly k ujasnění koncepce této práce i k řešení dílčích problémů, i za její ochotu vždy vysvětlovat, zodpovídat dotazy a číst a komentovat postupně vznikající verze textu.

Martinu Plátkovi vděčím za mnohé impulsy, které mi umožnily nahlédnout problémy z jiné perspektivy, a děkuji za jeho snahu přimět mě k přesnému vyjadřování a jednoznačné formulaci lingvistických pozorování i z nich vyplývajících závěrů.

Eva Hajičová a Petr Sgall mi spolu s Jarmilou Panevovou podkryli složitost i krásu práce s jazykem, a poskytli mi k ní teoretické základy.

Tomáš Holan a Vladislav Kuboň dali ochotně k dispozici výsledky své práce, díky čemuž jsem mohla testovat své obecné závěry v konkrétním prostředí automatické syntaktické analýzy.

Díky za to, že tato práce mohla vzniknout, patří i všem ostatním kolegům, rodině a přátelům.

V neposlední řadě děkuji také Dominikovi a Benediktovi, kteří se museli smířit s tím, že máma doma pracuje a ‚zase má v počítači nějaký stromy‘.

OBSAH

Úvod	1
Část I. Motivace	3
I.1. Motivační příklady	3
I.2. ‚PP-attachment‘ v angličtině	9
Část II. Teoretický rámec lingvistického popisu strukturálních významů	13
II.1. FGD a definice homonymie	15
II.2. Klasifikace homonymních konstrukcí	19
II.2.1. Syntaktická homonymie	19
II.2.2. Morfemická homonymie	20
Homonymie volných doplnění	21
Homonymie vyjádření valenčního a volného doplnění	23
II.2.3. Fonologická homonymie	25
II.2.4. Fonologická homonymie v lexiku	28
II.2.5. Eliminace homonymie	28
Fonologická homonymie a její eliminace na vyšších rovinách	29
Fonologická homonymie v lexiku a její eliminace na vyšších rovinách ...	30
II.3. Homonymie předložkových skupin	31
Shrnutí	33
Část III. Automatické vyhledávání syntaktických struktur ‚podezřelých‘ vzhledem k homonymii předložkových skupin	35
III.1. Definování ‚podezřelých‘ struktur	37
III.1.1. ‚Podezřelé‘ struktury v příkladech	38
‚Cesta ke kořeni stromu‘ – typy struktur (a) a (b)	38
‚Testování sesterských uzlů‘ – typ struktury (c)	41
III.2. Návrh maker	45
Makro typu A	45
Makro typu B	47
Shrnutí	49
Část IV. Lingvistická kritéria	51
Předpoklady	52
Metoda redukční analýzy (RA)	53
IV.1. Slovosled	57
IV.1.1. Základní (prototypické), projektivní konstrukce s Pg	57
IV.1.2. Pg v netypických pozicích a neprojektivní konstrukce	63
Slovesně jmenné výrazy	65
Substantiva odvozená od sloves	67
Neprojektivní rozvíjení adjektiva	68
Kombinace jevů s uvolněnými slovoslednými podmínkami	69
Shrnutí poznámek k slovosledu	71

IV.2.	Valenční rámce	73
IV.2.1.	Valenční rámce sloves	74
	Rovina podkladové reprezentace	74
	Rovina povrchové syntaxe	76
	Kvazivalenční doplnění sloves	79
IV.2.2.	Valenční rámce substantiv	82
	Tektogramatická rovina	83
	Povrchová rovina	83
	Kvazivalenční doplnění substantiv	89
IV.2.3.	Slovesně jmenné výrazy	91
IV.2.4.	Valenční rámce adjektiv	95
	Kvazivalenční doplnění adjektiv	98
IV.2.5.	Vyhodnocování navržených závislostí	99
	Klasifikace jednotlivých typů valenčních požadavků	100
	Shrnutí poznámek k valenci	103
IV.3.	Sémantika	105
	Metoda pro návrh sémantických rysů substantiv	105
IV.3.1.	Návrh sémantických rysů pro substantiva	106
	Primární funkce Pg podle sémantických rysů substantiv	107
IV.3.2.	Návrh pravidel využívajících sémantických rysů substantiv	110
	Shrnutí	116
IV.4.	Slovosled a formální kritéria	119
IV.4.1.	„Princip separace“	119
IV.4.2.	Pravidlo o postavení příklonek ve větě	122
	Shrnutí	124
Část V.	Uplatnění a evaluace navržených kritérií	125
V.1.	Uspořádání kritérií v jednotlivých krocích analýzy založené na principu redukce a vyhodnocení získaných závislostí	127
	Řazení pravidel v jednom kroku analýzy	127
	Vyhodnocení získaných závislostí	128
V.2.	Uplatnění kritérií ve vybraných příkladech	131
	Shrnutí	140
Část VI.	Testování navržených kritérií	143
VI.1.	Gramatika a její interpret	143
VI.1.1.	Úpravy gramatiky	144
	Úpravy meta-pravidel gramatiky	145
	Vyloučení neadekvátních závislostí a evaluace získaných analýz	148
VI.2.	Příklady automatické analýzy	149
	Shrnutí	154
Závěr	155
Literatura	159
Příloha: Použité valenčních rámce		

ÚVOD

Tématem této práce je strukturální analýza syntakticky homonymních předložkových skupin (Pg, prepositional group). Budeme se tímto problémem zabývat z hlediska možnosti popisu homonymních konstrukcí pro účely automatického zpracování přirozeného jazyka, jmenovitě češtiny. Lze tedy říci, že se zabýváme problémem, který je ve složkově orientovaných popisech nazýván PP-attachment, tedy problémem hledání větných členů, které (potenciálně) daná Pg rozvíjí. Cílem práce je formulace kritérií pro identifikaci řídicího členu předložkové skupiny, resp. pro rozhodnutí o vytvoření a vyhodnocení několika syntaktických struktur v případě homonymních konstrukcí.

V první části uvádíme několik motivačních příkladů.

Druhá část je věnována specifikaci problému: základním paradigmatickým naším popisem jazyka (FGD, Functional Generative Description). Uvádíme klasifikaci homonymních konstrukcí založenou na určování, prvky které roviny homonymii do zkoumané věty vnášejí. Pro Pg je charakteristická především tzv. morfemická homonymie, tedy skutečnost, že konkrétní Pg může být v konkrétní větě považována za rozvíjení různých větných členů.

Kritéria pro automatickou disambiguaci (zjednoznačnění) budeme formulovat jednak na základě ručně nashromážděných vět, jednak na základě ‚podezřelých‘ vět vybraných z Pražského závislostního korpusu (PDT, Prague Dependency Treebank) a jejich modifikací. Ve třetí části popisujeme nástroje navržené pro automatické vyhledávání vět, jejichž povrchová struktura (lépe: struktura na analytické rovině) odpovídá některé z vytipovaných ‚podezřelých‘ struktur.

Čtvrtá, klíčová část je věnována návrhu kritérií vhodných pro automatické zpracování předložkových skupin při (lingvisticky podloženém) určování větných členů, které jimi mohou být rozvíjeny. Přiblížíme zde také metodu redukční analýzy (RA), která tvoří metodologický rámec našeho zkoumání.

V páté části předkládáme návrh uspořádání jednotlivých typů kritérií a možnost jejich využití při větném rozboru založeném na principech redukční analýzy. Velmi podstatná je možnost evaluace získaných závislostních struktur podle typu použitých kritérií.

V poslední, šesté části popisujeme experiment, ve kterém byla navržena kritéria pro určení (potenciálních) závislostí předložkových skupin použita v rámci konkrétní automatické procedury, za jejíž základ byla zvolena gramatika vyvíjeného Robustního parseru pro češtinu.

ČÁST I.

MOTIVACE

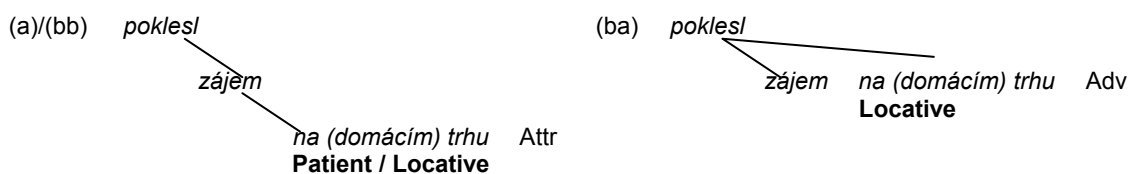
I.1. MOTIVAČNÍ PŘÍKLADY

Zabýváme-li se automatickou analýzou přirozeného jazyka, brzy narazíme na problém homonymie předložkových skupin¹ (homonymie Pg). Z čistě syntaktického hlediska může totiž Pg v konkrétní větě rozvíjet jakékoliv sloveso, podstatné jméno nebo adjektivum, případně i adverbium. Přitom ovšem významy (čtení) dané věty, které odpovídají různým analýzám (tj. analýzám lišícím se výběrem větného členu, který je rozvíjen zkoumanou Pg), mohou být různé (což znamená, že se mohou lišit jejich pravdivostní podmínky). Skutečnost, že v některých případech označují stejnou kognitivní situaci, je dána specifickým kontextem užití věty, který je založen na mimojazykové znalosti, a ekvivalenci tudíž nelze považovat za obecně platnou. Proto je potřeba zabývat se otázkou stanovení skutečného řídicího členu Pg, případně – jde-li o homonymní výskyt Pg – určení všech větných členů, které mohou být modifikovány danou Pg.

Podívejme se na následující větu, která obsahuje Pg sestávající z předložky *na* a (syntaktického) substantiva² v lokálu (dále budeme značit *na+Loc*):

Příklad:

(I.1) *V období, kdy prudce poklesl zájem na domácím trhu, dokázala továrna část výroby exportovat.* (Prague Dependency Treebank,³ bmd03zua.fs #4, zkráceno)



Obr.1: Možné analýzy Pg *na (domácím) trhu*⁴

Z hlediska analýzy Pg *na (domácím) trhu* je podstatný valenční rámec substantiva⁵ *zájem*, který sestává ze dvou doplnění, aktoru a pacientu (stejně jako rámec slovesa *zajímat se*, od kterého je

¹ Termíny jako homonymie a předložková skupina budou vysvětleny v kapitole II.1.

² Substantivem nebo (podstatným) jménem dále rozumíme syntaktické substantivum, tj. takovou třídu slov, které mají syntaktické vlastnosti substantiva (kromě substantiv jde hlavně o zájmena a ‚nominalizovaná‘ adjektiva).

³ Pražský závislostní korpus (PDT), základní zdroj podkladových dat, přiblížíme v kapitole III., kde jsou také odkazy na příslušnou literaturu.

⁴ Struktura vět je znázorňována závislostním stromem, jehož vlastnosti jsou popsány v oddíle II.1.

⁵ Valence je jedním z kritérií pro určení homonymie Pg – věnujeme jí kapitolu IV.2.

odvozeno). Patient má dvě možné povrchové realizace – předložku *na* a jméno v lokálu nebo předložku *o* a jméno v akuzativu. (Aktor lze vyjádřit podstatným jménem v genitivu nebo přivlastňovacím adjektivem, běžné je zaplnění jeho pozice všeobecným aktantem.)

zájem ... Act (Gen_{subj} / Adj_{pos}) Pat (*na*+Loc / *o*+Acc)⁶

Předložkovou skupinu *na (domácím) trhu* lze tedy analyzovat několika způsoby:

- (a) lze ji považovat za valenční doplnění substantiva *zájem* (Patient);
- (b) nebo ji lze považovat za volné doplnění místa (Locative)⁷ – a to
 - (ba) za doplnění slovesa *poklesnout*, nebo
 - (bb) za nevalenční atribut substantiva *zájem*.

Dvě čtení věty, kde je Pg analyzována jako volné doplnění, tedy (ba) a (bb), označují stejnou kognitivní situaci, v obou případech Pg *na (domácím) trhu* jednoznačně určuje místo, ke kterému se vztahuje výpověď věty. Valenční pozice substantiva *zájem* odpovídající patientu je obsazena všeobecným aktantem, který je specifikován v druhé části věty – jde o *zájem o výroby*.

Naproti tomu analýza (a) má jiný význam, který lze parafrázovat výběrem jiného povrchového vyjádření patientu (I.1'), příp. jiným lexikálním obsazením (I.1''):

(I.1') *V období, kdy prudce poklesl zájem o domácí trh, dokázala ...*

(I.1'') *V období, kdy prudce poklesl zájem na získání domácího trhu, dokázala ...*

Dvojice (a)-(ba) a (a)-(bb) mají tedy různý význam (různé pravdivostní podmínky). □

Na tomto místě je potřeba demonstrovat, co vlastně rozumíme **syntaktickou analýzou** přirozeného jazyka. Je zřejmé, že při porozumění větě člověk kromě znalosti gramatiky a slovníku daného jazyka využívá také znalostí sémantických a pragmatických (svého povědomí o světě i znalosti širokého kontextu promluvy) – jak o tom svědčí i předchozí příklad. Takto komplexní informaci o světě popisovaném zkoumanou větou (nechceme-li se omezit na velmi úzkou a přesně vymezenou oblast užívání jazyka) nelze plně zahrnout do žádného nástroje pro plně automatické zpracování přirozeného jazyka. Je proto třeba stanovit, které typy informací (s ohledem na účel konkrétního nástroje) lze s úspěchem využít a na které je nutno rezignovat.

Podívejme se na tento problém i z druhé strany – co to znamená **čistě syntaktická analýza**? Kuboň (1999) uvádí následující příklad:

Příklad:

(I.2) *Banky snižují úroky z ekonomických důvodů.* (Kuboň (1999))

(I.2') *Banky snižují úroky z krátkodobých půjček.*

Z čistě syntaktického hlediska (nebereme-li v úvahu svou znalost okolního světa) jsou tyto dvě věty strukturovány zcela stejně – skládají se z určitého tvaru slovesa *snižovat*, jeho dvou aktantů (aktoru *banka* a patientu *úrok* v patřičných morfologických formách) a ze členu, který je morfematically vyjádřen předložkovou skupinou *z+Gen*. Tato Pg může rozvíjet buď sloveso (věta

⁶ Za názvem aktantu (jeho zkratkou – Act, Pat, Orig, Addr, Eff) uvádíme v kulatých závorkách jeho možná povrchová vyjádření (pád substantiva, Pg, vedlejší věta), oddělená lomítkem /. Fakultativní valenční doplnění označujeme závorkou. U sloves oddělujeme tzv. levointenční aktant svislíci |.

⁷ Přidržíme se názvů funktořů, které jsou používány v PDT.

(I.2)), nebo předcházející substantivum (věta (I.2')). Tyto dvě možnosti nelze rozlišit, aniž bychom brali v úvahu sémantiku. □

Takováto analýza založená na syntaktické strukturaci věty v mnoha případech úplně dostačuje (jmenujme například rozpoznávání správnosti, tj. analýzu pro grammar-checker); navíc je opodstatněná i značným urychlením výpočtu a nezanedbatelným zmenšením prostorové náročnosti, což je velmi podstatné pro jakékoli praktické využití. Přesto nás z jazykového hlediska zcela neuspokojuje. Proto se snažíme najít další (už ne v přísném slova smyslu syntaktická) kritéria, která by se dala využít pro přesnější určení slova/větného členu, který je rozvíjen předložkovou skupinou (v závislostní terminologii: pro určení větného členu, na kterém daná Pg závisí).

Je nepochybné, že se – chceme-li automaticky zpracovávat přirozený jazyk – neobejdeme bez valenčních rámců sloves. Podívejme se na možné analýzy následující věty.

Příklad:

(I.3) *Malým znamenkem plus ohodnotil minulý rok prezident Jelcin ve včerejším každoročním projevu.* (MF Dnes, 1996)

Při analýze věty (I.3) podstatným způsobem pomůže **valenční rámec slovesa ohodnotit**:

ohodnotit ... Act (Nom) | Pat (Acc)

Větný člen *rok* bude primárně (na základě znalosti valenčního rámce) na povrchové rovině analyzován jako objekt odpovídající hloubkovému patientu.⁸ Druhá možná analýza, která slovo *rok* rozebírá jako příslovečné určení času (Kdy?), zachycuje situaci, kdy ve větě chybí člen odpovídající (v zásadě nevypustitelnému) patientu. Takový člen ovšem může být v určitém užití věty doplněn z předchozího kontextu, proto nelze ani tuto analýzu zcela vyloučit – náš výzkum směřuje k tomu, aby taková analýza byla alespoň upozaděna. □

Zde je potřeba zdůraznit, že i parsery navržené pro zpracování českých vět, které se deklarují jako čistě syntaktické, pracují se zjednodušenými valenčními rámci sloves. Pro tuto práci je podstatný především přístup V. Kuboně, autora gramatiky tzv. robustního parseru, viz Kuboň (1999). **Robustní parser pro češtinu**⁹ byl vyvíjen pro potřeby grammar-checkeru, tj. automatického gramatického korektoru. Bylo proto potřeba, aby byl schopen analyzovat nejen správně utvořené (well-formed) české věty, ale také věty obsahující syntaktickou chybu, lépe řečeno věty se syntaktickou nekonzistencí (ill-formed). Proto se autor rozhodl uvažovat omezené valenční rámce sloves obsahující pouze valenční doplnění v prostých, bezpředložkových pádech. Valenční doplnění realizovaná předložkovými skupinami se zpracovávají stejným způsobem jako volná doplnění. Analýza následující věty (syntakticky zcela jistě správně utvořené, i když poněkud stylisticky neobratné) ilustruje možnosti zpracování Pg pomocí robustního parseru.

Příklad:

(I.4) *Adresa Úřadu pro řešení důsledků pobytu sovětských vojsk na území ČSFR je Slezská 9, 120 29 Praha 2.* (PDT, bl104js.fs #2)

Z čistě syntaktického pohledu je předložková skupina *na území (ČSFR)* homonymní (jde o morfematickou homonymii, jak ji definuje Panevová (1981), zde oddíl II.1.) – Pg může být analyzována jako nominální doplnění rozvíjející substantiva *adresa, úřad, řešení, důsledky, pobyt* a *vojska* a také jako volné doplnění slovesa *být*. Robustní parser tedy vytvoří sedm

⁸ O rozdílech mezi hloubkovou a povrchovou strukturou věty viz kapitola II.1., vztahu valenčního rámce a jeho realizace na povrchové rovině se dotýkají kapitoly IV.2.1. a IV.2.2.

⁹ Gramatika robustního parseru je přibližena v části VI.

syntaktických závislostí dané Pg. (K této větě se vrátíme ještě v oddíle III.2., kde budeme ilustrovat činnost vyhledávacích maker (označena (III.10a)).) □

Je vidět, že rozhodnutí analyzovat Pg jako volné doplnění vede k vytváření velkého množství homonymních konstrukcí. Kuboň obchází tento nechtěný nárůst počtu výsledných syntaktických struktur technickým rozhodnutím: předložkové skupiny jsou analyzovány jednak jako větné členy modifikující nejbližší předcházející jméno nebo sloveso, jednak jako větné členy modifikující nejbližší následující sloveso. To znamená, že zkoumaná Pg se ‚zavěšuje‘ jednak na ten uzel závislostního grafu, který reprezentuje nejbližší jméno nebo sloveso předcházející v povrchové realizaci věty tuto Pg (pokud takový uzel existuje), jednak na uzel reprezentující nejbližší následující sloveso (pokud existuje). Přitom ovšem předpokládá, že (ze syntaktického hlediska) Pg může záviset i na všech dalších slovesných a jmenných uzlech ležících na cestě mezi hlavním slovesem a uzlem reprezentujícím vybrané řídicí slovo. Tyto analýzy lze kdykoli dogenerovat. (Výsledkem tohoto rozhodnutí jsou tedy dvě – alespoň co se týká analýzy Pg – syntaktické struktury věty (I.4), se závislostmi *vojsk na území (ČSFR)* a *na území (ČSFR) je.*) Takovéto řešení ostatně koresponduje s možnými technickými řešeními, kterých se v obdobných případech používá při syntaktické analýze anglických vět – vrátíme se k nim v následujícím oddíle.

Robustní parser můžeme pokládat za příklad syntaktického přístupu k automatické analýze češtiny. Kromě parserů založených na statistických a pravděpodobnostních metodách (jako např. Ribarov (1996), Zeman (1997)), které nejsou pro naši práci relevantní, byly vyvinuty pilotní implementace několika dalších parserů pro češtinu využívajících pravidel člověkem psaných gramatik. Jde například o parser navržený v rámci projektu automatického překladu RUSLAN, viz Oliva (1989), o systém popsany Z. Kirschnerem (1994) nebo grammar-checkery navržené pro projekt Lateslav (Language Technologies for Slavic Languages), viz Plátek et al (1996) a Oliva (1996). Systémy navržené Olivou předpokládají využití souboru **sémantických rysů**, jejich lingvistický popis však nebyl publikován. Jedním z našich cílů je diskuse o možnostech, které s sebou – z hlediska automatického odhalování a zpracování homonymie předložkových skupin – nese využívání sémantických rysů podstatných jmen, předložek a sloves v lexikonu.

Základní strategií našeho přístupu ke zkoumání možných závislostí předložkových skupin, k určení a vyhodnocení jejich homonymie, případně k jejich disambiguaci, je metoda postupného zjemňování pravidel, kterou lze shrnout do tří bodů. 1. Nejprve zkoumáme předložkové skupiny v prototypických, projektivních strukturách, omezujeme se přitom na jednoduchá slovosledná pravidla a pravidla využívající valenčních informací. 2. V druhém kroku tuto základní analýzu zjemňujeme, připouštíme neprojektivní závislosti, využíváme bohatší slovníkové informace (např. údajů o víceslovných výrazech, případně též sémantických údajů), podrobnější pravidla. 3. Poslední fáze je charakterizována komplexními pravidly, na jejichž základě jsou ‚prořezávány‘ získané struktury, jsou vyloučeny nepřípustné tvary závislostních stromů. Výsledné závislosti předložkových struktur jsou vyhodnoceny podle typu pravidla, které vedlo k jejich vytvoření..

Zmíňme zde základní charakteristiku metody, se kterou přistupujeme ke zkoumaným větám. Jde o metodu **redukční analýzy** (RA).¹⁰ V principu se RA shoduje s přístupem, který

¹⁰ Principy RA budou popsány v části IV.

používají při rozboru českých vět žáci základní školy, když při určování tzv. větných dvojic postupně zjednodušují rozebíranou větu. Na použitelnost tohoto přístupu jako podkladu pro automatickou analýzu přirozeného jazyka poukazuje Novotný (1988) (pojmu RA ve své práci ovšem nepoužívá), dále jej rozvíjí Plátek (viz např. Jančar et al (1999)). Výběr RA, která spočívá v postupném zjednodušování věty (vypouštění větných členů) tak, aby její správnost zůstala zachována (upřesnění viz dále), má zásadní vliv na formulaci kritérií pro určování homonymie Pg. Předpokládáme totiž, že se všechna pravidla, definovaná na základě lingvistických kritérií, budou aplikovat opakovaně, v každém kroku analýzy. Zdá se, že to odpovídá i intuitivní představě o postupu při rozboru komplikované věty: opakovaně užíváme své znalosti syntaxe k postupnému zjednodušování analyzované věty.

K popisu syntaxe češtiny se tradičně používá **závislostní formalismus** založený na klasickém zpracování syntaxe (viz např. Šmilauer (1966)), jehož základním konceptem je relace závislosti mezi větnými členy (členem řídícím a členem závislým). Za závislý člen se považuje ten, který lze vynechat, aniž věta ztratí svou syntaktickou správnost (přesněji: toto kritérium platí, je-li zohledněno pro slovní druhy, viz např. Sgall et al (1986a), str.120). Kromě ekonomičnosti a přehlednosti popisu (všechny důležité informace o větném členu jsou obsaženy v komplexním ohodnocení jednoho uzlu) je podstatný další rys závislostního formalismu, a to možnost jednoduchého zachycení vět s neprototypickým slovosledem, včetně neprojektivních konstrukcí. Tato výhoda se uplatňuje především pro tzv. jazyky s volným slovosledem, a není tedy zřejmě náhoda, že závislostní přístupy jsou charakteristické i pro jiné takové jazyky (zejména pro ruštinu, viz Mel'čuk (1988), Dikovský a Modina (2000)). Tento přístup je také východiskem pro metodu redukční analýzy.

Analýzou syntaktické platnosti předložkových skupin (i bezpředložkových nepřímých pádů a příslovcí) v české větě se zabývala už v 70. letech J. Panevová, která (spolu s E. Buráňovou, A. Bémovou a dalšími) vypracovala dílčí algoritmickou proceduru pro rozlišení předmětu a příslovečných určení několika druhů. Z algoritmů tehdy byla publikována jen malá část (viz Panevová a Sgall (1979), (1980), Bémová (1979)). Dnes, kdy je k dispozici rozsáhlý soubor textů z Českého národního korpusu včetně syntaktických anotací (na tzv. analytické rovině) v Pražském závislostním korpusu, jsou ovšem výchozí podmínky pro takový výzkum daleko lepší.

Než se budeme zabývat možnostmi zachycení a odhalení homonymie předložkových skupin v češtině, všimněme si přístupů k Pg (resp. k předložkovým frázím), které se používají při automatické syntaktické analýze anglických vět.

I.2. ‚PP-ATTACHMENT‘ V ANGLIČTINĚ

Při automatickém zpracování anglických vět je potřeba řešit stejný problém připojování předložkových frází, tzv. PP-attachment, jako v češtině. Anglické parsery (jmenujme například parser PEG, Heidorn (1975), který využívá principu ‚last attachment‘) běžně pracují se dvěma typy technických řešení, pomocí kterých se vyhýbají nutnosti využívat specifických syntaktických a sémantických vlastností jednotlivých lexikálních jednotek. Jde o tzv. ‚minimal attachment principle‘ a ‚last attachment principle‘.

‚Minimal attachment principle‘, tedy princip minimálního připojení, byl navržen pro gramatiky využívající bezprostředně složkový formalismus, kde se jednotlivé syntaktické struktury popisující zkoumanou větu mohou lišit počtem uzlů. Kritériem pro výběr řídicího uzlu, tj. ‚head‘ (z několika možných, vybraných na základě ‚globálně‘ platných pravidel) je celkový počet uzlů ve výsledné struktuře – Pg je připojována tak, aby se výsledná syntaktická struktura skládala z co nejmenšího počtu uzlů. Zdá se, že tento princip lze parafrázovat tak, aby byl využitelný i v závislostně orientovaných gramatikách – za jeho dobu je možno považovat pravidlo preference struktury s menším počtem průchodů ‚vyššími‘ uzly.

Druhým používaným principem je ‚last attachment principle‘, podle kterého je zkoumaná Pg připojována k nejbližšímu podstatnému jménu nebo slovesu, které této Pg předchází v povrchové realizaci věty.

Jak už jsme poznamenali dříve, jsou taková řešení pouze východisky z nouze, která zcela neodpovídají lingvistické intuici, a to ani pro angličtinu (viz příklady (I.6) a (I.6')), ani pro češtinu. Slouží k disambiguaci Pg, neumožňují přiměřeně zachytit skutečnou homonymii předložkových skupin. Srovnajme následující věty obsahující substantivum rozvíjené (jedním nebo několika) neshodným atributem v genitivu, po kterém následuje Pg.

Příklady:

(I.5a) *Nedávné kontakty prostějovského výrobce kabel, batohů, míčů a celé škály sportovních potřeb na všech úřadech práce ČR a SR byly zbytečné.* (PDT, bcd11eba.fs #24)

(I.5b) *V letech 1991-1993 jsem absolvovala kurzy řízení a marketingu prestižní Collège Bart v kanadském Québecu.* (PDT, bce17zua.fs #8, pozmeněno)

Ve větě (I.5a) je substantivum *kontakty* rozvíjeno jmennou skupinou skládající se z podstatného jména v genitivu (*prostějovského výrobce*, které je samo rozvíjeno koordinační skupinou [*kabel, batohů, míčů a celé škály sportovních potřeb*] (opět jména v genitivu). Pg *na (všech) úřadech (práce ČR a SR)* je vhodné (vzhledem k obsahu věty) analyzovat jako neshodný přívlástek jména *kontakt* (ačkoli je v povrchové realizaci věty ‚daleko‘, neodpovídá tedy principu ‚last attachment‘).

Na druhou stranu, ve větě (I.5b) lze jistě Pg v (*kanadském Québecu*) chápat jako atribut (v povrchovém zápisu) předcházejícího jména *Collège (Bart)*. Lze ji ovšem považovat i za člen rozvíjející substantivum *kurzy* nebo sloveso *absolvovat* (s jinými pravdivostními podmínkami; poslední možnost odpovídá analýze podle principu ‚minimal attachment‘).

K oběma větám se ještě vrátíme v kapitole IV.1.1., při určování slovosledných vzorců vzhledem k homonymii Pg (viz věty (IV.4a) a (IV.4b)). □

Problémem disambiguace předložkových skupin ve zpracování přirozeného jazyka se zabývá např. Allen (1987). Zmiňuje zde sémantický přístup založený na klasifikaci sloves. Pro každé sloveso je definována množina předložek, které mohou tvořit předložkovou skupinu rozvíjející toto sloveso; vyskytne-li se v odpovídající pozici předložková skupina uvozená jinou předložkou, je považována za substantivní doplnění. Allen uvádí následující příklad.

Příklad:

(I.6) *I gave the book to Sue.* (Allen (1987))

(I.6') *I gave the book from Sue.* (Allen (1987))

Ve větě (I.6) Pg *to Sue* rozvíjí sloveso *to give*, protože vyjadřuje jeden z participantů události *dávání*. Naopak v (I.6') musí být Pg *from Sue* analyzována jako neshodný atribut substantiva *book* (nemá smysl ji považovat za člen modifikující událost *dávání*). Jinými slovy, předložka *to* patří do množiny předložek, které mohou uvozovat doplnění slovesa *to give*, zatímco předložka *from* podle Allena do této množiny nepatří. □

Myšlenka stanovení takových množin předložek (pro jednotlivé participanty zkoumaných sloves) koresponduje s obecnějším konceptem zachycení valenčních rámců sloves a určením jejich možné povrchové formy, s nímž pracuje Funkční generativní popis češtiny.¹¹

Většina z novějších přístupů, které se zabývají předložkovými skupinami, je založena na statistických metodách a metodách využívajících (různě bohatě anotované) korpusy. Tyto přístupy si kladou za cíl disambiguaci Pg, ale nezkoumají v úplnosti jejich případnou homonymii. Obvykle přitom řeší zjednodušenou úlohu – rozhodují pouze mezi závislostí Pg na bezprostředně předcházející jmenné frázi a na slovesné frázi (obvykle se takový problém popisuje čtveřicí (V, N_1, P, N_2), kde V je ‚head verb‘ slovesné fráze, N_1 je ‚head noun‘ jmenné fráze, P předložka a N_2 ‚head noun‘ předložkového doplnění). Úspěšnost nejlepších metod se blíží 90 procentům (viz Stetina a Nagao (1997)).

Následující poznámky ilustrují komplexnost nástrojů, které mohou tyto přístupy k disambiguaci Pg využívat.

Zajímavý přístup k disambiguaci předložkových skupin v angličtině, založený na kumulaci heuristických pravidel, popisují Jensen a Binot (1987). Navrhují zde postup a nástroje pro využití informací obsažených v on-line výkladových slovnících (konkrétně Longman Dictionary of Contemporary English). Pro výchozí analýzu zkoumaných vět a pro analýzu slovníkových definic předpokládají použití už zmíněného syntaktického parseru (PEG). PEG připojuje Pg na principu ‚last attachment‘, s vyznačením dalších možných připojení. Na základě dostupných slovníkových informací (definice, příklady, synonyma, poznámky a komentáře) jsou definovány vzorce popisující ‚cílový uzel‘ pro Pg. Pro každou předložku je stanoven systém heuristických pravidel, která se jednak odvolávají na naplnění vzorců, jednak pracují s hierarchickými řetězci danými slovníkovými definicemi. Tato pravidla vedou k vyhodnocení možných připojení Pg.

Z novějších prací zabývajících se disambiguací předložkových frází lze dále jmenovat přístup založený na výpočtu ‚souvýskytových‘ frekvencí (lexical associations) sloves a substantiv s předložkami (Hindle a Rooth (1993)). Trénovací data jsou získána extrakcí všech frází (V, N_1, P, N_2) z velkého (automaticky) syntakticky zpracovaného korpusu (použit FIDDITCH partial parser).

Ratnaparkhi (1998) přibližuje statistické modely pracující s heuristickými pravidly pro disambiguaci. Tyto modely využívají korpus označovaný POS-taggerem, v němž jsou

¹¹ Základní myšlenky FGD jsou popsány v kapitole II.1.

detekovány možná řídicí slova (head words) jmenných frází pomocí tzv. chunkeru. Poté je pomocí heuristických pravidel určeno head word předložkové fráze. Podstatná je skutečnost, že trénovací data jsou získána z textu, který není syntakticky analyzován, přičemž je zcela ignorována možnost homonymního výskytu předložkových frází (homonymie testovacích dat je řešena na základě jednoznačných trénovacích dat).

V příspěvku Pantela a Dekang Lee (2000) je popisován přístup využívající iterativní metody pro získání trénovacích dat z automaticky zpracovaného korpusu. PP je disambiguována metodou lineární kombinace rysů, je použita aproximace pomocí ‚kontextuálně podobných slov‘ (contextually similar words).

Náš přístup se svou metodou zásadně odlišuje od zmíněných statistických a korpusových metod. Je založen na zkoumání lingvistických pravidel, chceme formulovat lingvisticky podložená pravidla pro detekci adekvátních větných členů, které mohou být rozvíjeny předložkovou skupinou (na nichž Pg závisí). Množina ‚podezřelých‘ sloves, substantiv, ale také adjektiv, je širší než předpokládají výše zmíněné přístupy. Ve flexivním jazyce s tzv. volným slovosledem je navíc analýza Pg komplikována různými zdroji homonymie (k morfematické homonymii přistupuje i víceznačnost slovních tvarů). Využití lingvistických pravidel umožňuje vyhodnocení stanovených závislostí, určení jejich ‚spolehlivosti‘.

ČÁST II.

TEORETICKÝ RÁMEC LINGVISTICKÉHO POPISU STRUKTURÁLNÍCH VÍCEZNAČNOSTÍ

Homonymie náleží spolu se synonymií k podstatným rysům jazyka. Pojem homonymie lze vztáhnout k různým úrovním popisu jazyka, k různě komplexním jednotkám. Mluví se např. o homonymii jednotlivých morfémů, o homonymii slovních tvarů a homonymii slov, o homonymii vět.

Homonymii slov vymezil např. B. Trnka pomocí 'úhrnného významu slov' – homonyma jsou v jeho pojetí slova, která mají ve všech svých tvarech společnou slovní realizaci, přitom mají různý úhrnný význam (viz Trnka (1990)). Pro českou deklinaci ji v pojetí FGD důkladně zpracoval Sgall (1967). Z kvantitativního hlediska studovala homonymii slovních forem Těšitelová (1966), která rozlišuje homonymii slovních forem (pro splývání některých slovních forem jediného lexému používá termín morfologická homonymie), na rozdíl od lexikální homonymie, kde dochází ke splývání všech tvarů různých lexémů.¹² Syntaktickou homonymií se zabývá v rámci odvozování vět z gramatických větných vzorců Daneš (1964) – klasifikuje ji podle typu operace, při které homonymie vzniká. Panevová (1981) navrhla klasifikaci homonymie podle toho, které prostředky (na které jazykové rovině) ji do vět vnášejí.

V této části nejprve uvedeme definici homonymie a její klasifikaci (oddíl II.1.), v následujícím oddíle se na základě konkrétních vět soustředíme na podrobnější rozbor a explikaci jednotlivých typů homonymie. Posléze do tohoto systému zařadíme homonymii předložkových skupin.

¹² Těšitelová nezavádí termín pro takové případy homonymie, kdy splývají jen některé tvary různých lexémů (typ *jedu, ženu*).

II.1. FGD A DEFINICE HOMONYMIE

Teoretický základ našeho výzkumu je dán **Funkčním generativním popisem** češtiny (FGD, Functional Generative Description), který popisuje např. Sgall (1967) a Sgall et al (1986a). Jde o závislostní typ formalismu, který – ačkoli byl navržen pro účely teoretického popisu struktury českých vět pomocí generativní procedury, jak ostatně napovídá už jeho název – je ‚direction independent‘; to znamená, že může sloužit stejně dobře jako podklad pro analytické procedury i pro procedury syntetické, generativní.

Základní charakteristikou FGD je stratifikační přístup k popisu jazyka – popis jazyka je rozčleněn do několika rovin. Každá z rovin je množinou zápisů vět, každá má svou syntax. Nejvyšší rovina odpovídá významovému plánu jazyka, nejnižší plánu grafematickému či fonetickému. Jednotky jednotlivých sousedních rovin jsou ve vzájemném vztahu formy a funkce (jednotka vyšší roviny je funkcí jednotky nižší roviny, jednotka nižší roviny je její formou).

V ‚klasické‘ verzi (viz např. Sgall (1967), Panevová (1980)) FGD pracuje s pěti rovinami: rovinou podkladové reprezentace (tzv. tektogramatická rovina, TGR) (v anglické terminologii: level of underlying representation nebo tectogrammatical level), rovinou povrchové syntaxe (RPS) (level of surface structure), rovinou morfematickou (morphemic level) a rovinami morfonologickou (morphophonemic level) a fonetickou (phonetic level) (nebo grafematickou).

V novějším pojetí FGD, viz např. Sgall (1992), je počet rovin zredukován. Je zpochybněna teoretická opodstatněnost roviny povrchové syntaxe, rozlišují se pouze tři roviny: podkladová, morfematická a fonologická. Nicméně pro automatické a poloautomatické systémy (jako je automatická syntaktická analýza nebo syntaktická anotace korpusu, viz kapitola III.) je vhodné využívat povrchové roviny nebo nějaké její obdoby alespoň jako technického nástroje. Pro PDT je pro tyto účely definována tzv. analytická rovina (viz Bémová et al (1997)), zde pracujeme s rovinou povrchové syntaxe, RPS, která je rovině analytické velmi blízka¹³.

n = 4	_____	rovina podkladové reprezentace (tektogramatická, TGR)
n = 3	_____	rovina povrchové syntaxe (RPS)
n = 2	_____	morfematická rovina
n = 1	_____	fonologická rovina

Obr.2: Roviny popisu jazyka (pro automatické systémy)

FGD využívá závislostního formalismu – na tektogramatické a na povrchové rovině je věta zachycena jako závislostní strom, což je spojitý acyklický orientovaný graf s kořenem, kde do každého uzlu kromě kořene vede právě jedna hrana, uzly jsou lineárně uspořádány. Hrany stromu reprezentují vztah závislosti. K popisu věty na nižších rovinách stačí lineární struktura, tedy řetězce.

¹³ Analytická rovina a RPS se liší zejména zpracováním pomocných slov (předložky, spojky, pomocná slovesa), která jsou na analytické rovině reprezentována vlastním uzlem, zatímco na RPS jim odpovídají pouze složky komplexních hodnot slov plnovýznamových. Na analytické rovině jsou zachyceny zvláštními uzly také interpunkční znaménka, navíc přistupuje i uzel reprezentující celou větu, kořen závislostního stromu.

Pro potřeby teoretického popisu je vhodné rozlišovat dva základní pojmy: homonymii a víceznačnost. Tyto pojmy souvisejí s asymetrií výrazového a významového plánu jazyka, kdy jedinému prvku výrazového plánu odpovídají dva prvky (nebo i více prvků) plánu významového (viz Panevová (1981)).

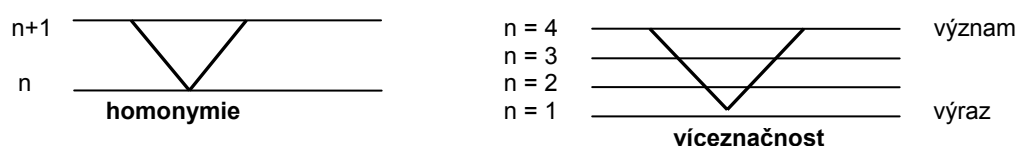
Označme roviny jazykového systému přirozenými čísly 1,...,n, přičemž nejnižší rovina (fonologická) má číslo 1.

Definice (homonymie na rovině n):

Homonymií na rovině n chápeme vztah, kdy alespoň dvěma výrazům (funkcím) na rovině (n+1) odpovídá jediný prostředek (forma) na rovině n. □

Homonymií (v užším smyslu) se tedy rozumí vztah mezi jednotkami sousedních rovin (tato homonymie prostředků, kdy jedna forma má dvě nebo více funkcí, může být ovšem kontextem eliminována); tento jev je významný pro teoretický popis.

Víceznačnost chápeme jako homonymii celých vět a konstrukcí. □



Obr.3: Homonymie a víceznačnost

Typy homonymie	Roviny popisu jazyka
n = 4 syntaktická homonymie	rovina podkladové reprezentace (tektogramatická, TGR)
n = 3 morfemická homonymie	rovina povrchové syntaxe (RPS)
n = 2 fonologická homonymie	morfemická rovina
n = 1	fonologická rovina

Obr.4: Typy homonymie

Podle Panevové tedy rozlišujeme tyto typy homonymie:

Syntaktická homonymie ... věta má různé hloubkové struktury (TGR), které jsou reprezentovány jediným povrchovým stromem (RPS).

Morfemická homonymie ... dvě funkce na RPS splývají na morfemické rovině, mají stejnou formu (tj. stejnou kombinaci sémat).

Fonologická homonymie ... liší se na morfemické rovině, splývá jejich fonologická (případně fonetická) reprezentace. Zvláště vyčleňujeme fonologickou homonymii v lexiku.

Mluvíme-li dále o homonymii, máme vždy na mysli homonymii na určité rovině jazykového popisu.

Řekněme zde ještě, co míníme termínem předložková skupina.

Vysvětlení pojmu (předložková skupina (Pg)):

Termínem **předložková skupina** zde míníme předložkový pád spolu se členy na něm přímo nebo nepřímo závislými, tedy ekvivalent toho, co je v bezprostředně složkových gramatikách označováno jako předložková fráze. (Protože tento pojem je úzce spjat s formalismem

bezprostředně složkovým, a navíc je v češtině termín fráze terminologicky obsazen, pokládáme za vhodné používat jiný termín.)

V nejjednodušším případě jde o pouhou předložku a substantivum, jehož pád je dán rekcí předložky.

V případě, kdy je Pg tvořena předložkou a jmennou skupinou, je obvykle rekcí předložky dáno povrchové vyjádření řídicího substantiva jmenné skupiny. (Je-li jmenná skupina tvořena jediným substantivem, dostáváme předchozí případ.) Na základě redukční analýzy lze za Pg považovat její **jádro** tvořené předložkou a řídicím substantivem jmenné skupiny (větné členy rozvíjející toto substantivum jsou redukovány před vyšetřováním možných závislostí Pg; k jádru přitom nepatří valenční členy substantiva). Toto jádro Pg označujeme též jako **redukovanou Pg**. (RA umožňuje pomocí přepisování zpracovat i případy, kdy pád řídicího substantiva v neredukované Pg není dán rekcí předložky, např. ve výrazech s číslovkami.) □

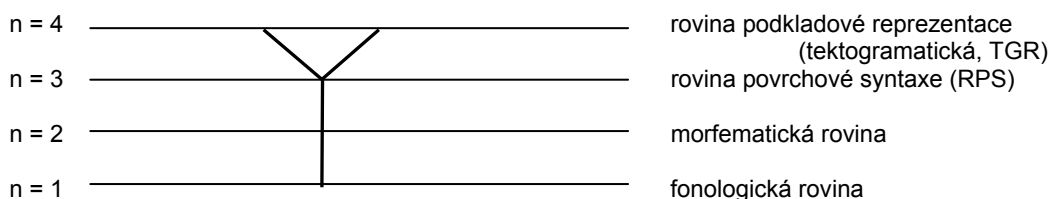
II.2. KLASIFIKACE HOMONYMNÍCH KONSTRUKCÍ

Zde se chceme na základě konkrétních vět soustředit na podrobnější rozbor a explifikaci jednotlivých typů homonymie.

Kromě známých a v lingvistické literatuře často uváděných příkladů čerpáme v tomto oddíle převážně z tisku (MF Dnes, LN, Respekt), dále z literárních časopisů (Revolver Revue, Kritická Příloha RR), několik příkladů pochází z české i překladové prózy.

II.2.1. Syntaktická homonymie

Pokud pracujeme se dvěma rovinami větné stavby (např. při anotování korpusu), je věta reprezentována jediným zápisem – závislostním stromem – na rovině povrchové syntaxe, kterému odpovídají dva zápisy na rovině tektogramatické. Jak vyplývá z následujících příkladů obsahujících syntaktickou homonymii, splynutí reprezentací na RPS je dáno počtem a výběrem rozlišovaných jednotek (v případě korpusu tedy spíše technickými parametry roviny odpovídající povrchové syntaxi), syntaktickou homonymii lze zachytit již jako víceznačnost morfémů.



Obr.5: Syntaktická homonymie

Rozeberme zde tyto možnosti na několika konkrétních větách.

Homonymie vedlejších vět

Příklady:

(II.1a) *Mluvili o tom, kdo u nich bude bydlet.* (Panevová (1981))

(II.1b) *Užovky, které mají jedové zuby uloženy vpředu v horní čelisti, zařazujeme do skupiny hadů jedovatých.* (Staněk – Velký obrazový atlas zvířat)

V případě věty (II.1a) jde o homonymii věty obsahové (s odkazovacím zájmenem spojujícím větu řídicí s větou závislou) a věty přívlastkové (s anaforickým zájmenem).

V souvětí (II.1b) nelze nerozlišit, zda se jedná o restriktivní či nerestriktivní (deskriptivní) vedlejší větu přívlastkovou. □

Homonymie daná nejednoznačným vztahem negace a plurálu, či nejasnou koreferencí

Příklady:

(II.2a) *Všichni tomu nevěřili.* (Daneš (1964))

(II.2b) *Zahlédli ji až při odchodu z divadla.* (Daneš (1964))

Věta (II.2a) má dvě různá čtení, která lze parafrázovat takto (jde zde ovšem o nejasný vztah mezi homonymií a obsahovou neurčitostí):

Ne všichni tomu věřili. / Každý z nich tomu nevěřil.

Věta (II.2b) lze číst třemi možnými způsoby, výběr adekvátního čtení závisí na daném kontextu:

Zahlédli ji při jejím odchodu / při jejich odchodu / když se odcházelo z divadla. □

Homonymie genitivu subjektového a genitivu objektového

Do této skupiny se řadí následující příklady:

Příklady:

(II.3a) *Spotřeba motorů stoupá.* (Sgall (1967))

(II.3b) *Kritika brazilského delegáta byla ostrá.* (Daneš (1964))

FGD ovšem rozlišuje na rovině povrchové syntaxe přívlastek subjektový a přívlastek objektový (viz Panevová (1980)). To znamená, že je tento typ homonymie posunut o rovinu níž – jde tedy o homonymii morfematickou.

Zajímavý posun nastává u substantiva kritik:

(II.3b') *Kritici Tváře Lopatka a Doležal ...* (KP RR)

Spojení *kritici Tváře* můžeme chápat stejně jako sousloví *členové skupiny* (tj. přináleží ke skupině, App), nebo objektově, jako *průvodci skupiny cizinců* (Pat). □

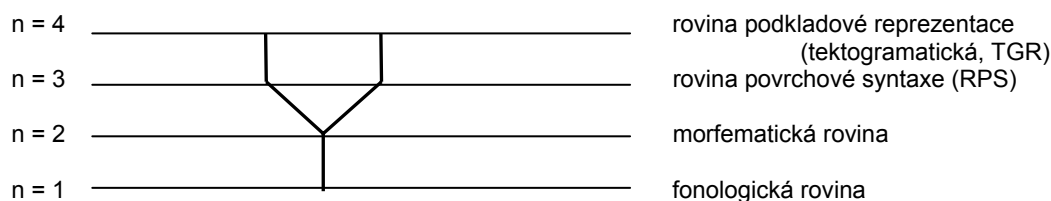
II.2.2. Morfematická homonymie

Tento typ homonymie je charakterizován situací, kdy věta má alespoň dva zápisy na RPS a jediný na rovině morfematické (závislý větný člen se může vztahovat k různým členům řídícím).¹⁴

Velmi podstatný je zde slovosled – změnou slovosledu se totiž většinou jedno čtení eliminuje – viz např. (II.4a), (II.4a') nebo (II.5b), (II.5b'), (II.5b'').

K tomuto typu homonymie patří věty, ve kterých existence dvojí struktury na rovině povrchové syntaxe v podstatě není sémanticky relevantní (II.4a), věty, které obvykle chápeme jednoznačně, ale může se vyskytnout kontext, ve kterém se uplatní jiné čtení ((II.6) a (II.6')), ale i věty, jejichž strukturu bez znalosti širšího kontextu (II.5b) nebo dokonce mimojazykové znalosti (II.8b) nelze vytvořit.

¹⁴ Tento typ homonymie se v anglicky psané literatuře často označuje jako syntaktická homonymie, zde se držíme české terminologie J. Panevové (1981).



Obr.6: Morfematická homonymie

Morfematickou homonymií můžeme rozdělit do několika skupin podle toho, kterých větných členů se týká. Zvláštní pozornost věnujeme větám, ve kterých je homonymní valenční a volné doplnění slovesa.

Homonymie volných doplnění

Homonymie přívlastku neshodného a příslovečného určení

Často se vyskytují větné členy, které lze analyzovat dvojím způsobem – buď jako doplnění substantiva, tedy (neshodný) atribut, nebo jako různé typy příslovečných určení. Ve větách (II.4a)-(II.4g) jde určení místa, v (II.5a)-(II.5b) o určení prostředku, či prospěchu (II.6).

Příklady:

(II.4a) *Žáci dělají chyby ve skloňování.* (Hausenblas (1958))

(II.4a') *Žáci dělají ve skloňování chyby.*

(II.4b) *Jirka maloval tužkou na obočí.* (Panevová (1986))

(II.4c) *(Zazdění oken.) Pokušení pozorovat ruch na ulici ... zmizelo s bočními okny.* (E. Canetti – Pochodeň v uchu)

(II.4d) *Vidím dívku z vlaku.* (Panevová (1986))

(II.4e) *... (úřad) hledá policisty na hranici.* (MF Dnes, 1996)

(II.4f) *Televizní síť CNN bude mít první pobočku v Havaně.* (MF Dnes, 1996)

(II.4g) *Opatrnost byla vůbec v týdnu před Solanovou cestou do Moskvy v sídle NATO pokynem číslo jedna.* (MF Dnes, 1996)

Z našeho hlediska je zajímavá zejména věta (II.4e), protože obsahuje homonymii morfematickou i fonologickou. Větný člen *na hranici* můžeme považovat jednak za přívlastek neshodný rozvíjející substantivum *policista*, jednak za příslovečné určení, a to buď účelu (Aim, v akuzativu), nebo místa (Locative, Kde? v lokálu) (viz též oddíl II.3.).

Podobně se ve větě (II.4g) člen *v sídle NATO* může vztahovat k predikátu *být*, může rozvíjet (jako přívlastek neshodný) substantivum *Moskva*, ale i větné členy *cesta* a *týden*. (Homonymie předložkové skupiny může být eliminována až s ohledem na mimojazykové znalosti.) □

Příklady:

(II.5a) *Při své poslední vyjíždě, kterou jsem s Dalamánkem na Sazi podnikl, přijeli jsme ...* (V. Burian – Veselé vyprávění dětem)

---> *Při vyjíždě, kterou jsem se psem na koni podnikl, ...*

(II.5b) *Chlapec honí pána na kole.* (Donald, 1997)

(II.5b') *Chlapec honí na kole pána.*

(II.5b'') *Pána na kole honí chlapec.*

(II.6) *Kupoval kalhoty pro kluka. Pro holku (protože jiné nechtěla).*

(II.6') *Kupoval kalhoty pro kluka.* (Hausenblas (1958))

Ke větě (II.5b) a jejím variantám (II.5b') a (II.5b'') se podrobněji vrátíme v kapitole IV.1.1. věnované povrchovému slovosledu a formulaci základních slovosledných vzorců vzhledem k homonymii Pg (viz příklad (IV.3)). □

Homonymie dvou různých druhů jediného typu příslovečného určení

Příklad:

(II.7) *Udělá to za dvě hodiny.* (Panevová (1980))

Pg za dvě hodiny je příslovečné určení času, které lze chápat jednak jako odpověď na otázku Kdy?, jednak jako odpověď na dotaz Za jak dlouho? □

Homonymie přiřazení přívlastku neshodného k několika členům řídicím

Věty (II.8a) a (II.8b) obsahují větný člen, který může být analyzován jako přívlastek, který potenciálně rozvíjí dvě různá podstatná jména.

Příklady:

(II.8a) *Kupoval ždímačku na prádlo nejnovějšího typu.* (Daneš (1964))

(II.8b) *Přinesl balík knih od přítele.* (Panevová (1986))

V (II.8a) přichází v úvahu závislosti ždímačka nejnovějšího typu a prádlo nejnovějšího typu, věta popisuje dvě různé situace.

Ve větě (II.8b) jde o knihy od přítele, balík od přítele a navíc o příslovečné určení přinést od přítele. □

Konkurence doplnění adjektiva a slovesa, resp. substantiva

Předložkové skupiny a příslovce mohou v některých slovosledných pozicích nabízet vždy více syntaktických struktur, kde jde o příslovečné určení, které rozvíjí buď sloveso, nebo adjektivum, příp. o konkurenci adjektivního a nominálního doplnění. Některé takové Pg a příslovce jsou homonymní, v ostatních případech vznikají tzv. ‚falešné větné dvojice‘.

(II.9a) *... je hotelem, ve kterém bydlí v Čechách (zatím téměř) neznámá skupina [squatterů].* (Panevová (1998a))

(II.9b) *V chůzi si ještě uhlazoval do fazonky rozcuchané, husté kudrny.* (Třešňák – Klíč je pod rohožkou)

(II.9c) *Přijela policie, evakuovala téměř tisíc přítomných lidí a bezpečně nastraženou bombu odpálila.* (MF Dnes, 1997)

(II.9d) *... rovná na ramínku vystavený kabát.* (Panevová (1998a))

(II.9e) *Ročního limitu přípustné dávky pro jednoho obyvatele dosáhne člověk až příjmem 100 000 BQ cesia 137 za rok.* (Panevová, 1998a)

Ve větě (II.9a) preferujeme analýzu Pg jako příslovečného určení rozvíjejícího adjektivum, v Čechách neznámá; Pg tvoří se slovesem ‚falešnou větnou dvojici‘.

Naopak ve větách (II.9b) a (II.9c) Pg *do fazonky*, resp. adverbium *bezpečně* rozvíjí slovesa.

Ke větě (II.9d), která je vícestupná (Pg *na ramínku* modifikuje jednak sloveso *rovnat*, jednak adjektivum *vystavený*), a jejím slovosledným variantám se vrátíme v souvislosti s povrchovým slovosledem v oddíle IV.1., příklad (IV.7).

Bližší pohled si zaslouží příklad (II.9e). Tuto větu lze chápat jako neprojektivní,¹⁵ člen *pro jednoho obyvatele* by potom (jako příslovečné určení způsobu – zřetele) rozvíjel adjektivum *přípustné* (genitiv). Zdá se ovšem, že srozumitelnější, „přirozenější“ by byla její modifikovaná, projektivní verze.

(II.9e') *Ročního limitu dávky přípustné pro jednoho obyvatele ...*

Pg *pro jednoho obyvatele* lze však analyzovat také jako přívlástek neshodný, který rozvíjí buď substantivum *dávka*, nebo substantivum *limit*. Z čistě syntaktického hlediska je nutno uvažovat i slovesné doplnění – *dosáhnout pro jednoho obyvatele*. □

Homonymie vyjádření valenčního a volného doplnění

Zásadní roli při stavbě závislostního stromu hrají valenční požadavky a jejich možná povrchová vyjádření.¹⁶ Mohou si přitom konkurovat valenční a volná doplnění různých větných členů (II.10) i různá doplnění jediného lexému (II.11a)-(II.11c).

Konkurence valenčního a volného doplnění různých větných členů

Příklad:

(II.10) *V souvislosti s přivtělením Rakouska k Německu je vyslýchán gestapem nemocný Trubeckoj a umírá v roce 1938.* (J. Černý – Dějiny lingvistiky)

Při rozboru věty (II.9e) pomůže valenční rámec slovesa *vyslýchat*:

vyslýchat ... Act | Pat (Acc)

Aktor má v pasivní větě (systémově) formu instrumentálu, substantivum *gestapo* (*gestapem_{ins}*) bude na základě valenčních požadavků určeno jako doplnění vyjadřující hloubkový aktor. Nabízí se také druhá zajímavá varianta *gestapem nemocný* – analogie spojení *tuberkulózou nemocný* (syntakticky stejně strukturováno, možno vyloučit až na základě obsahu). □

Konkurence valenčního a volného doplnění jediného větného členu

V následujících větách se setkáváme s možností analyzovat Pg jako valenční i volné doplnění slovesa (na povrchové rovině realizováno jako Pg s+Ins).

Příklady:

(II.11a) *Pestrá přihrbblá armáda. Jediná armáda, které pan P. nevyhlásil válku. Dokonce kdysi válčil s nimi.* (Třešňák – Klíč je pod rohožkou)

---> *Kdysi s nimi válčil.*

(II.11b) *Další armáda, které pan P. nevyhlásil válku, dokonce s nimi kdysi nějaký čas prohrával.* (Třešňák – Klíč je pod rohožkou)

¹⁵ Pojem neprojektivity přiblížíme v kapitole IV.1.2.

¹⁶ Valenci sloves, substantiv a adjektiv a jejichmu využití pro automatické zpracování Pg je věnována kapitola IV.2.

Z kontextu vyplývá, že správné je v tomto případě čtení *válčili*, resp. *prohrávali společně*, ovšem na základě jediné klauze rozhodnout nelze, věta je homonymní. Pro správnou analýzu těchto vět je podstatné zachycení valenčního rámce slovesa.

Uvažujeme tento valenční rámec slovesa:

válčit ... Act | Pat (s+Ins / proti+Dat)

Větný člen *s nimi* je homonymní, lze ho analyzovat dvojím způsobem. Jednak jako předmět, tedy realizaci hloubkového patientu (jako alternativu *válčit proti někomu*). Nebo ho lze považovat za příslovečné určení (adverbiale) vyjadřující volné doplnění – doprovod (s významem *válčit společně s někým*; v tomto případě zůstává patient na povrchové rovině nevyjádřen, valenční požadavek slovesa je naplněn tzv. všeobecným aktantem). Tyto dvě analýzy mají různé pravdivostní podmínky, a tedy různý význam.

Ještě složitější situace nastává u slovesa *prohrávat*. Uvažujeme valenční rámec

prohrávat ... Act | Pat (Acc) (Addr (s+Ins / proti+Dat))

Obligatorní patient je v povrchové struktuře věty vyjádřen všeobecným doplněním. Doplnění realizované předložkovou skupinou *s nimi* odpovídá jednak volnému doplnění – doprovodu, ve druhém čtení jde o vyjádření fakultativního adresátu (ve významu *proti nim*). □

Příklad:

(II.11c) *Teroristé se chtějí sejít s ručními granáty a revolvery v jednu hodinu u vchodu a čekat, až ministr vyjde* (L. Andrejev – O sedmi oběšených, přel. L. Dušková)

Pracujeme s následujícím valenčním rámcem slovesa *sejít se* (reflexivum tantum):¹⁷

sejít se ... Act^R | Pat^R (s+Ins)

Plurál substantiva *teroristé* ukazuje, že lze uvažovat o recipročním naplnění valenčního rámce slovesa – reflexivní zájmeno *se* může zaplnit pozici odpovídající hloubkovému patientu (viz Panevová (1999)). Potom lze předložkovou skupinu *s granáty* určit jako příslovečné určení odpovídající volnému doplnění slovesa – doprovodu. Povolíme-li neprojektivní konstrukce, je možné tuto Pg pokládat i za neshodný přívlastek substantiva *teroristé* (případně by bylo možné uvažovat o tom, že jde o (neshodný) doplněk).

Člen *s granáty* můžeme ovšem určit také jako předmět (na povrchové rovině) odpovídající hloubkovému patientu. □

Věta (I.1) (*zájem na domácím trhu*) ukazuje, že se možnost homonymie valenčního a volného doplnění neomezuje na slovesa.

Příklad:

(I.1) *V období, kdy prudce poklesl zájem na domácím trhu, dokázala továrna část výroby exportovat.* (PDT, bmd03zua.fs #4, zkráceno) □

Ostatní

Věta (II.12a) ilustruje homonymii predikátu jmenného se sponou a přívlastku shodného (spojenou s homonymií fonologickou):

¹⁷ Panevová (1999) navrhuje způsob, jak zachytit možnost recipročního i asymetrického užití slovesa v jediném rámci – aktanty, které mohou vstupovat do vztahu reciprocity, jsou označeny horním indexem R.

Příklady:

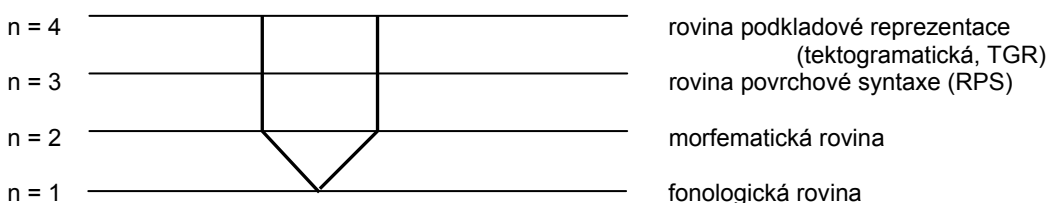
(II.12a) ... *bude těžké zkoušky provádět při nízké teplotě.* (Panevová a Sgall (1980)) □

(II.12b) *Výtahy se neslyšně vznáší mezi koberci vystlanými podlažími.* (MF Dnes, 1996)

Věta (II.12b) připomíná poměrně častý typ fonologické homonymie (viz kapitola II.2.3.), zde ovšem nejde o náhodnou shodu na fonologické rovině, ale o stejnou kombinaci sémat, tedy o splnutí na morfematické rovině (*koberci*_{ins,pl}, *podlažími*_{ins,pl}). □

II.2.3. Fonologická homonymie

Homonymie tohoto typu vzniká shodou reprezentace lexikálních jednotek na fonologické rovině (tj. skutečností, že jeden morf má několik funkcí). To ovšem znamená, že při jiném lexikálním obsazení nejednoznačnost zaniká.



Obr.7: Fonologická homonymie

„Klasické“ příklady

Panevová uvádí jako příklady fonologické homonymie (v naší terminologii) následující věty (II.15a)-(II.15c), větu (II.4e) jsme zmiňovali už v předcházející kapitole v souvislosti s morfematickou homonymií.

Příklady:

(II.15a) *Chlapci přinesou velkou rybu.* (Daneš (1964))

(II.15b) *Tu knihu prý chválí.* (Daneš (1964))

(II.15c) *Loni jsem chytil tlouště na višni.* (Panevová (1981))

(II.4e) ... *(úřad) hledá policisty na hranici.* (MF Dnes, 1996) □

Homonymie pádových tvarů předmětu a přívlastku shodného, resp. neshodného

„Falešné větné dvojice“ mohou vznikat i jako alternativní analýza větného členu, který splňuje valenční požadavky slovesa; zajímavým příkladem je konkurence valenčních požadavků slovesa a substantiva. Splnutím forem různých morfologických kategorií (tj. jsou-li vyjádřeny stejnými morfy) může dojít například k homonymii předmětu a přívlastku shodného, resp. neshodného.

Příklady:

(II.16a) *Pohladil Františkovi/y vlasy.* (Panevová (1986a))

(II.16b) *Benediktovi psi utekli do lesa.*

(II.16c) *Stejně jako podmínky pro práci Karla Flose příjemně překvapili i lidé, s nimiž bude spolupracovat.* (MF Dnes, 1997)

(II.16d) *Ve vsi se proslýchá, že mladý muž měl nastoupit do zaměstnání, ale dědeček budoucího zaměstnavatele na jeho problémy upozornil, takže mu místo nevyšlo.* (MF Dnes, 1995)

Ve (II.16a) jsou dvě možné alternativy graficky rozlišeny, v dalších větách už k tomuto rozlišení nedochází. Substantivum *František* (v dativu) může odpovídat volnému doplnění prospěchu na podkladové rovině (proto nám při adjektivním čtení tento člen v rámci slovesa *pohládit* nechybí). Obdobně u věty (II.16b).

Ve větě (II.16c) substantivum *Flos* buď odpovídá adresátu (rámec slovesa *překvapovat* viz (II.17i)), nebo rozvíjí (jako přívlastek neshodný) substantivum *práce* (potom je adresát všeobecný).

Ve větě (II.16d) uvažujeme valenční rámec slovesa *upozornit* a substantiva *dědeček*:

upozornit ... Act | Pat (na+Acc) (Addr (Acc))

dědeček ... App (Gen/Adj_{pos})

Větný člen (*budoucího*) *zaměstnavatele* lze analyzovat jednak jako neshodný atribut v genitivu rozvíjející substantivum *dědeček* (jde o vyjádření sémanticky obligatorního doplnění relačního jména – příslušnosti (Appurtenance), viz Piřha (1981), které je ovšem vypustitelné v povrchové realizaci věty), jednak se může jednat o doplnění slovesa, a to doplnění odpovídající (fakultativnímu) adresátu (akuzativ).

Je zřejmé, že je třeba upřednostnit valenční požadavky slovesa, nicméně jistě existují kontexty, ve kterých je adekvátní adnominální analýza Pg.¹⁸ □

Kombinace dvou lexikálních jednotek s homonymními morfologickými informacemi

Situaci, kdy kvůli kombinaci dvou lexikálních jednotek s homonymními morfologickými informacemi nejsme schopni (bez mimojazykových znalostí) provést morfologickou disambiguaci a rozhodnout, které z nich odpovídá kterému valenčnímu doplnění, budeme ilustrovat ještě následujícími příklady.

Příklady:

(II.17a) *Naše děti ohrožují divoká zvířata.* (Daneš (1964))

(II.17b) *Představil tchyni hospodyni.* (Panevová (1981))

(II.17c) *Kienovy kanibaly udusily růže.* (E. Canetti - Zaslepení)

---> *Kienovy kytky udusily růže.*

(II.17d) *Včera večer zasáhl osmipatrový dům naproti granát.* (MF Dnes, 1996)

(II.17e) *Koncert postižených financuje folkový festival.* (MF Dnes, 1996)

(II.17f) *Starý parlament Běloruska má OBSE za legitimního nástupce.* (MF Dnes, 1997)

(II.17g) *Učitelé se brání: děti novodobé dějiny nezajímají, mnohem raději mají starověk a středověk.* (MF Dnes, 1996)

¹⁸ Křížením valenčních požadavků a stanovením jejich preferencí se zabýváme v kapitole IV.2.5.

(II.17h) *Všichni vystavující byli tehdy pořadatelé představeni řediteli Gallery XX. Century.* (MF Dnes, 1997)

(II.17i) *Škola překvapuje nejen žáky, ale také kantory.* (MF Dnes, 1997)

Pokud věty na základě jejich obsahu analyzujeme, často není naplněn nepříznakový slovosled (tj. subjekt – sloveso – objekt), viz (II.17a), (II.17c), (II.17e) a (II.17f).

Ve větě (II.17i) substantiva *žáky* a *kantory* mohou být v instrumentálu (Kým překvapuje?), nebo v akuzativu (Koho překvapuje?). Uvažujeme valenční rámec slovesa překvapovat:

překvapovat ... Act | Pat (Ins) Addr (Acc)

Obě obligatorní valenční doplnění, patient i adresát, mají často hodnotu všeobecného participantu (General Participant). Formy *žáky*, *kantory* tedy mohou zaplňovat jak pozici odpovídající patientu (adresát všeobecný), tak pozici odpovídající adresátu (patient všeobecný).

□

Homonymie určení substantiva a jeho přívlastku neshodného

Následující věty charakterizuje dvojí možné určení řídicího členu jmenné skupiny – bez znalosti situace zde není jasné, který větný člen je řídicí a který člen má být chápán jako jeho neshodný atribut.

Příklady:

(II.18a) ... *chránit bezpečnostní zóny vyhlášené OSN.* (MF Dnes, 1997)

(II.18b) (*zazdění oken*) *Pokušení pozorovat ruch na ulici – čas ukrádající nešvar, se kterým člověk zřejmě přichází na svět – zmizelo s bočními okny.* (E. Canetti – Pochodeň v uchu)

(II.18c) *Když přičtete delegáty posvěcené paragrafy ve stanovách (umožňující snadno vyloučit oponenty), pak je na místě obava, zda ČSSD ...* (MF Dnes, 1997)

(II.18d) *Proto není selekční tlak determinující struktury a funkce kultury tak neúprosný, jako je tomu ve vývoji druhů.* (MF Dnes, 1998)

---> *Proto není selekční tlak determinující funkce kultury tak zásadní, jako ...*

Různé významy věty (II.18a) se vyznačují odlišnou závislostní strukturou – jde buď o *zóny*_{Acc, pl} (*vyhlášené*_{Acc, pl} (*OSN*_{Ins})), tj. *zóny*, které *vyhlásila OSN* (adjektivní přívlastek v postpozici), nebo o *zóny*_{Acc, pl} (*vyhlášené*_{Gen, sg} *OSN*_{Gen, sg}), tj. s preponovaným adjektivním přívlastkem.

Jmennou skupinu v (II.18b) lze analyzovat dvojím způsobem, buď jde o *nešvar*, který *ukrádá čas* (adjektivní přívlastek v prepozici rozvitý substantivem), nebo o *čas*, který *ukrádá nešvar* (adjektivní přívlastek v postpozici rozvitý substantivem). Bez mimojazykové znalosti zde nelze určit, který větný člen je řídicí a který člen má být chápán jako jeho neshodný atribut.

V (II.18c) jde buď o *paragrafy*, které jsou *posvěceny delegáty* (adjektivní přívlastek v prepozici rozvitý substantivem), nebo o *delegáty*, kteří jsou *posvěceni paragrafy* (adjektivní přívlastek v postpozici rozvitý substantivem). Bez pragmatické znalosti zde není jasné, který větný člen je řídicí a který člen má být chápán jako jeho neshodný atribut.

Obdobně ve větě (II.18d).

□

V následující větě vůbec nejsme schopni rozhodnout (dokonce ani v rámci širokého kontextu), co měl autor vlastně na mysli.

Příklad:

(II.18e) ..., stále si ještě zoufal a říkal, že veliký bol, kterým ho stihl Pán, jen umocňuje bolest, již mu působí lidé. (RR)

Tvary substantiv *bol* i *bolest* mohou vyjadřovat nominativ i akuzativ, větu tedy lze vyjádřit pomocí dvou morfematických zápisů:

Bol_{Nom} umocňuje $bolest_{Acc}$ / Bol_{Acc} umocňuje $bolest_{Nom}$

□

II.2.4. Fonologická homonymie v lexiku

Taková homonymie není z hlediska syntaktické analýzy příliš zajímavá – už z morfologické analýzy totiž získáme u každé homonymní lexikální jednotky několik různých čtení, se kterými potom musíme pracovat paralelně.

Uvedme zde pouze několik ilustrativních příkladů.

Příklady:

(II.19a) Sešit je na okraji. (Panevová (1981)) (substantivum / sloveso)

(II.19b) Chléb se solí. (Panevová (1981)) (předložka+substantivum / refl. částice+sloveso)

(II.19c) B. Doležal a J. Lopatka byli z celé skupiny autorů časopisu *Tvář* kritiky nejobávanějšími. (KP RR, Blažek)

(*kritiky* – buď genitiv sg substantiva *kritika*, nebo instrumentál pl substantiva *kritik*)

(II.19d) Byl (přímo) proti jaderné elektrárně. (adverbium / předložka)

(II.19e) Toto postižení lidé chápou. (adjektivum / substantivum)

(II.19f) Stanou se soudruhy. (refl. částice / předložka)

(II.19g) Tento kohoutek je nejmenší. (Sgall (1967))

(II.19h) Mlhovky stály místo ministra dopravy. (Panevová) (substantivum / předložka)

Tuto tzv. lexikální homonymii podle Těšitelové (1974) charakterizuje fonologické splývání různých lexémů (*proti*, *se*, *kohoutek*, *místo*). Do této třídy ovšem spadají i případy, kde splývají pouze některé tvary různých lexémů (*sešit* X *sešít*, *sůl* X *solit*, *kritika* X *kritik*, *postižený* X *postižení*). □

II.2.5. Eliminace homonymie

Zatím jsme se zabývali větami víceznačnými – tedy větami, ve kterých se vyskytuje nějaký (aspoň jeden) homonymní prostředek, přičemž homonymie daná tímto prostředkem způsobuje (alespoň) dvojí čtení celé větné konstrukce.

Na nižších rovinách jazykového systému může ovšem docházet (a také dochází) k situaci, kdy forma na rovině n má dvě (nebo víc) funkcí na rovině $n+1$ – jde tedy o homonymii na rovině n , ale jedno čtení (resp. všechna čtení až na jedno) je eliminováno na

vyšších jazykových rovinách. To se týká především fonologické homonymie. V takovém případě nás zajímá, nejen kde (na které rovině) homonymie vzniká, ale také kterým prostředkem je vyloučena, resp. z jakého důvodu jsou některé analýzy upozaděny.

Fonologická homonymie a její eliminace na vyšších rovinách

Příklady:

(II.20a) *Profesionálnímu oku fotografie získané z filmu ani vzdáleně nepřipomínají skutečnou M.M. (MF Dnes, 1998)*

(II.20b) *Před vlastním zásahem policie neprodyšně uzavřela okolí kostela. (MF Dnes, 1997)*

(II.20b') *Před vlastním zásahem policie [hlídka] neprodyšně uzavřela okolí kostela.*

(II.20c) *(chlapec polykal korále) Když sestra toho chlapce pohřešila korále, mohla si oči vyplakat.*

(II.20d) *Bez důkladného studia (určité) věci pochopit nelze.*

(II.20e) *Další odpůrci indického guru obvinili, že ... (RR 35, Brezina)*

(II.20f) *Trosky zaměstnance ... zavalily. (Panevová)*

(II.20g) *A soud pana Bittnera osvobozuje. (MF Dnes, 1996)*

(II.20h) *Zřejmě nejslavnější skandály opředený román Nahý oběd ... W. S. Borroughse ... vydalo nakladatelství ... (MF Dnes, 1996)*

Ve větách (II.20a) a (II.20b) jde o **homonymii podmětu a přívlastku neshodného**, ve větách (II.20c)-(II.20g) o **homonymii předmětu a přívlastku neshodného**, přičemž větná struktura s přívlastkem je správná pouze v (II.20c).

Tento typ vět bývá v literatuře nazýván „garden path sentences“.

Ke správné analýze těchto vět na tektogramatické rovině často pomůže valenční rámec sloves – při nesprávné analýze totiž nelze tento rámec zaplnit (koncept valenčních rámců ovšem nelze automaticky uplatnit na povrchové rovině, kde zřejmě lze – alespoň v některých kontextech, jako je např. odpověď na otázku – vypustit libovolný valenční člen).

Ve větách (II.20a) a (II.20b) chybí aktor – ten je ovšem vždy vypustitelný. Ve větě (II.20a) lze tedy substantivum *fotografie* analyzovat buď jako subjekt, tj. větný člen odpovídající hloubkovému aktoru, nebo jako přívlastek neshodný rozvíjející substantivum *oko*. V obou případech se členy *získané z filmu* identifikují jako přívlastek neshodný rozvíjející substantivum *fotografie*. Ovšem druhé čtení je zřejmě v obvyklých kontextech z hlediska mluvčího málo přijatelné.

Varianta (II.20b') ukazuje kontext, který dovoluje analyzovat substantivum *policie* jako neshodný atribut substantiva *zásah*.

Jiná situace nastává ve větě (II.20c): analyzuje-li se člen (*tento*) *chlapec* jako patient, věta se stává nekonsistentní (člen *korále* nelze určit).

Ve větách (II.20d)-(II.20g) při chybném, lépe řečeno málo pravděpodobném čtení chybí objekt (jemu odpovídající člen valenčního rámce není ani v jednom případě teoreticky vypustitelný).

Ve větě (II.20h) se chybná analýza přeruší v okamžiku, kdy se ve větě objeví tři substantiva v nominativu (a přitom nejde o jediný komplexní větný člen jako je např. koordinace nebo řetězec typu titul + jméno). □

Fonologická homonymie v lexiku a její eliminace na vyšších rovinách

Všímáme si zde pouze homonymie, která se týká různých slovních druhů a může tedy narušit stavbu věty.

Příklady:

(II.21a) *Věděla (Nad'a), že žena je zlo a že jí ublíží.* (J. Topol – Anděl)

(II.21b) *Vysvětlila své vlastní pojetí problematiky.*

(II.21c) *... má kladný vliv na zdraví.* (LN, 1995)

(II.21d) *Stěny a klenby domů ničí voda.* (MF Dnes, 1994)

Morfologická analýza nabízí několik možností. Při chybném výběru lexikálních jednotek většinou vůbec nelze postavit závislostní strom. V druhé části (II.21a) (*jí* – V/Pron), v (II.21b) (*vlastní* – V/Adj) a v (II.21c) (*zdraví* – V/N/Adj) se při špatném čtení objevují dvě určitá slovesa v jediné větě klauzi, naopak první část (II.21a) (*je* – V/Pron) a (II.21d) (*ničí* – V/Pron) neobsahuje ani jedno. □

Příklady:

(II.22a) *V níž položených oblastech ...* (MF Dnes, 1997)

(II.22b) *Podle jejich slov ...*

(II.22b') *Podle jejich slov zneužil.*

(II.22b'') *Podle jejich slov bylo vše jinak.*

(II.22c) *Po zabití 20_{Nom/Gen} Srbů poblíž Sarajeva ...* (MF Dnes, 1998)

(II.22c') *Po zabití 20_{Gen} draků poblíž jeskyně rytíři odjeli.*

(II.22c'') *Po zabití 20_{Nom} draků poblíž jeskyně strašilo.*

Věta (II.22a) obsahuje homonymní tvar *níž* (Adv/Pron), (II.22b) slovo *podle* (Adv/Prep), ve větě (II.22c) nemá číslovka psaná číslicí specifikovaný pád (v úvahu přichází nominativ a genitiv). Pro obě čtení existují kontexty, kontext pro méně pravděpodobné čtení se ovšem hledá poměrně obtížně. □

Závěr:

Nutno říci, že jsme nenalezli vhodné víceznačné věty, které by ilustrovaly syntaktickou homonymii. Věty, které byly prezentovány, byly buď věty nejasné na úrovni obsahové (nerozlišitelné na základě jazykové struktury), nebo byla jejich víceznačnost způsobena jevy, které je vhodné zachytit už na morfematické rovině.

Uvedené příklady napovídají, že naše práce se soustředí především na morfematickou homonymii, případně se bude dotýkat homonymie fonologické. Fonologickou homonymií v lexiku se dále zabývat nebudeme. □

II.3. HOMONYMIE PŘEDLOŽKOVÝCH SKUPIN

Pro homonymii Pg je příznačná především **morfematická homonymie**, tedy možnost, že zkoumaná Pg může být v konkrétní větě považována za rozvití různých větných členů.

Z čistě syntaktického hlediska může předložková skupina v principu modifikovat:

- (a) jakékoliv syntaktické jméno (slovo, které má syntaktické vlastnosti substantiva) předcházející v povrchové reprezentaci věty zkoumanou Pg; jakýkoliv tvar slovesa slouží jako levá zarážka. (Existují ovšem konstrukce, při jejichž analýze je třeba uvolnit tuto podmínku – jde především o substantiva odvozená od sloves, která mají ve svém rámci doplnění realizovaná Pg, a o slovesně jmenné výrazy (verbonominal collocations); podrobněji se jimi budeme zabývat v kapitole IV.2.3.);
- (b) jakékoliv sloveso (významové sloveso, modální sloveso stejně jako spona);
- (c) jakékoliv přídavné jméno;
- (d) některá adverbia (tato možnost není v této práci zkoumána).

Věty (I.1) a (I.3) ilustrují typy závislosti (a) a (b). Podívejme se nyní ještě na možnost (c), tj. na předložkovou skupinu, která rozvíjí přídavné jméno.

Příklad:

(II.23) *I parlamentní demokracie musí dostát nárokům **kladeným na demokracii** jako takovou.* (PDT, bm227zua.fs #22, zkráceno)

Přídavné jméno *kladený* – odvozené od slovesa *klást*, které společně s podstatným jménem *nárok* tvoří slovesně jmenné spojení *klást nároky na něco* – je rozvíjeno Pg *na demokracii (jako takovou)*. Podrobnější analýze této věty se budeme věnovat v příkladu (IV.27) a také v kapitole o valenčních rámcích adjektiv (IV.2.4.). □

Na tomto místě je potřeba zmínit skutečnost, že není podstatné, že v příkladu (II.23) jde o deverbativní adjektivum, tj. adjektivum odvozené od slovesa. Také původní adjektiva mohou být rozvíjena předložkovými skupinami, jak o tom svědčí příklad (II.24).

Příklad:

(II.24) *Na první pohled **hrůzostrašný** sjezd **není** nebezpečnější než jiné lyžařské disciplíny.* (Kuboň (1999))

Pg *na (první) pohled* může být analyzována dvěma způsoby – jako adverbialní rozvití přídavného jména *hrůzostrašný*, nebo jako doplnění sponového slovesa *být*. □

Jak jsme se snažili ukázat, morfematická homonymie je pro předložkové skupiny typická. Některé české předložky, jak už bylo uvedeno, však mají dvojí (výjimečně i trojí) možnou rekcí (pojí se se dvěma, výjimečně se třemi pády, Šmilauer (1966)), proto se v souvislosti s Pg (při vhodném lexikálním obsazení) může objevit i **fonologická homonymie**. Jako příklad můžou posloužit předložky *na* a *v* (obě se pojí jednak s akuzativem, jednak s lokálem) a předložka *za* (vyžaduje genitiv, *za vlády císaře*, akuzativ, *za koho mě máš*, nebo instrumentál, *za domem*).

Z tohoto hlediska se podrobněji vrátíme k větě (II.4e).

Příklad: (pokračování)

(II.4e) ... (*úřad*) **hledá policisty na hranici**. (MF Dnes, 1996)

Předložková skupina *na hranici* splňuje podmínky dvou typů homonymie:

- (a) morfemtické homonymie – Pg rozvíjí
 - (aa) buďto sloveso *hledat*,
 - (ab) nebo podstatné jméno *policisty*;
- (b) fonologické homonymie – Pg může rozvíjet sloveso i substantivum
 - (ba) buďto jako volné doplnění místa (Locative, *na+Loc*), odpověď na otázku Kde?
 - (bb) nebo jako volné doplnění účelu (Aim, *na+Acc*).

Zatímco možnost výskytu morfemtické homonymie je dána obecně platným syntaktickým pravidlem, fonologická homonymie je způsobena souběhem charakteristik dvou lexikálních jednotek (předložky vyžadující akuzativ nebo lokál a substantiva s homonymními tvary pro akuzativ a lokál jednotného čísla). □

Podobná situace nastává i v následující větě.

Příklad:

(II.25) *Brankář Barda a trenér Haber byli jako jediní schopni zabezpečit přípravu na MS 1993 ve Švédsku*. (PDT, bmc26aba.fs #18, zkráceno)

Předložková skupina *na MS* splňuje podmínky dvou typů homonymie:

- (a) morfemtické homonymie – Pg rozvíjí
 - (aa) sloveso *zabezpečit* (jako volné určení místa, odpověď na otázku Kde?),
 - (ab) substantivum *příprava* (odpověď na otázku Kde?);
- (b) fonologické homonymie – Pg může rozvíjet podstatné jméno *příprava*
 - (ba) jako valenční doplnění, patient (s povrchovým vyjádřením *na+Acc*),
 - (bb) jako neshodný nevalenční atribut s lokálním významem, odpověď na otázku Kde?

Určení Pg jako homonymní je dáno jednak syntaktickým pravidlem pro připojování Pg, jednak vlastnostmi konkrétního lexikálního naplnění, a to nesklonnou zkratkou (dva různé pády (= morfémy) jsou vyjádřeny jedinou fonologickou jednotkou, jde tedy o fonologickou homonymii). □

Z hlediska automatického určování homonymních větných členů je zajímavá hlavně morfemtická homonymie, kde je potenciálně homonymní větný člen určován na základě obecných syntaktických pravidel. (Z opačného pohledu: obecná syntaktická pravidla umožňují vyloučení neadekvátních závislostí, případně upozadění málo pravděpodobných významů, čtení věty.)

Zatímco potenciální morfemtická homonymie je zpracovávána během syntaktické analýzy, možnost fonologické homonymie je detekována již při morfologické analýze – zkoumané lexikální jednotky v takovém případě nesou nejednoznačnou morfologickou informaci. V takovém případě se syntaktická analýza rozpadá do několika větví, ve kterých se postupně

zpracovávají všechny možné kombinace vstupních morfologických charakteristik jednotlivých slovních forem. Pokud dvě (nebo více) větví syntaktické analýzy s různou kombinací těchto charakteristik vedou k úspěšnému rozboru věty, je konstatována víceznačnost věty daná splnutím fonologické realizace některých jednotek vyšší roviny.

Toto je také důvod, proč je potřeba – z hlediska automatické syntaktické analýzy – soustředit se zejména na morfematickou homonymii a na návrh obecných syntaktických pravidel, které by ji umožnily zachytit. Jak bylo naznačeno, tato pravidla se opírají zejména o povrchový slovosled, o valenční vlastnosti sloves, substantiv a adjektiv, případně o sémantické rysy jednotlivých lexikálních jednotek.

Shrnutí

Na základě analýzy motivačních příkladů můžeme tedy shrnout: zatímco morfematická homonymie předložkových skupin je jevem systémovým, daným obecnými syntaktickými pravidly, výskyt fonologické homonymie je způsoben více méně náhodným souběhem vlastností konkrétních lexikálních jednotek (při změně lexikálního obsazení lze víceznačnost eliminovat).

Předmětem našeho zkoumání se proto stanou případy morfematické homonymie předložkových skupin. Naším úkolem je navrhnout relevantní kritéria (použitelná pro účely automatického zpracování přirozeného jazyka) umožňující odhalení a popis tohoto typu homonymie Pg, tj. kritéria dovolující rozhodnout, které slovo či slova věty mohou být považována za větné členy rozvíjené zkoumanou Pg. Budeme zkoumat kritéria čistě syntaktická, kritéria sémantická i kritéria využívající kombinovanou informaci.

Zopakujme zde stručně, co nás vede – kromě důvodů čistě teoretických, tj. snahy blížit se ke zcela adekvátnímu popisu jazyka – k zájmu o správné určení větných členů rozvíjených předložkovou skupinou. Obecně totiž věta s Pg analyzovaná jako atributivní (nominální) doplnění, a stejná věta s touž Pg analyzovaná jako adverbální (slovesné) doplnění nemají stejný význam (mohou mít různé pravdivostní podmínky). Proto je důležité určit skutečný/skutečné řídicí člen/členy Pg v závislostním stromu reprezentujícím zkoumanou větu. Možnost, že různá čtení takové věty ve specifickém kontextu sdílejí kognitivní obsah (mají stejné pravdivostní podmínky), je náhodná a není obecně dána. Často ji lze určit pouze na základě mimojazykových znalostí.

ČÁST III.

AUTOMATICKÉ VYHLEDÁVÁNÍ SYNTAKTICKÝCH STRUKTUR ‚PODEZŘELÝCH‘ VZHLEDEM K HOMONYMII PŘEDLOŽKOVÝCH SKUPIN

V první fázi práce bylo potřeba získat dostatečné množství podkladových dat umožňujících návrh kritérií pro automatickou disambiguaci předložkových skupin, respektive pro určení jejich (morfematické) homonymie. Jako základního zdroje (kromě ručního sbírání vět) jsme použili Pražský závislostní korpus (PDT, Prague Dependency Treebank). V této části chceme popsat nástroje navržené pro automatické vyhledávání ‚podezřelých‘ vět, tj. vět, jejichž povrchová struktura (přesněji řečeno: struktura na analytické rovině) odpovídá některé z vytipovaných ‚podezřelých‘ struktur.

Pražský závislostní korpus je český korpus s bohatým anotačním schématem (popis viz např. Hajič (1998)), který je strukturován do třech rovin: obsahuje plnou morfologickou analýzu na nejnižší úrovni, syntaktickou anotaci na prostřední, tzv. analytické rovině a anotaci na úrovni jazykového významu. PDT se skládá z obecných novinových článků, ekonomických zpráv a analýz, z článků populárně vědeckých časopisů a z textů z oboru informačních technologií.

Pro vyhledávání ‚podezřelých‘ vět, tedy vět, jejichž syntaktická struktura signalizuje možný výskyt homonymie Pg, je využívána anotace na tzv. analytické rovině. Jde o pomocnou rovinu (nemá oporu v teorii FGD), kterou lze považovat za aproximaci roviny povrchové. Závislostní formalismus je zde zachován. Struktura věty je zachycena pomocí souvislého orientovaného acyklického grafu, kde každý uzel kromě kořene má právě jednoho otce (jde tedy o strom). Uzly stromu jsou ohodnoceny komplexními symboly (páry atribut – hodnota). Každému slovu ve větě (až na drobné výjimky) odpovídá jeden uzel stromu (plus jeden uzel, kořen stromu, reprezentující celou větu). Lineární uspořádání uzlů odpovídá povrchovému slovosledu. Předložková skupina je zachycena jako podstrom závislostního stromu, na jehož vrcholu je uzel reprezentující předložku.

Vyvstávají otázky, proč jako zdroje dat využíváme PDT, k čemu potřebujeme jeho syntaktické struktury (které už dále nevyužíváme), proč se neomezíme na řetězce s morfologickými značkami (a tedy informace dané morfologickou analýzou a povrchovým slovosledem). Pokusme se zde na tyto otázky odpovědět.

Předložkové skupiny prototypicky splňují určitá slovosledná omezení,¹⁹ na základě slovosledu se obvykle dají vytipovat všechny členy věty, na kterých Pg potenciálně závisí (které

¹⁹ Podrobně se těmito omezeními budeme zabývat v oddíle IV.1.

mohou být Pg rozvíjeny); lze definovat ‚zarážky‘ pro hledání řídicích členů Pg. To by mluvilo pro využití morfologicky označovaného korpusu, syntaktickou strukturu vět dále nepotřebujeme. Nicméně existují věty (a byly nalezeny v PDT), ve kterých taková slovosledná omezení splněna nejsou. Potom k určení možných řídicích členů Pg poslouží např. syntaktická struktura věty, kterou poskytuje anotace na analytické rovině PDT.

Protože se dále chceme zabývat (také) jazykovými jevy, ve kterých je vhodné uvolnit slovosledná omezení (kapitola IV.1.2.) a které mohou vést například k neprojektivnímu doplnění substantiva realizovanému předložkovou skupinou, využíváme pro shromáždění podkladového materiálu syntaktické struktury PDT.

S tímto záměrem také souvisí rozhodnutí nebrat při vyhledávání ‚podezřelých‘ vět v úvahu povrchový slovosled (k tomuto rozhodnutí se ještě vrátíme v závěru kapitoly).

III.1. DEFINOVÁNÍ ‚PODEZŘELÝCH‘ STRUKTUR

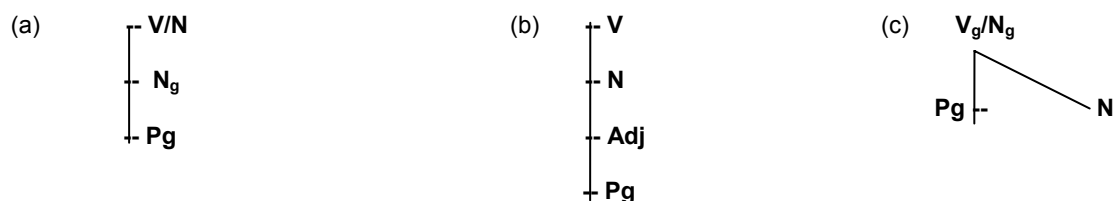
Věty obsažené v PDT jsou disambiguované, pro každou větu byla (ručně) vybrána jediná syntaktická struktura, kterou anotátor považoval (vzhledem k širšímu kontextu i ke své mimojazykové znalosti) za adekvátní. Vyskytují se zde ovšem uzly s tzv. zdvojenými analytickými funkcemi – Atr/Adv, Adv/Atr, Atr/Obj, Obj/Atr a Atr/Atr. Tyto zdvojené funkce zachycují dvojí možnou syntaktickou strukturu věty, avšak pouze v rámci jediného významu věty (pro který se, s přihlédnutím ke kontextu, anotátor rozhodl). Platí obecná zásada, podle které se uzel s několika možnými funkcemi určuje jako rozvíetí jména (N_g), přičemž jeho zdvojená analytická funkce poukazuje na možnost chápat ho i jako rozvíetí uzlu řídicího toto N_g – tj. slovesa (Atr/Adv, Adv/Atr, Atr/Obj, Obj/Atr) nebo jiného jména (Atr/Atr). Vícečetná homonymie zůstává nezachycena.

Zkoumání kritérií, která by umožnila zachycení homonymie předložkových skupin, nelze stavět pouze na větách s P_g se zdvojenou analytickou funkcí – právě proto, že věty v PDT jsou (vzhledem k významu) disambiguované. Syntaktické struktury zde nepostihují ta možná čtení věty, která se liší pravdivostními podmínkami od čtení vybraného anotátorem. Taková čtení jsou ovšem pro popis homonymie P_g stejně důležitá jako čtení, která mají (díky specifickému kontextu) význam stejný. To vedlo k navržení vlastních nástrojů pro identifikaci ‚podezřelých‘ struktur vzhledem k homonymii P_g .

Na základě zkoumání vět, ve kterých se objevuje homonymní předložková skupina – jde především o homonymii typu II. (viz oddíl II.2.) – a na základě vzorku vět z PDT, které obsahují předložkovou skupinu nesoucí zdvojenou analytickou funkci, jsou navrženy tři základní ‚podezřelé‘ struktury (tj. syntaktické struktury, které obsahují potenciálně homonymní předložkovou skupinu).

Definice (základní ‚podezřelé‘ syntaktické struktury vzhledem k homonymii P_g):

Definujeme tři základní ‚podezřelé‘ syntaktické struktury vzhledem k homonymii P_g (linearizovaná forma tu nemusí odpovídat slovosledu):



Obr.8: Základní ‚podezřelé‘ syntaktické struktury vzhledem k homonymii P_g

(a) syntaktická struktura, kde je sloveso nebo jméno (V/N)²⁰ rozvíjeno jiným jménem (N_g), které je samo rozvíjeno zkoumanou P_g (tj. $V/N (N_g (P_g))$ v linearizované formě);

(b) syntaktická struktura, kde je jméno (N) rozvíjeno adjektivem (Adj), které je samo rozvíjeno zkoumanou P_g , celý podstrom rozvíjí sloveso (V) (tj. $V (N (Adj (P_g)))$ v linearizované formě);

²⁰ Jménem je zde míněno syntaktické substantivum, viz poznámka 2.

(c) syntaktický podstrom, kde je sloveso nebo jméno (V_g/N_g) rozvíjeno zkoumanou P_g ; jiné jméno (N) se objevuje jako sesterský uzel této P_g (tj. $V_g/N_g (P_g N)$ v linearizované formě).

□

Poznámka:

Toto jsou skutečně pouze základní ‚podezřelé‘ struktury vzhledem k homonymii P_g .

- U struktur (a) a (b) je nutné hledat potenciální řídicí jméno nebo sloveso na celé ‚cestě‘ od testované předložky ke kořeni stromu, resp. ke kořenu podstromu reprezentujícímu klauzi.

- U struktury (c) je potřeba testovat všechny sestry zkoumané předložky.

- Je potřeba počítat s tím, že se všechna jména mohou vyskytovat v prostých nebo předložkových pádech (což má na analytické rovině podstatný vliv na tvar hledaného podstromu; jako P_g označujeme jen tu předložkovou skupinu, která je právě předmětem analýzy).

- Na libovolném místě podstromu se můžou objevit koordinované větné členy (což má opět na analytické rovině podstatný vliv na tvar podstromu).

□

III.1.1., ‚Podezřelé‘ struktury v příkladech

Rozeberme na konkrétních větách z ručně sebraného materiálu, které obsahují předložkové skupiny $na+Acc$, $na+Loc$, $v+Acc$ a $v+Loc$, jak na analytické rovině PDT vypadají struktury, které nás budou zajímat. a to s jednou výjimkou – ve schématech nezachycujeme povrchový slovosled, který na analytické rovině zachycen je.

‚Cesta ke kořeni stromu‘ – typy struktur (a) a (b)

Základní struktury



Obr.9: ‚Podezřelé‘ syntaktické struktury vzhledem k homonymii P_g – typy (a) a (b)

Najdeme testovanou předložku a její řídicí uzel (N_g); procházíme cestu ke kořeni stromu a zkoumáme, zda se na ní neobjeví další sloveso (V) nebo jméno (N),²¹ které by (z čistě syntaktického hlediska) mohlo být řídicím uzlem předložkové skupiny. Resp. najdeme předložkovou skupinu rozvíjející adjektivum (Adj) a hledáme sloveso (V), které by mohlo být jejím řídicím uzlem.

²¹ Pro úplnost: hledané i řídicí jméno může být v prostém (N) nebo předložkovém pádu (N_p), přesněji bychom tedy měli psát N/N_p , resp. $N_g/(N_p)_g$. Protože tento fakt nemá vliv na vyhledávání, používáme jednodušší označení N, resp. N_g .

Příklady:

- (III.1) *Chlapci honí kamarády na kolech.* (kamarády na kolech X honí na kolech)
(III.1') *Mluvil o prázdninách na Šumavě.* (prázdninách na Šumavě X mluvil na Šumavě)
(III.2) *Neschopnost ... způsobila skluz v realizaci autonomie na území Golanských výšin.*
(autonomie na území (Golanských výšin)) X realizace na území (Golanských výšin)
X skluz na území (Golanských výšin) X způsobila na území (Golanských výšin)
(III.3) *Rovná na ramínku vystavený kabát.* (na ramínku vystavený X rovná na ramínku)

Předpokládáme, že věty (III.1) a (III.1') by měly v PDT přiřazenu strukturu $V(N_g(Pg))$, jde tedy o typ odpovídající ‚podezřelé‘ struktuře (a). Věta (III.2) se strukturou $N(N_g(Pg))$ odpovídá rovněž typu (a), věta (III.3) se strukturou $V(N(Adj(Pg)))$ koresponduje s typem (b). V závorce za větou uvádíme ve formě tzv. skladebných dvojic možné závislosti Pg (jako první vždy tu strukturu, o které předpokládáme, že by byla (primárně) zachycena v PDT).

□

‚Podezřelé‘ struktury nezachycují pochopitelně všechny typy struktur, ve kterých se může vyskytnout (potenciálně) homonymní Pg. Snažili jsme se popsat ty případy, které se vyskytují často. Například následující věta by nebyla automaticky označena jako potenciálně homonymní (lépe řečeno: neobsahuje strukturu vytipovanou z hlediska homonymie Pg jako zajímavou).

Příklad:

- (III.4) *Neschopnost dohodnout se na jednání způsobila zhoršení vzájemných vztahů.*
(dohodnout se na jednání X na jednání způsobila)

Předpokládáme, že věta (III.4) by v PDT měla strukturu $V(N(V_g(Pg)))$. Na základě ‚podezřelých‘ struktur typu (a) a (b) by nebyla vybrána jako věta obsahující potenciálně homonymní Pg, přestože homonymní nepochybně je.

□

Tyto dvě základní ‚podezřelé‘ struktury se poněkud komplikují při výskytu koordinace (koordinace je zahrnuta do struktur kvůli ‚širšímu záběru‘ připravovaných maker, tj. pro rozmnožení hledaného materiálu). Koordinovaný přitom může být kterýkoliv jejich uzal.

V následujících odstavcích ukážeme výše navržené ‚podezřelé‘ struktury obohacené o koordinaci. (Koordinaci značíme pro přehlednost hranatými závorkami.)

Předložka řídí koordinovaná jména

Tato skutečnost nemá vliv na další zpracování, uveďme jako příklady modifikace vět (III.1)-(III.3).

Příklady:

- (III.1a) *Chlapci honí kamarády na kolech a tříkolkách.*
(kamarády na [kolech a tříkolkách] X honí na [kolech a tříkolkách])
(III.2a) *Neschopnost ... způsobila skluz v realizaci autonomie na obsazeném břehu i přilehlé vrchovině.*

(autonomie na [(obsazeném břehu) i (přilehlé vrchovině)])

X realizace na [(obsazeném břehu) i (přilehlé vrchovině)]

X skluz na [(obsazeném břehu) i (přilehlé vrchovině)]

X způsobila na [(obsazeném břehu) i (přilehlé vrchovině)])

(III.3a) Rovná na ramínkách a figurínách vystavené kabáty.

(na [ramínkách a figurínách] vystavené X rovná na [ramínkách a figurínách]) □

Pg je členem koordinační skupiny

Pg se často vyskytují v koordinaci (se dvěma stejnými nebo s různými předložkami). V takových větách pracujeme s celou koordinační skupinou, hledáme její potenciální řídicí uzel – jméno nebo sloveso.

Příklady:

(III.1b) Chlapci honí kamarády na kolech a na tříkolkách.

(kamarády [(na kolech) a (na tříkolkách)]) X honí [(na kolech) a (na tříkolkách)])

(III.2b) Neschopnost ... způsobila skluz v realizaci autonomie na území Golanských výšin a v Jeruzalémě.

(autonomie [(na území (Golanských výšin)) a (v Jeruzalémě)])

X realizace [(na území (Golanských výšin)) a (v Jeruzalémě)]

X skluz [(na území (Golanských výšin)) a (v Jeruzalémě)]

X způsobila [(na území (Golanských výšin)) a (v Jeruzalémě)])

(III.3b) Rovná na ramínkách a na figurínách vystavené kabáty.

([(na ramínkách) a (na figurínách)] vystavené

X rovná [(na ramínkách) a (na figurínách)]) □

Pg rozvíjí celou koordinační skupinu, resp. koordinované Ng/Adj

(i) Předložková skupina může rozvíjet celou koordinační skupinu složenou z Ng/Adj. Potom je možnost, že – kromě nadřazeného jména, resp. slovesa – rozvíjí také některé jednotlivé členy této koordinace.

(ii) Předložková skupina, která rozvíjí Ng/Adj, může v jistých případech rozvíjet i celou koordinační skupinu, ke které toto Ng/Adj patří. V tomto případě předpokládáme, že v PDT je primárně zachycena struktura uvedená v příkladech jako druhá.

Příklady:

(III.1c) Chlapci honí kamarády a kamarádky na kolech.

([kamarády a kamarádky] na kolech X [kamarády a (kamarádky na kolech)]) X honí na kolech)

(III.2c) Neschopnost ... způsobila skluz v zajištění pořádku a práva na území Golanských výšin.

([pořádku a práva] na území (Golanských výšin))

X [pořádku a (práva na území (Golanských výšin))] X zajištění na území (Golanských výšin)

X skluz na území (Golanských výšin) X způsobila na území (Golanských výšin)

(III.3c) Rovná na ramínkách vystavené a naaranžované šaty.

((na ramínkách [vystavené a naaranžované]) šaty

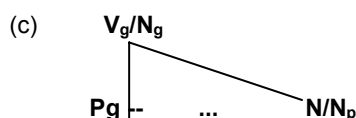
X [(na ramínkách vystavené) a naaranžované] šaty X rovná (na ramínkách) □

Poznámka:

Různé typy koordinace se můžou kombinovat. Jinou otázkou zůstává, jak ‚vysoko‘ je vhodné prohledávat závislostní strom reprezentující větu, zda existuje (z hlediska lingvistiky) pravidlo omezující vertikální vzdálenost potenciálního řídicího uzlu (tj. délku cesty v závislostním grafu). □

‚Testování sesterských uzlů‘ – typ struktury (c)

Základní struktury



Obr.10: ‚Podezřelá‘ syntaktická struktura vzhledem k homonymii Pg – typ (c)

Najdeme zkoumanou předložku a její řídicí uzlu (N_g/V_g); prohledáváme ‚sestry‘ této předložky v závislostní reprezentaci věty a testujeme, zda nejde o jméno (v prostém nebo předložkovém pádu, N/N_p)²² – to by se mohlo stát řídicím uzlem předložkové skupiny.

Testování sesterských uzlů je přínosné zejména pro V_g .

Příklady:

(III.5) Změna vychází z požadavků rozpočtu na příští rok.

(požadavků na (příští rok) X rozpočtu na (příští rok))

(III.5') Bojují s nárůstem loupeží na našem území.

(nárůstem na (našem území) X loupeží na (našem území))

(III.6) Navštívil hřbitov v Praze na Vyšehradě.

(hřbitov na Vyšehradě X v Praze na Vyšehradě)

(III.6') Miloval procházky v lesích na břehu řeky.

(procházky na (břehu řeky) X v lesích na (břehu řeky))

(III.7) Úřad hledá policisty na hranici. Účel (Aim), Kde?

(hledá na hranici X policisty na hranici)

²² Zde má skutečnost, zda je potenciální řídicí jméno v prostém nebo předložkovém pádu, vliv na návrh vyhledávacího makra, proto v tomto případě rozlišujeme N a N_p . Jméno, pokládané v PDT za řídicí člen P_g , označujeme N_g bez ohledu na to, zda se vyskytuje s předložkou nebo bez ní.

(III.8) *Bydlel v Praze na Hradčanech.*

(bydlel na Hradčanech X v Praze na Hradčanech)

(III.8') *Rád pobýval v lesích na břehu řeky.*

(pobýval na (břehu řeky) X v lesích na (břehu řeky))

Předpokládáme, že věty (III.5) a (III.5') by měly v PDT přiřazenu strukturu $N_g(N Pg)$, věty (III.6) a (III.6') strukturu $N_g(N_p Pg)$. Věť (III.7) odpovídá struktura $V_g(N Pg)$ a větám (III.8) a (III.8') struktura $V_g(N_p Pg)$. Všechny tudíž splňují podmínky pro ‚podezřelou‘ strukturu typu (c). (V tento okamžik nebereme v potaz skutečnost, že některé tyto věty odpovídají i ‚podezřelé‘ struktuře (a) – konkrétně věty (III.5), (III.5') a (III.6) a (III.6') –, a že tedy Pg může rozvíjet i nadřizené sloveso.) V závorce za větou uvádíme ve formě skladebných dvojic možné závislosti Pg (jako první vždy tu strukturu, o které předpokládáme, že by byla (primárně) zachycena v PDT).²³ □

Jak už jsme poznamenali, snažíme se postihnout zejména takové případy homonymie Pg, které se vyskytují často. Proto se omezujeme skutečně pouze na testování ‚sester‘ zkoumané předložky v závislostní reprezentaci věty. Porovnejme toto omezení s následujícím příkladem.

Příklad:

(III.9) *Povinnost pracovat na něm zanechala stopy.*

Předpokládáme, že v PDT bude primárně zachycena závislost *zanechat na něm*, tedy věta bude mít přiřazenu strukturu $V(N(V_{inf})PgN)$. Jako další potenciální skladebné dvojice budou (neadekvátně)²⁴ určeny dvojice *povinnost na něm* a *stopy na něm*. Skutečně možné druhé čtení věty – *pracovat na něm* – navrženo nebude, tato možnost není zahrnuta v navržených ‚podezřelých‘ strukturách. (Pokud bychom považovali toto druhé čtení za primárně zachycené v PDT, větná struktura $V(N(V_{inf}(Pg))N)$ by odpovídala ‚podezřelé‘ struktuře (a) a k problémům by nedošlo.) □

Některé případy této základní struktury jsou zajímavé zejména v kombinaci s koordinací. Zde se ovšem situace komplikuje (koordinace je opět zahrnuta pro rozmnožení hledaného materiálu). Koordinovaný může být kterýkoli uzel základní struktury, různé typy koordinace se můžou kombinovat.

Předložka řídí koordinovaná jména

Tato skutečnost nemá vliv na další zpracování, uveďme jako příklady modifikaci vět (III.5)-(III.8).

Příklady:

(III.5a) *Změna vychází z požadavků rozpočtu na příští rok i další léta.*

(požadavků na [(příští rok) i (další léta)] X rozpočtu na [(příští rok) i (další léta)])

(III.6a) *Navštívil kostely v Praze na Starém Městě a Malé Straně.*

(kostely na [(Starém Městě) a (Malé Straně)])

X v Praze na [(Starém Městě) a (Malé Straně)])

(III.7a) *Úřad hledá policisty na křižovatky i pochůzky městem. Aim*

²³ Konvence pro značkování PDT předpokládá rozhodování na základě obsahu.

²⁴ Důvody, proč není reflektován povrchový slovosled, viz výše.

(hledá na [křižovatky i (pochůzky městem)])

X policisty na [křižovatky i (pochůzky městem)])

(III.8a) *Bydlel v Praze na Hradčanech a Malé Straně.*

(bydlel na [Hradčanech a (Malé Straně)] X v Praze na [Hradčanech a (Malé Straně)])

□

Předložková skupina je členem koordinační skupiny

Pg se může vyskytovat v koordinaci (se dvěma stejnými nebo s různými předložkami), potom hledáme potenciální řídicí uzel celé koordinační skupiny – jméno nebo sloveso, jak ukazují následující příklady.

Příklady:

(III.5b) *Změna vychází z požadavků rozpočtu na příští rok i na další léta.*

(požadavků [(na (příští rok)) i (na (další léta))])

X rozpočtu [(na (příští rok)) i (na (další léta))]

(III.5b') *Bojují s nárůstem loupeží na našem území i ve světě.*

(nárůstem [(na (našem území)) i (ve světě)] X loupeží [(na (našem území)) i (ve světě)])

(III.6b) *Navštívil kostely v Praze na Starém Městě a na Malé Straně.*

(kostely [(na (Starém Městě)) a (na (Malé Straně))])

X v Praze [(na (Starém Městě)) a (na (Malé Straně))]

(III.7b) *Úřad hledá policisty na křižovatky i na pochůzky městem. Aim*

(hledá [(na křižovatky) i (na (pochůzky městem))])

X policisty [(na křižovatky) i (na (pochůzky městem))]

(III.8b) *Bydlel v Praze na Hradčanech a na Malé Straně.*

(bydlel [(na Hradčanech) a (na (Malé Straně))])

X v Praze [(na Hradčanech) a (na (Malé Straně))]

□

Předložková skupina rozvíjí celou koordinační skupinu

Předložková skupina může rozvíjet celou koordinační skupinu složenou z Ng/Adj. Potom je možné, že rozvíjí také větný člen, který je v závislostním stromu reprezentován sesterským uzlem zkoumané předložky.

Příklady:

(III.5c') *Bojují s nárůstem a stoupající brutalitou loupeží na našem území.*

(s [nárůstem a (stoupající brutalitou)] na (našem území) X loupeží na (našem území))

(III.6c) *Navštívil kostel a hřbitov v Praze na Vyšehradě.*

([kostel a hřbitov] na (Vyšehradě) X v Praze na Vyšehradě)

(III.7c) *Úřad hledá a nenachází policisty na křižovatky. Aim*

([hledá a nenachází] na křižovatky X policisty na křižovatky)

(III.8c') *Rád pobýval a toulal se v lesích na břehu řeky.*

([pobýval a toulal se] na (břehu řeky) X v lesích na (břehu řeky))

□

Předložková skupina rozvíjí N_g/V_g modifikované koordinovanými jmény

Předložková skupina, která rozvíjí N_g/V_g , může v jistých případech rozvíjet také celou koordinální skupinu řízenou tímto N_g/V_g , nebo jen některé její členy.

Příklady:

(III.5d') *Bojují s nárůstem loupeží a vražd na našem území.*

(nárůstem na (našem území) X [loupeží a vražd] na (našem území)

X [loupeží a (vražd na (našem území))])

(III.6d') *Miloval procházky v lesích a v hájích na břehu řeky.*

(procházky na (břehu řeky) X [v lesích a v hájích] na (břehu řeky)

X [v lesích a (v hájích na (břehu řeky))])

(III.7d) *Úřad hledá pohraničnický a policisty na hranici. Aim, Kde?*

(hledá na hranici X [pohraničnický a policisty] na hranici

X pohraničnický a [policisty na hranici])

(III.8d) *Bydlel na vesnici i ve městě na břehu řeky.*

(bydlel na (břehu řeky) X [(na vesnici) i (ve městě)] na (břehu řeky)

X [na vesnici i ((ve městě) na (břehu řeky))])

(III.8d') *Rád pobýval v lesích a v hájích na břehu řeky.*

(pobýval na (břehu řeky) X [v lesích a v hájích] na (břehu řeky)

X [v lesích a (v hájích na (břehu řeky))])

□

Předložková skupina rozvíjí N_g/V_g modifikované jinou Pg' s koordinovanými jmény

Předložková skupina, která rozvíjí N_g/V_g , může v jistých případech rozvíjet také celou koordinální skupinu, která je součástí jiné předložkové skupiny, nebo jen některé její členy. Tento případ se týká pouze vět (III.6') a (III.8').

Příklady:

(III.6e') *Miloval procházky v lesích a hájích na břehu řeky.*

(procházky na (břehu řeky) X (v [lesích a hájích]) na (břehu řeky)

X [v lesích a (hájích na (břehu řeky))])

(III.8e') *Rád pobýval v lesích a hájích na břehu řeky.*

(pobýval na (břehu řeky) X (v [lesích a hájích]) na (břehu řeky)

X [v lesích a (hájích na (břehu řeky))])

□

III.2. NÁVRH MAKER

Pro práci s korpusem je využíván program Graph Editor. Tento editor umožňuje – kromě standardních funkcí jako třeba prohlížení vět nebo vyhledávání uzlů s určitými hodnotami některých atributů – vytváření maker, pomocí nichž lze v korpusu vyhledávat vytipované struktury.

„Podezřelé“ struktury popsané v předchozím oddíle vedly k napsání dvou maker pro Graph Editor. Tato makra automaticky vyhledávají věty z PDT, jejichž reprezentace na analytické rovině (tedy závislostní strom) obsahuje jako podstrom některou z „podezřelých“ struktur. Makra jsou psána pro předložkovou skupinu $na+N_{Acc}$, resp. $na+N_{Loc}$, ale jednoduchým zadáním jiné předložky je lze použít pro vyhledávání „podezřelých“ vět obsahujících tuto předložku (testováno pro předložku *v*).

Makro typu A

Makro typu A pokrývá vytipované „podezřelé“ struktury (a) a (b). Najde předložkovou skupinu, najde její řídicí uzel a zkontroluje, zda jde o N_g , V_g ²⁵ nebo Adj. Poté prochází cestu od tohoto řídicího uzlu ke kořenu stromu (uzel s analytickou funkcí AuxS), resp. k hranici klauze (AuxC, AuxX, Apos)²⁶ a testuje, zda se na ní neobjeví sloveso nebo jméno, které by mohlo být řídicím členem předložkové skupiny (N/V). Přitom pokrývá i struktury obsahující koordinaci na všech úrovních (jak bylo popsáno v III.1.1.).

Makro se volá klávesovou zkratkou Shift+G pro aktuální podstrom, resp. Ctrl+Shift+G pro celý soubor. Makro označí zadanou předložku (atribut reserve1 – „Jsem PREP“), její řídicí uzel („Jsem govern“) a potenciální řídicí uzel („PREP lze pověsit sem“). Potenciální řídicí uzel lze najít podle hodnoty % atributu reserve2.

A-makro je zatím odzkoušeno na 25 souborech PDT (po 50 větách). Nejlepších výsledků dosahuje varianta makra pro N_g – jak ostatně vyplývá už z návrhu „podezřelých“ struktur. Bylo jím vybráno cca 160 „podezřelých“ vět.

Příklady:

(III.10a) *Adresa Úřadu pro řešení důsledků pobytu sovětských vojsk na území ČSFR je Slezská 9, 120 29 Praha 2.* (PDT, bl104js.fs #2)

(III.10b) *Noví majitelé domů potřebují peníze na jejich opravu.* (PDT, bl104js.fs #15)

(III.10c) *V přímém přenosu přivítají například i půlnoc na Staroměstském náměstí.* (PDT, bl110js.fs #18)

(III.10d) *Zostřování sociálního klimatu na Slovensku není žádným překvapením.* (PDT, bm207zua.fs #7)

(III.10e) *V Gale nepředpokládají, že by se trh práce na Prostějovsku a v okolí do dvou let zásadně změnil.* (PDT, bcd11eba #48)

²⁵ A-makro tedy pokrývá i strukturu s P_g rozvíjející sloveso; nejde tu však o typický výskyt homonymie P_g .

²⁶ Analytická funkce AuxC označuje podřadicí spojku, AuxX čárku (ne však hlavní koordinační), Apos znamená hlavní uzel apozice.

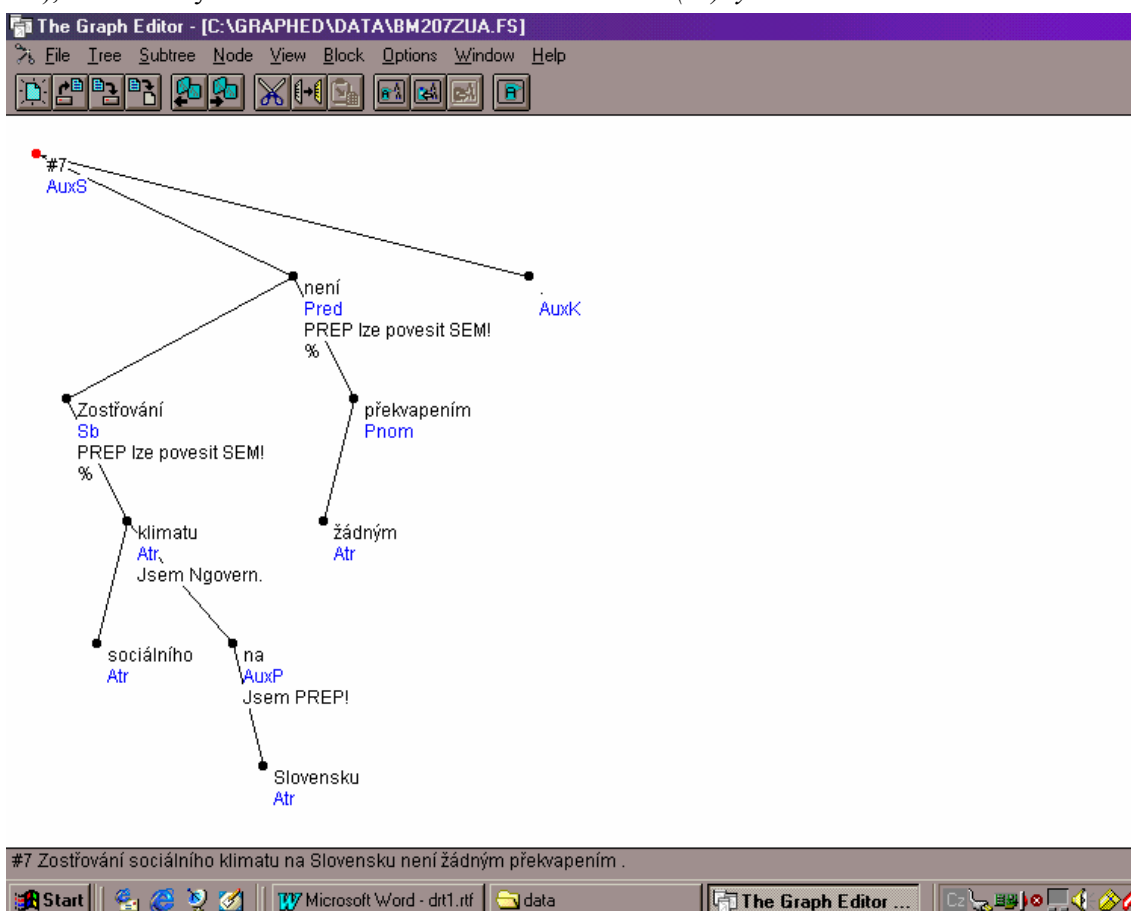
(III.10f) *Cestující v taxíku může kdykoli změnit svou instrukci o cíli jízdy.* (PDT, bm227zua.fs #25B)

Ve větě (III.10a) je Pg *na území* anotována jako atribut substantiva *pobyt* (analytická funkce Atr). A-makro navrhne jako potenciální řídicí uzly jména *důsledky, řešení, Úřad, adresa* a sloveso *být*.

Ve větě (III.10b) Pg *na opravu* rozvíjí substantivum *peníze*, analytická funkce AdvAtr zachycuje možnost považovat ji za rozvíjí slovesa *potřebovat* (se stejným významem), což odpovídá výsledku A-makra.

Obdobně v (III.10c) je Pg *na náměstí* označena jako přívlastek podstatného jména, resp. příslovečné určení (analytická funkce AtrAdv), A-makro nabízí sloveso *přivítat*.

Ve (III.10d) je Pg *na Slovensku* určeno jako člen rozvíjející jméno *klima* (analytická funkce Atr), A-makro vybere substantivum *zostřování* a sloveso *(ne)být*.



Obr.11: Výsledek vyhledávání pomocí makra A pro větu (III.10d)

(III.10e) zachycuje případ, kdy je N_g rozvíjeno koordinovanými předložkovými skupinami. V PDT mají koordinované Pg analytickou funkci AtrAdv_Co, rozvíjejí substantivum *práce*, resp. sloveso *změnit*. A-makro nabízí navíc substantivum *trh*.

(III.10f) je příkladem ‚podezřelé‘ věty vybrané pro předložku *v*: v PDT je primárně určena závislost *cestující v taxíku* (analytická funkce AtrAdv), A-makro vybere jako potenciální řídicí uzly sloveso *moci* (odpovídá ‚zdvojené‘ analytické funkci). □

Makro typu B

Makro typu B hledá ‚podezřelou‘ strukturu (c). Najde předložkovou skupinu, najde její řídicí uzel (N_g nebo V_g) a poté všechny sestry uzlu reprezentujícího předložku. Testuje, zda se mezi nimi neobjeví sloveso (V), jméno v prostém (N) nebo předložkovém pádu, tj. jiná předložková skupina (N_p) – takové větné členy jsou potenciálně rozvíjeny zkoumanou předložkovou skupinou. Strom je prohledáván pouze do hloubky 1 (nepočítá-li se koordinace), (tj. skutečně se testují jen sesterské uzly zadané předložky), resp. do hloubky 2 u případných dalších předložkových skupin nebo vyskytnou-li se ve struktuře koordinované uzly, případně do hloubky 3, kombinuje-li se koordinace se jménem / jmény v předložkovém pádu.

Makro se volá klávesou G pro aktuální podstrom, resp. Ctrl+G pro celý soubor. Stejně jako makro A i makro B označí zadanou předložku, její řídicí uzel a potenciální řídicí uzel (reserve1). Potenciální řídicí uzel lze najít podle hodnoty \$ atributu reserve2.

B-makro je zatím odzkoušeno na stejném vzorku – cca 1250 vět – jako A-makro. Podstatných výsledků dosahuje makro pro V_g , což opět vyplývá už z návrhu ‚podezřelých‘ struktur. Získali jsme cca 150 vět obsahujících vytipovaný podstrom. Oba typy ‚podezřelých‘ podstromů se často kombinují v jedné větě.

Příklady:

(III.11a) *Zápas začal náš tým se třemi muži v přední linii, **Kuku na hrotu doplňovali Hapal s Dubovským.*** (PDT, bmc26aba.fs #25)

(III.11b) *V letech 1991-1993 jsem **absolvovala kurzy řízení a marketingu na Collège Bart** v kanadském Québecu.* (PDT, bce17zua.fs #8)

(III.11c) ***Stát neposkytne na nákup** bytů žádné slevy, ani žádnou jinou finanční pomoc (původně se zvažovalo, že do ceny bytu se promítne řada faktorů, zejména však amortizace).* (PDT, bcb01aba.fs #16)

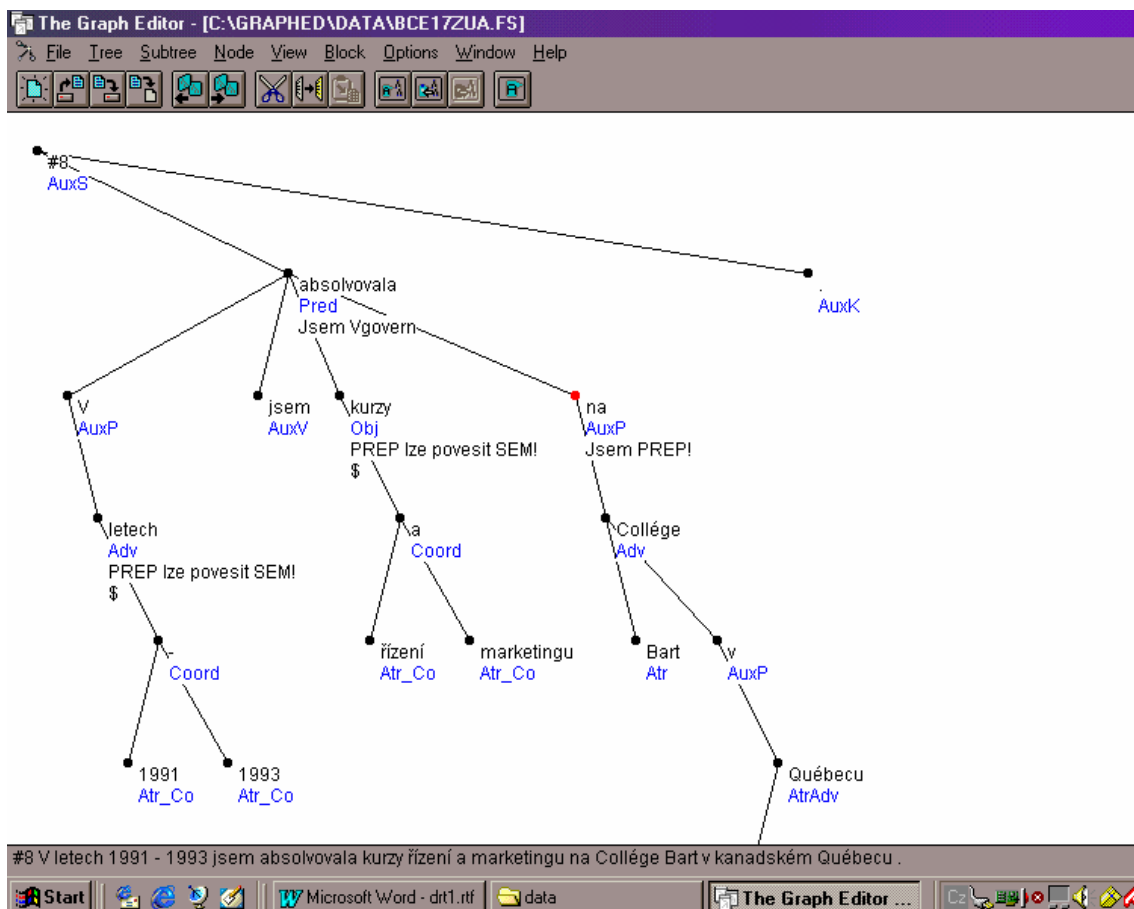
(III.11d) ***Cenu J. Seiferta, ..., zřídila Nadace Charty 77 ve Stockholmu v lednu 1986 k uctění** památky prvního čs. nositele Nobelovy ceny za literaturu.* (PDT, bmb09eba.fs #19)

Ve větě (III.11a) je P_g na hrotu analyzována jako adverbiale ke slovesu *doplňovat* (analytická funkce Adv). B-makro navrhne jako potenciální řídicí členy jména *Kuka* (adekvátně) a *Hapal* (neadekvátně).

Modifikaci věty (III.11b) jsme použili jako větu (I.5b) pro ilustraci principu ‚last attachment‘. V PDT je P_g na *Collège (Bart)* anotována jako adverbiale ke slovesu *absolvovat* (analytická funkce Adv). B-makro vybere větné členy *kurzy* (adekvátně) a *letech* (nevhodně).

(III.11c) ukazuje situaci, kdy P_g rozvíjí sloveso modifikované koordinovanými jmény. P_g na *nákup* (analytická funkce Obj) přísluší ke slovesu *neposkytnout*. B-makro nabídne jako možné řídicí členy substantiva *slevy* (OK), *stát a pomoc* (KO).

Poslední příklad zachycuje výsledek B-makra pro předložku *v*. P_g ve *Stockholmu i v lednu (1986)* (analytické funkce Adv) rozvíjejí sloveso *zřídít*, B-makro nabízí podstatná jména *Cena* (KO), *Nadace* (OK), *uctění* (KO) a dále vždy jméno tvořící druhou P_g . (Substantivum *Charty (77)* by bylo vybráno A-makrem.)



Obr.12: Výsledek vyhledávání pomocí makra B pro větu (III.11b) □

Na větách v předcházejících příkladech je vidět, že některé uzly jsou vybrány jako potenciální řídicí uzly testované předložkové skupiny neadekvátně (týká se všech tří ‚podezřelých‘ struktur). To je způsobeno rozhodnutím **nebrat při návrhu maker v úvahu povrchový slovosled** (viz výše), které vede k nárůstu počtu vět vybraných jako ‚podezřelé‘ vzhledem k homonymii Pg (ovšem vzhledem k účelu maker – pomoci při získání co možná největšího počtu vět obsahujících homonymní Pg k dalšímu lingvistickému zpracování – toto není příliš na závadu). Na druhou stranu takové rozhodnutí umožňuje zachytit velmi důležitý jev: odhlédneme-li od povrchového slovosledu, získáme pomocí navržených maker věty obsahující neprojektivní²⁷ výskyty doplnění realizované pomocí předložkových skupin. To se týká především valenčních doplnění substantiv odvozených od sloves a slovesně jmenných vyjádření.²⁸ Získáme tedy příležitost popsat poměrně častý zdroj neprojektivních konstrukcí, kdy je substantivum rozvíjeno Pg v pozici, která není typická pro nominální doplnění. Jako příklad takového postavení Pg uveďme modifikaci věty (III.10b).

Příklad:

(III.10b') *Na opravu domů noví majitelé potřebují peníze.* (PDT, bl104js.fs #15, pozměněno)

²⁷Pojem neprojektivity je vysvětlen v kapitole IV.1.1.

²⁸Věnujeme jim zejména kapitoly IV.2.2. a IV.2. 3.

Nominální doplnění obvykle následují za svým řídícím jménem. To ale neplatí pro slovesně jmenná vyjádření – Pg *na opravu (domů)* intuitivně rozvíjí jak sloveso *potřebovat*, tak podstatné jméno *peníze*. Ačkoli Pg předchází substantivum, lze ji chápat jako jeho valenční doplnění (více viz kapitolu IV.2.3. o slovesně jmenných výrazech). □

Možnost zachytit neprojektivní rozvití substantiva předložkovou skupinou bereme jako dostatečný důvod pro odhlédnutí od omezení daných povrchoým slovosledem.

Shrnutí

Pražský závislostní korpus, přesněji anotace na analytické rovině, slouží jako základní zdroj materiálu pro zkoumání homonymie předložkových skupin. Pro vyhledávání ‚podezřelých‘ vět byla navržena dvě makra reflektující tři ‚podezřelé‘ syntaktické struktury.

Výběr syntakticky anotovaného korpusu (přestože v další práci už struktury vět nejsou využívány) umožňuje získávat věty nespňující slovosledná omezení typická pro Pg, včetně vět s neprojektivním doplněním ve formě Pg. Popis takových jazykových konstrukcí tvoří jedno z témat čtvrté části této práce.

Sekundárním výsledkem této části je ilustrace možnosti vyhledávání relativně komplikovaných (povrchových) struktur v PDT pomocí maker.

ČÁST IV.

LINGVISTICKÁ KRITÉRIA

Jak už jsme poukázali dříve, adekvátní syntaktická analýza věty, a zejména jejích částí vyjádřených předložkovými skupinami, je obvykle založena nejen na čistě syntaktických pravidlech, často zahrnuje i širší, věcné, jazykově explicitně nevyjádřené znalosti. V této části se pokusíme ukázat, jaké typy informace lze brát v úvahu při návrhu kritérií umožňujících (lingvisticky podložené) automatické určování všech členů české věty, které mohou být rozvíjeny zkoumanou předložkovou skupinou. Zabýváme se tedy možností automatického určení morfematické homonymie předložkových skupin.

Přiblížíme zde metodu redukční analýzy, která tvoří metodologický rámec našeho zkoumání.

Při formulaci kritérií pro určování možných závislostí předložkových skupin postupujeme metodou jejich postupného zjemňování. Nejprve zkoumáme předložkové skupiny v prototypických, projektivních strukturách (zabýváme se především jednoduchými slovoslednými pravidly a pravidly využívajícími valenční informace). V následující fázi tuto základní analýzu zjemňujeme, rozvolňujeme slovosledná omezení, připouštíme neprojektivní závislosti. Využíváme přitom bohatší slovníkové informace (např. údajů o víceslovných výrazech, případně též sémantických údajů), zpřesňujeme pravidla. V posledním kroku pracujeme s komplexními pravidly, na jejichž základě jsou vyloučeny nepřijatelné neprojektivní tvary závislostních stromů získaných v druhé fázi. Výsledné závislosti předložkových struktur jsou vyhodnoceny podle typu pravidla, které vedlo k jejich vytvoření.

Kritéria pro určení možných závislostí Pg jsou rozděleny do čtyř skupin. Prvním typem kritérií jsou kritéria založená na slovosledných pozorováních, na jejichž základě jsou definovány „jednoznačné“ a „podezřelé“ slovosledné vzorce vzhledem k homonymii Pg. Všímáme si zejména Pg v běžných pozicích, ale i konstrukcí, ve kterých Pg tvoří neprojektivní struktury (oddíl IV.1). Základním zdrojem informací pro disambiguaci předložkových skupin jsou valenční rámce sloves, podstatných jmen a přídavných jmen – a to jednak standardní valenční rámce (jak jsou popisovány v teoretických studiích), jednak informace o kvazivalenčních doplněních (jde o „běžně užívaná“ doplnění). V souvislosti s valencí jednotlivých slovních druhů jsou zkoumány i tzv. slovesně jmenné výrazy (oddíl IV.2.). V oddíle IV.3. se zabýváme kritérii, která využívají sémantické klasifikace jednotlivých lexikálních položek, především sémantických rysů podstatných jmen, sloves a předložek. Dále ilustrujeme omezení daná pravidly, která reflektují vzájemné slovosledné pozice (valenčních i

volných) doplnění lexikální jednotky, na jejichž základě lze vyloučit některé nepřipustné neprojektivní tvary závislostního stromu (oddíl IV.4.).

V části V., věnované uplatnění navržených kritérií, navrhneme možné uspořádání pravidel a způsob jejich použití pro určení větných členů, které mohou být rozvíjeny předložkovou skupinou, v mechanismu analýzy založené na redukci. Stanovíme spolehlivost získaných závislostí.

Předpoklady

Shrňme zde nutné předpoklady, na kterých stavíme kritéria pro automatické určení možných závislostí předložkových skupin a pro odhalení jejich případné homonymie. Týkají se jednak slovníku, jednak předzpracování vstupního textu. Omezujeme se přitom na zpracování jedné klauze věty.

Slovník:

- Předpokládáme existenci bohatého českého slovníku, který by obsahoval jednak podrobné morfologické údaje (včetně informací o odvozování), jednak údaje o valenčních rámcích sloves, podstatných jmen a adjektiv, a to jak jejich standardních valenčních rámců (tj. jejich aktantů a obligatorních volných doplnění), tak i informací o ‚kvazivalenčních‘ doplněních²⁹ zachycujících obvyklé užití slova.

- Dále předpokládáme, že slovník bude obsahovat informace o tzv. slovesně jmenných spojeních a jejich valenci – tím míníme ustálená spojení (primárně) slovesa a jména, která jsou rozvíjena jako celek (viz kapitola IV.2.3., ve které uvádíme i množství příkladů).

- Slovník zachycuje víceslovné výrazy jako nevlastní předložky, frazémy a idiomy.

- Fakultativně – pokud by byly takové informace ve slovníku zachyceny – je možno pro ohodnocení jednotlivých získaných struktur s předložkovými skupinami využít i sémantických rysů substantiv a sloves, spolu s možnými sémantickými funkcemi předložek. □

Vymezení slovesně jmenných výrazů ovšem zdaleka není triviální a vyžaduje další zkoumání. Proto navrhuje dva přístupy k vyšetřování závislostí předložkových skupin. Základní přístup využívá slovníkové informace o slovesně jmenných výrazech a jejich valenci; tento předpoklad lze oslabit a pracovat pouze s valenci jednotlivých slov.

Na tomto místě je nutno přiznat, že tak obsáhlý slovník zatím není pro češtinu k dispozici. Nicméně postupně vzniká pro účely jednotlivých projektů zabývajících se automatickým zpracováním češtiny. Návrh formátu jednotlivých druhů lexikálních jednotek lze nalézt například u Skoumalové (1994) a (1997). Tuto práci lze chápat také jako příspěvek k diskusi o tom, jak bohatou informaci je vhodné ve slovníku uchovávat.

Předzpracování vstupního textu:

- Požadujeme, aby vstupní text byl morfologicky zpracován, tedy aby k jednotlivým slovům analyzované věty byla přiřazena jejich úplná, nedisambiguovaná morfologická analýza. Navrhovaná kritéria využívají lemata jednotlivých slovních forem a postupně zpracovávají všechny možné kombinace určených morfologických kategorií (sémat).

- Navíc předpokládáme, že ve vstupním textu jsou provedeny některé úpravy lokálního charakteru – jsou složena ustálená spojení (frazémy, idiomy) a některá další slovní spojení (zejména nevlastní předložky – *bez ohledu na, na základě* ...), jsou identifikovány a redukovány

²⁹ Pojem kvazivalenčního doplnění přiblížíme v oddíle IV.2.

analytické tvary sloves, případně jsou vyhledána slovesně jmenná spojení (typu *mít nárok/šanci/právo, podat odvolání/ stížnost, ...*). Dále jsou redukována zdůrazňovací slova (částice *bohudík, prý, snad, také, totiž, však, ...*). Charakteristická je ovšem skutečnost, že v určitých fázích analýzy se vracíme k původnímu znění věty (viz formální kritéria, uplatnění navržených kritérií).

- Jsou identifikovány a složeny dvojice závorek, dvojice slov uvozující souřadná spojení (*jak – tak, jednak – jednak, ...*), jsou určeny některé apozice (*resp., totiž, tj., a to ...*).
- Pracujeme s předpokladem, že jsou určeny hranice klauze. □

Metoda redukční analýzy (RA)

Obecným schématem, které tvoří metodologický rámec našeho přístupu k analýze závislosti českých předložkových skupin – k jejich automatické disambiguaci, případně k určení homonymie Pg – je metoda redukční analýzy, kterou ilustrují např. Jančar et al (1999).

Charakteristika metody redukční analýzy (RA):³⁰

Redukční analýza sestává z postupného zjednodušování rozvitě věty tak, aby při jednotlivých krocích zůstala zachována syntaktická (ale právě jen syntaktická, o odchylkách viz níže) správnost věty. Redukční analýza pracuje s povrchoвым zápisem věty, její základní operací je operace vypuštění:

- Každý krok zjednodušování rozvitě věty je realizován **vypuštěním** jednoho větného členu a případně přepsáním dalších slov (v zájmu zachování syntaktické správnosti věty). K vypuštění větného členu dochází na základě lingvisticky podložených pravidel opřených o závislostní syntax.

Proces postupného zjednodušování věty je nedeterministický, v každém kroku redukce lze vypustit jakýkoliv ze (syntakticky) vzájemně nezávislých členů věty. □

Na redukční analýzu lze pohlížet jako na model přibližující (intuitivní) postup člověka při rozboru věty, na formální popis takového postupu.³¹ Umožňuje formulovat základní kritéria pro zpracování přirozeného jazyka, které lze dále uplatnit při tvorbě konkrétních gramatik.

Aplikace RA na automatické zpracování Pg

Analýza věty přirozeného jazyka (češtiny) založená na metodě redukční analýzy (vztažené k povrchové syntaxi) tedy postupuje v krocích, v každém kroku je **redukován (vypuštěn) jeden větný člen**. Z charakteru redukční analýzy vyplývá skutečnost, že všechna závislá slova jsou zpracována dříve, než je redukováno jejich slovo řídící – v opačném případě, pokud by byl dříve vypuštěn řídící člen, nebylo by možno aplikovat pravidla, která by vedla k redukci členu závislého (viz níže). Otázku určení směru závislostní relace,³² tedy rozlišení

³⁰ Pracujeme zde na rovině povrchové syntaxe. Velmi zajímavá je i možnost uplatnit metodu redukční analýzy na tektogramatické rovině, kde její aplikace slibuje nový pohled na problematiku valenčních rámců.

³¹ Zdůrazňujeme, že nám nejde o skutečný psychologický model simulující porozumění člověka přirozenému jazyku. Popisujeme pouze možný pohled na postupné zjednodušování věty, tak jak se je učí žáci ve školách.

³² Stanovením syntaktických kritérií pro rozlišení řídícího a závislého členu závislostní relace se zabývá také např. I. Mel'čuk (1988), kde pracuje s termíny pasivní syntaktické valence, nebo distribuce.

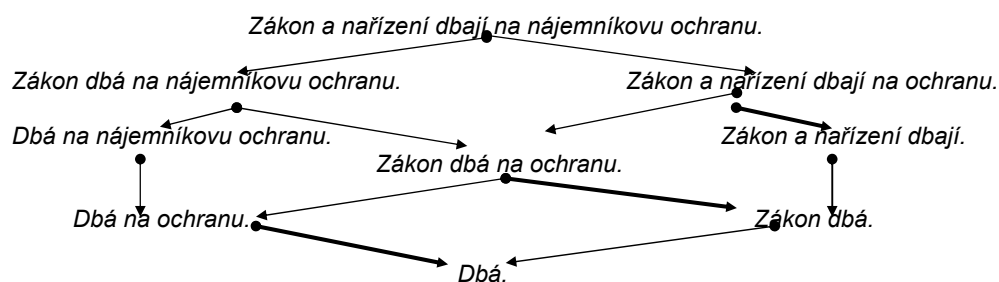
řídícího a závislého členu této relace, zodpovídá Sgall (1998), který navrhuje využít principu analogie na úrovni slovních druhů.³³

Možnosti **přepisování** slov během analýzy se omezují na změny v morfologických kategoriích čísla, rodu, příp. pádu a to pouze tak, aby byla při vypouštění zachována existující gramatická shoda (např. shoda podmět – přísudek nebo změny při zpracování číslovek).

Mechanismus postupného vypouštění během analýzy založené na principu redukce a možnost přepisování ilustruje následující příklad:

Příklad:

(IV.1) *Zákon a nařízení dbají na nájemníkovu ochranu.* (PDT, bl101js.fs #18, pozměněno)



Obr.13: Větve analýzy věty (IV.1) založené na redukci

Na Obr. 13 jsou zachyceny větve, na které se „rozpadne“ analýza věty založená na redukci. Tyto větve vedou k úspěšné adverbální analýze Pg *dbát na (nájemníkovu) ochranu* (jsou přitom využita pravidla popsaná v následujících kapitolách; větve, ve kterých není zachována syntaktická správnost věty, ve schématu zachyceny nejsou).

Při postupné redukci větných členů může být v prvním kroku redukováno slovo *nájemníkovu*, původní věta je zjednodušena na *Zákon a nařízení dbají na ochranu*. Druhý krok analýzy může sestávat ze zjednodušení koordinace – vypuštění *a nařízení* (tím dojde ke změně čísla podmětu) a přepsání tvaru slovesa *dbají* na *dbá* tak, aby (morfologické) číslo slovesa bylo ve shodě se změněným číslem podmětu. (Obdobně i v dalších dvou větvích analýzy, kde dochází ke zjednodušení koordinace.)

Pg *na (nájemníkovu) ochranu* tedy může být zpracovávána ve třech větvích RA – kroky, ve kterých je Pg redukována, jsou ve schématu označeny tučnou šipkou. □

V každém kroku RA lze vypustit jakékoliv ze vzájemně (syntakticky) nezávislých slov rozebírané věty, celý proces analýzy je **nedeterministický**. Zpracování věty se obvykle rozpadá na několik větví. Jediným kritériem pro výběr větných členů, které mohou být v daném kroku analýzy redukovány, je zachování (relativní) syntaktické správnosti věty (viz poznámka 33). K analýze možných závislostí předložkové skupiny a k její redukci tedy obvykle dochází v několika větvích RA.

Člověk při analýze vět pracuje s významem, a tedy nepřipustí analýzu, která by změnila smysl původní věty (resp. vnášela smysl, který původní věta nemá). Pro automatickou disambiguaci Pg, případně pro určení homonymie Pg oslabujeme předpoklad **zachování správnosti** během redukce. Analýzu věty pokládáme za správnou (a jednotlivé redukční kroky za oprávněné), pokud je úspěšná, tj. pokud vede ke zjednodušení zkoumané rozvíte věty na větu o jednom členu (predikátu u slovesných vět). Takové oslabení předpokladu ovšem může vést

³³ Zjednodušeně, pro ilustraci takového postupu v případě aktoru: protože existují slovesa, která nevyžadují aktor (*prší*), je vhodné považovat aktor vždy za větný člen závislý na slovese. Obdobně je i nevypustitelný objekt (např. u slovesa *potkat*) v redukční proceduře vypouštěn (neboť existují slovesa, která obligatorní objekt nemají).

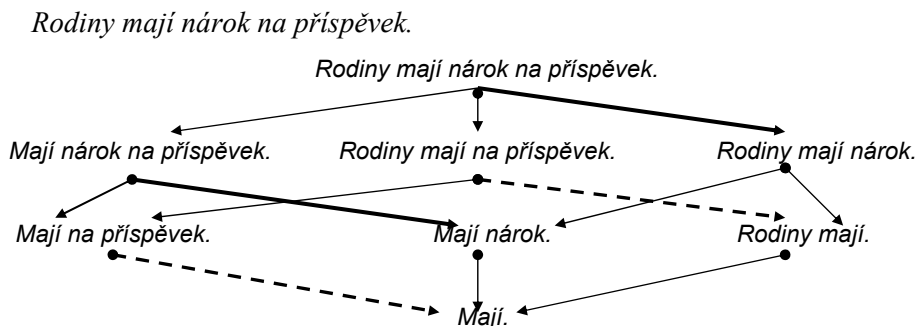
k potvrzení neadekvátní analýzy věty – může se totiž stát, že během RA je vytvořena syntakticky správná věta, která má jiný význam než věta původní. Tuto možnost ilustruje příklad (IV.2). V současné době ovšem nemáme dostatek podkladů pro formulaci kritérií, která by takovému potvrzení neadekvátní analýzy zabránila (pokud je vůbec možné taková kritéria najít).

Rozbor věty (IV.2), založený na principu vypouštění, ukazuje přínos **klasifikace** jednotlivých typů **pravidel** pro automatické řešení homonymie Pg.

Příklad:

(IV.2) *Tyto rodiny mají nárok na státní vyrovnávací příspěvek.* (na základě PDT, bl103zu.fs #42)

Pro lepší přehlednost předpokládejme, že v prvních krocích analýzy byly vypuštěny vzájemně nezávislé shodné přívlasky *tyto* (rozvíjí substantivum *rodiny*), *státní* a *vyrovnávací* (,zavěšeno‘ na podstatném jménu *příspěvek*). Rozebíráme tedy zjednodušenou větu:



Obr.14: Větvě analýzy věty (IV.2) založené na redukci

Analýza věty se rozpadá do několika větví, na zpracování Pg má vliv skutečnost, jak bohaté informace jsou k dispozici.

Pokud pracujeme se slovesně jmennými výrazy, může být Pg redukována ve dvou větvích RA (viz schéma na Obr.14, silné šipky) – Pg *na příspěvek* je určena (na základě pravidel pro řešení homonymie Pg a jejich ohodnocení) jako valenční doplnění substantiva *nárok* (s vysokou preferencí).

Pokud se slovesně jmennými výrazy nepracujeme, může dojít k redukci Pg v dalších dvou krocích analýzy (čárkované šipky ve schématu) – dostaneme navíc závislost *mít na příspěvek*, s Pg určenou jako doplnění slovesa *mít* (s nižší preferencí).

Ve všech případech je splněna podmínka zachování správnosti – obě redukováné věty (*(Rodiny) mají nárok.* / *(Rodiny) mají na příspěvek.*) jsou syntakticky správné a analýza vede k predikátu *Mají*. Ovšem analýza *mít na příspěvek* má jiný význam než původní věta. Její význam je svázán se slovesně jmenným výrazem *mít peníze/prostředky ... (na něco)*, jehož jmenná část může být vypuštěna. Správnost tohoto rozboru je tedy dána lexikálním obsazením.

Náš výzkum směřuje ke stanovení preferencí pro jednotlivé rozboru věty – ohodnocení jednotlivých typů pravidel umožňuje ,vzdvihnout‘ některé analýzy a ,upozadit‘ jiné.

V příkladu (IV.2) má adekvátní analýza Pg jako valenčního doplnění substantiva vyšší ohodnocení než zavádějící analýza Pg jako volného, příp. kvazivalenčního doplnění slovesa. (Podrobněji se k tomuto příkladu vrátíme v kapitole věnované slovesně jmenným výrazům, IV.2.3.) □

Naším cílem je tedy definovat **lingvisticky podložená pravidla** pro automatické zpracování předložkových skupin, případně pro určení jejich homonymie. Předpokládáme, že tato pravidla bude možno aplikovat ve všech větvích analýzy, v každém jejich kroku, ve kterém může být zpracována předložková skupina. Mohou se přitom vztahovat jednak k aktuálnímu stavu věty, tj. ke větě už (částečně) zjednodušené, jednak k originálnímu znění věty. Na základě těchto pravidel budou určeny větné členy, které mohou být rozvíjeny zkoumanou Pg, a Pg bude redukována. Pokud daná větev RA povede k úspěšné analýze (k větě o jednom členu), budeme pokládat získané připojení, příp. získaná připojení za adekvátní.

Významnou možností je **ohodnocení lingvistických kritérií** podle stupně spolehlivosti získaných závislostí (jak ukazoval rozbor věty (IV.2)). To se týká jednak jednotlivých typů kritérií, ale i pravidel jednoho typu (jde zejména o klasifikaci jednotlivých typů valenčních doplnění, viz IV.2.5.). Výsledné analýzy věty budou uspořádány podle typu pravidel, která byla použita pro připojení Pg – v této fázi práce předpokládáme částečně automatické vyhodnocování získaných závislostních stromů, pro které postačuje prosté stanovení pořadí jednotlivých typů pravidel (viz část V.).

Přirozeně, čím více informací je možno využít, tím spolehlivější je i získaná analýza. Nicméně příklad (IV.2) ukazuje, že automatická procedura může poskytovat dostatečně spolehlivé výsledky i při neúplné vstupní informaci, což je umožněno ohodnocením výsledných struktur podle jejich spolehlivosti. Tento fakt je velmi důležitý vzhledem ke skutečnosti, že vstupní informace pro jakoukoli automatickou analýzu přirozeného jazyka jsou nutně neúplné – znalostí, na jejichž základě člověk rozumí větě přirozeného jazyka, zahrnují nejen jevy náležející morfologii a syntaxi, ale i širokou škálu jevů povahy sémantické i pragmatické, a jako takové nemohou být plně pokryty žádným automatickým systémem.

Při zkoumání kritérií pro automatické zpracování předložkových skupin a pro zachycení jejich případné homonymie pracujeme s principem **přesouvání** slov ve větě. Metodu redukční analýzy uplatňujeme jednak na podkladových větách nalezených v PDT, jednak na jejich slovosledných permutacích (tj. na větách se změněným povrchovým slovosledem).

Princip přesouvání slov ve větě (a následný rozbor takto získaných vět založený na redukční analýze) je využíván zejména pro stanovení slovosledných omezení relevantních pro adekvátní výběr větných členů, které mohou být rozvíjeny předložkovou skupinou. Významně se uplatňuje i při vyhodnocování nalezených závislostí, při určování jejich preferencí. (V kapitole IV.2.5. se vrátíme k modifikacím věty (IV.1) a k jejím slovosledným variantám – na příkladu (IV.31a) ukážeme vliv přesouvání na vyhodnocení navržených závislostí *zákon na (nájemníkovu) ochranu a dbát na (nájemníkovu) ochranu.*)

K principům obecné metody redukční analýzy na rovině povrchové syntaxe a k jejímu vlivu na formulaci pravidel pro analýzu možných závislostí Pg se budeme vracet u těch typů zkoumaných kritérií, ve kterých se odráží zvolený přístup – výběr RA ovlivňuje především stanovení a využití tzv. slovosledných vzorců, kterým je věnován oddíl IV.1.

Závěr:

Cílem naší práce je vytvoření modelu automatické syntaktické analýzy předložkových skupin, který respektuje základní principy redukční analýzy. Stanovíme posloupnost kroků, ve kterých jsou zachycena lingvistická kritéria pro zpracování předložkových skupin a pro zachycení jejich případné homonymie. □

IV.1. SLOVOSLED

Povrchový slovosled věty je hlavním čistě syntaktickým prostředkem pro určování větných členů, které mohou být rozvíjeny zkoumanou předložkovou skupinou (na kterých P_g závisí; ve stromové reprezentaci povrchové struktury věty jde o možné řídicí uzly uzlu reprezentujícího P_g). Pro některé úlohy (jako je např. rozpoznávání, tj. kontrola gramatické správnosti věty, grammar checking) se při zpracování předložkových skupin obvykle lze omezit na pravidla daná povrchovým slovosledem (Kuboň et al (1997)). Pro úplnou analýzu češtiny, jazyka, který se vyznačuje značnou volností slovosledu, a zejména při snaze o zachycení homonymie předložkových skupin, je potřeba slovosledná kritéria uplatňovat společně s dalšími typy kritérií, zejména s pravidly využívajícími konceptu valence.

Při rozboru vět z PDT obsahujících potenciálně homonymní P_g (získaných pomocí maker A a B popsaných v části III.) se potvrdilo, že P_g se vyskytují především v projektivních konstrukcích, nicméně mohou tvořit také konstrukce neprojektivní.

Rozdělíme proto zkoumání povrchového slovosledu a jeho vlivu na jednoznačnost, případně víceznačnost předložkové skupiny na dvě fáze. V první fázi se omezíme na základní, projektivní jazykové konstrukce, ve druhé fázi se budeme zabývat i konstrukcemi méně typickými a neprojektivními.

IV.1.1. Základní (prototypické), projektivní konstrukce s P_g

Projektivní jazykovou konstrukcí zde (neformálně řečeno) myslíme takovou část věty, jejíž závislostní reprezentace na povrchové rovině má vlastnosti projektivního stromu.

Charakteristika projektivního stromu:

Závislostní strom je **projektivní** právě tehdy, když pro každou dvojici jeho uzlů $[n_g, n_d]$, kde n_d bezprostředně závisí na n_g , platí podmínka projektivnosti.

Podmínka projektivnosti je splněna právě tehdy, pokud všechny uzly stromu n_i , které leží (v uspořádání daném povrchovým slovosledem) mezi n_g a n_d , závisí (přímo nebo zprostředkovaně) na řídicím uzlu n_g .

Není-li pro nějakou dvojici uzlů $[n_g, n_d]$ splněna podmínka projektivnosti, je **závislostní strom neprojektivní**. □

Problémy související s projektivitou jsou studovány od šedesátých let (viz např. Marcus (1965), Nebeský (1972), Dikovský a Modina (2000)). Formální aparát umožňující klasifikaci neprojektivních vět, míru jejich neprojektivity, lze nalézt např. u Holana et al (1998). Neprojektivitou českých vět se zabývala Uhlířová (1972), konstrukci řady českých vět s vysokou mírou neprojektivity uvádějí Holan et al (2000).

Podívejme se v prvé řadě na P_g v prototypických konstrukcích. Metodologický rámec redukční analýzy, spolu s principem přesouvání, je vhodný k rozboru vět obsahujících potenciálně homonymní předložkové skupiny. Formulujeme pět slovosledných „pozorování“, která vedou ke stanovení základních slovosledných vzorců (WOP, word order pattern)

vzhledem k homonymii Pg – WOP s (primárně) jediným čtením, základních ‚podezřelých‘ WOP a WOP s preferovaným čtením.

Pozorování:

- (1) Předložková skupina chápaná jako doplnění substantiva rozvíjí (prototypicky) pouze jména, která v povrchové reprezentaci věty této Pg předcházejí. To znamená, že řídicí jméno musí předcházet svému atributivnímu doplnění realizovanému předložkovou skupinou.
- (2) Jakákoli slovesná forma slouží jako ‚zarážka‘ při hledání jména, které může být chápáno jako větný člen rozvíjený zkoumanou Pg. Jinak řečeno, předložková skupina primárně nerozvíjí jméno, které je od ní v povrchové realizaci věty odděleno libovolnou formou slovesa, tj. žádné sloveso (ani žádná část analytického tvaru slovesa) se nemůže objevit mezi substantivem a jeho (neshodným) přívlastkem realizovaným předložkovou skupinou.
- (3) Jakákoli slovesná forma slouží také jako ‚zarážka‘ pro Pg analyzovanou jako doplnění adjektiva – předložková skupina nemůže rozvíjet žádné adjektivum, které je od ní v povrchové realizaci věty odděleno slovesem (jakoukoli jeho formou, jakoukoli částí analytického tvaru slovesa).
- (4) Předložková skupina, která rozvíjí adjektivum s funkcí shodného přívlastku, může v povrchové reprezentaci věty tomuto adjektivu předcházet i za ním následovat, nemůže však od něj být oddělena jménem, které je tímto adjektivem modifikováno – netvoří obvykle neprojektivní konstrukci.
- (5) Shodný přívlastek realizovaný shodným adjektivem, které samo není dále rozvíjeno, obvykle v češtině předchází své řídicí jméno (s některými regulárními výjimkami jako je např. specifická kontrastivní pozice adjektiva (*jedna kapsa prázdná, druhá vysypaná*) nebo vědecké názvosloví). Tudíž fakt, že za jménem následuje shodné adjektivum, napovídá, že se dá očekávat, že samo toto adjektivum bude rozvíjeno – často následující předložkovou skupinou.

□

K pozorování (2) a (3) poznamenejme, že se v této práci nebudeme zabývat rozбором analytických tvarů sloves (předpokládáme, že jejich ‚spojení‘ je realizováno modulem morfologické analýzy). Způsob ‚zavěšování‘ volných doplnění v povrchové analýze do jisté míry závisí na úzu přijatém pro konkrétní úlohu, pro nás je podstatná pouze pozice jednotlivých slovesných tvarů v povrchovém vyjádření věty. Symbol **V** v následujících vzorcích označuje libovolnou část analytického tvaru slovesa, resp. jeho ‚stopu‘ po redukci na základě RA. (V podkladové struktuře věty je sloveso reprezentováno jediným uzlem.)

Na základě uvedených pozorování vymežeme základní slovosledné vzorce vzhledem k homonymii Pg, platné pro prototypické větné konstrukce.

Definice (základní ‚jednoznačné‘ WOP):

Definujeme tři **základní ‚jednoznačné‘ slovosledné vzorce vzhledem k homonymii Pg**, které se týkají povrchové reprezentace věty a platí pro prototypické struktury:

- sloveso, resp. jméno v neslovesné větě stojící na začátku klauze, je bezprostředně následováno předložkovou skupinou; označme tento typ **VPg**, resp. **NPg** (takové sloveso, resp. jméno, je rozvíjeno zkoumanou Pg);
- předložková skupina je na začátku klauze, typ **PgV** (Pg prototypicky rozvíjí sloveso);
- adjektivum a předložková skupina jsou odděleny slovesem, typy **AVPg** a **PgVA** (Pg rozvíjí sloveso).

Každé syntaktické jméno může být buď v prostém pádu (tj. bez předložky) nebo v předložkovém pádu.

Je-li naplněn jeden ze základních ‚jednoznačných‘ WOP, je na jeho základě vytvořena jediná závislost Pg (s ‚váhou‘ volného doplnění). □

Definice (základní ‚podezřelé‘ WOP):

Definujeme čtyři **základní ‚podezřelé‘ slovosledné vzorce vzhledem k homonymii Pg**, které se týkají povrchové reprezentace věty a platí pro prototypické struktury:

- řetězec jmen na začátku klauze je bezprostředně následován předložkovou skupinou, typ **NNPg** (všechna jména jsou potenciálně rozvíjena Pg);
- řetězec sestávající ze slovesa, jména a předložkové skupiny (v tomto pořadí), tedy typ **VNPG** (Pg potenciálně syntakticky rozvíjí sloveso i jméno);
- předložková skupina stojí mezi jménem a slovesem (v tomto pořadí), jde tedy o typ **NPgV** (Pg potenciálně syntakticky rozvíjí sloveso i jméno);
- předložková skupina stojí mezi slovesem a adjektivem, typ **VPgA**, nebo je předložková skupina následována adjektivem a slovesem (v tomto pořadí), tedy typ **PgAV** (Pg potenciálně syntakticky rozvíjí sloveso i adjektivum).

Díky zvolenému přístupu, analýze založené na principu redukce, může být v každém z těchto čtyř vzorců jediné substantivum nahrazeno řetězcem jmen, přičemž každé z nich může být v prostém nebo v předložkovém pádu a každé je potenciálním řídicím uzlem Pg.

Je-li naplněn některý ze základních ‚podezřelých‘ WOP, jsou na jeho základě postupně vytvořeny všechny příslušné závislosti Pg (se stejnou ‚váhou‘). □

Definice (základní WOP s preferovaným čtením):

Definujeme **slovosledné vzorce s preferovaným čtením – Pg rozvíjí adjektivum**, které se týkají povrchové reprezentace věty a platí pro prototypické struktury:

- řetězec sestávající ze substantiva, s ním shodného adjektiva a předložkové skupiny (v tomto pořadí), tedy typ **N_iA_iPg**³⁴ (preferované čtení s adjektivem rozvíjeným Pg);
- předložková skupina stojí mezi jménem a s tímto jménem shodným adjektivem, tedy typy **N_iPgA_i** a **A_iPgN_i** (preferované čtení s adjektivem rozvíjeným Pg).

Je-li naplněn některý ze základních WOP s preferovaným čtením, je na jeho základě vytvořena závislost Pg na adjektivu, a to s vysokou preferencí.³⁵ □

Na tomto místě je potřeba vysvětlit, jak chápat tyto základní slovosledné vzorce. Jejich formulace a uplatnění úzce souvisí s obecnou metodou, se kterou přistupujeme k rozboru vět, tedy s metodou redukční analýzy, popsanou v úvodu této části.

Předpokládáme, že všechna pravidla navržená pro zpracování Pg se budou uplatňovat opakovaně, ve všech fázích analýzy Pg. Proto, vyšetřujeme-li možnou závislost Pg na substantivu či slovesu, se můžeme omezit na slovosledné vzorce obsahující pouze relevantní členy věty, tedy jméno/jména a sloveso/slovesné formy – to odpovídá těm větším postupné redukce věty, kdy jsou větné členy nepodstatné z hlediska závislosti Pg na N nebo V už redukovány. Na druhé straně, zabýváme-li se možnou závislostí Pg na adjektivu (tj. možností, že Pg rozvíjí adjektivum), jsou relevantní slovosledné vzorce obsahující adjektivum. To

³⁴ Index i znamená shodu v rodě, čísle a pádu mezi substantivem a adjektivem.

³⁵ Ostatní závislosti, které vzniknou v jiných větvích RA, budou mít preferenci nižší.

znamená, že potenciální doplnění adjektiva, v našem případě realizované jako Pg, musí být analyzováno (a vypuštěno) dříve, než je redukováno jeho řídicí adjektivum (v souladu s obecným principem analýzy.) (V příkladu (IV.6a), viz níže, ilustrujeme tento postup pro odhalení možné závislosti Pg na adjektivu.)

Vlivu povrchového slovosledu na homonymii Pg jsme se dotkli už v kapitole II.2.2. věnované morfematické homonymii, kde jsme uváděli následující jednoduchý, nicméně ilustrativní příklad ((II.5b), zde (IV.3)):

Příklad: (pokračování)

(IV.3) *Chlapec honí pána na kole.*

(IV.3') *Chlapec honí na kole pána.*

(IV.3'') *Pána na kole honí chlapec.*

Původní věta (IV.3) je homonymní, Pg *na kole* lze považovat jednak za atribut substantiva *pán*, jednak za volné doplnění slovesa *honit*. Vzhledem k navržené klasifikaci naplňuje ‚podezřelý‘ slovosledný vzorec **VNPg**.

Její slovosledná varianta (IV.3') je už ovšem jednoznačná (jde o ‚jednoznačný‘ WOP typu **VPg**, se závislostí *honí na kole*). Varianta (IV.3'') je opět homonymní, její povrchový slovosled souhlasí s ‚podezřelým‘ WOP **NPgV**.

Ani v původní větě, ani v její variantě (IV.3') nelze považovat zkoumanou Pg za neshodný atribut rozvíjející substantivum *chlapec* – takové analýze brání sloveso *honit*, které (v povrchovém řetězci věty) od sebe toto podstatné jméno a Pg odděluje. (S dalšími variantami věty (IV.3) lze pracovat obdobně.) □

Problém určení adekvátního řídicího jména pro Pg následující za substantivem, které je rozvíjeno (jedním nebo několika) neshodným atributem v genitivu, jsme ilustrovali už ve větách (I.5a) a (I.5b). Doplníme zde jejich WOP (označeny (IV.4a) a (IV.4b)).

Příklady: (pokračování)

(IV.4a) *Nedávné kontakty prostějovského výrobce kabel, batohů, míčů a celé škály sportovních potřeb na všech úřadech práce ČR a SR byly zbytečné.* (PDT, bcd11eba.fs #24)

(IV.4b) *V letech 1991-1993 jsem absolvovala kurzy řízení a marketingu prestižní Collège Bart v kanadském Québecu.* (PDT, bce17zua.fs #8, pozměněno)

Slovosled ve větě (IV.4a) odpovídá WOP ‚podezřelého‘ typu **NPgV** s řetězcem jmen *kontakty_{Gen} (prostějovského) výrobce_{Gen} kabel_{Gen}, batohů_{Gen}, míčů_{Gen} a (celé) škály_{Gen} (sportovních) potřeb_{Gen}*. Na základě čistě syntaktických pravidel jsou jako potenciální řídicí členy Pg *na (všech) úřadech (práce ČR a SR)* určena substantiva *kontakty*, *výrobce* a sloveso *být*. (Ale i celá koordinační skupina [*kabely, batohy, míče a celé škála sportovních potřeb*] a také její poslední redukovaný člen, tedy jméno *potřeby*; analýza koordinačních skupin ovšem přesahuje téma této práce.)

Ve větě (IV.4b) je realizován WOP typu **VNPg**, s řetězcem *kurzy [řízení a marketingu]_{Gen} (prestižní) Collège_{Gen} (Bart)*. Pg *v (kanadském) Québecu* tedy může syntakticky záviset na slovese *absolvovat* a na jménech *kurzy* a *Collège*. (A opět na celé koordinační skupině [*řízení a marketing*] a na jejím posledním členu *marketing*. Necháváme stranou rozbor apozice *Bart*.³⁶) □

³⁶ Jde o cizí vlastní jméno, a tedy jeho morfologické kategorie nejsou spolehlivě určeny.

Podívejme se na další věty a jejich klasifikaci podle WOP.

Příklady:

(IV.5a) *Na některých místech přitom státní mateřské školky praskají ve švech.* (PDT, bm102zua #1)

(IV.5b) *Na přelomu let 1991/1992 dostala Gala tvrdý direkt.* (PDT, bcd11eba.fs #29)

(IV.5c) *Pracovníci ... ji k dokumentaci použili poprvé při demonstraci na Letné.* (PDT, bla09jk.fs #38)

(IV.5d) *Japonsko a EU budou spolupředsedat mnohostranným rozhovorům o míru na Blízkém východě.* (PDT, bl103js.fs #34, zkráceno)

(IV.5e) *Kolik diváků na celém světě bralo Městečko Twin Peaks vážně.* (PDT, bla14jk.fs #38)

(IV.5f) *Pietní vzpomínka u hrobu Antonína Dvořáka na vyšehradském hřbitově v Praze byla včera věnována 90. výročí skladatelova úmrtí.* (PDT, bla06jk.fs #34)

Slovosled ve větách (IV.5a) a (IV.5b) odpovídá ‚jednoznačnému‘ WOP typu **PgV**, Pg tedy prototypicky rozvíjí sloveso – jsou vytvořeny závislosti *na (některých) místech praskat* a *na přelomu (let 1991/1992) dostat* (v obou případech ve shodě s ‚úplným‘ rozbořením vět).

Ve větách (IV.5c) a (IV.5d) je naplněn ‚podezřelý‘ WOP **VNPg**, v (IV.5c) ve své základní podobě (se jménem v předložkovém pádu – *při demonstraci*), v (IV.5d) s řetězcem jmen (*mnohostranným) rozhovorům o míru*. V obou případech jde o věty morfematically homonymní, všechny navržené větné členy jsou potenciálně rozvíjeny zkoumanými Pg.

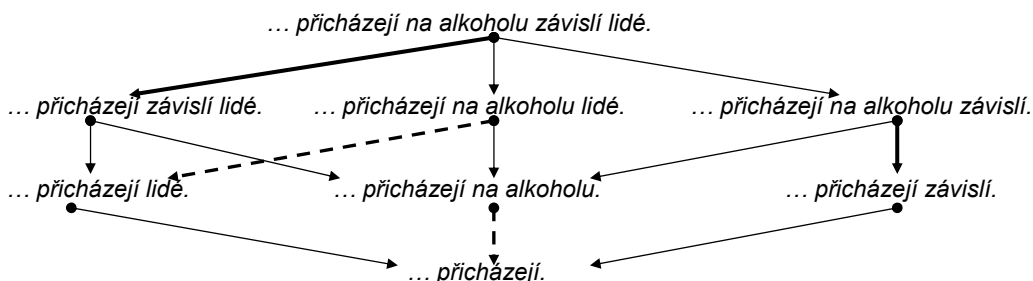
Věty (IV.5e) a (IV.5f) mohou sloužit jako příklady použití ‚podezřelého‘ WOP **NPgV**, v jeho základní podobě a v podobě s řetězcem substantiv. Ve druhém případě je k adekvátnímu určení substantiv rozvíjených Pg potřeba analýza jmenné fráze. □

Obraťme nyní pozornost k WOP, která obsahují adjektiva. Výše jsme uvedli, že pro určení možné závislosti Pg na adjektivu jsou relevantní ty stavy analýzy, ve kterých adjektivum ještě nebylo redukováno. Ilustrujme to nyní na následujících příkladech.

Příklad:

(IV.6a) *Do léčebny přicházejí na alkoholu závislí lidé.* (Panevová)

V neredukované větě (IV.6a) je naplněn ‚podezřelý‘ WOP typu **VPgA** – na jeho základě je stanovena jednak závislost Pg *na alkoholu* na adjektivu *závislý* (která je navíc podporována valenčním požadavkem tohoto adjektiva,³⁷ a jako taková získá vysokou preferenci), jednak závislost *přicházet na alkoholu* (s nízkou preferencí volného doplnění).



Obr.15: Větné analýzy věty (IV.6a) založené na redukcii

³⁷ Poznámkám k valenci deverbativních adjektiv je věnována kapitola IV.2.3.

Ve schématu na Obr.15 je vidět, že k redukci Pg může docházet i v jiných větvích RA. Silnou šipkou jsou značeny kroky, ve kterých je slovosled (redukován) věty typu **VPgA**. Čárkovaně jsou zachyceny dva kroky, ve kterých je naplněn WOP typu **VPg**, a které tedy vedou k závislosti Pg *na alkoholu* na slovese *přicházet* (bude mít nízkou preferenci volného doplnění).

(Popis aplikace jednotlivých kritérií pro WOP s adjektivem je podán v příkladu (V.2).)

V rozboru věty dochází k neadekvátní analýze Pg jako volného doplnění slovesa. Tuto analýzu nelze eliminovat pouze na základě slovosledu, lze ji pouze klasifikovat jako méně spolehlivou (viz i homonymní větu (IV.7) se stejným WOP, pro kterou jsou oba rozborů zcela adekvátní).

□

Deverbativní adjektiva s Pg zaplňující pozici ve valenčním rámci, která následují jako shodný přívlastek za svým řídicím substantivem, představují častý případ, který dokládá platnost WOP s preferovaným čtením. Ukazuje to i následující věta.

Příklad:

(IV.6b) *Neschopnost způsobila skluz v realizaci autonomie **omezené na území Golanských výšin**.* (PDT, blc02zu.fs, slovosledná varianta)

V této větě je adjektivum *omezený*, které je odvozeno od slovesa *omezit*, rozvíjeno Pg *na území (Golanských výšin)*, ve shodě s preferencí danou WOP typu **N_iA_iPg**. Toto čtení je podporováno i valenčním rámcem deverbativního adjektiva (odvození tohoto rámce ukážeme v kapitole IV.2.4.); ostatní možné závislosti, *způsobit na území (Golanských výšin)*, *skluz na území (Golanských výšin)*, *realizace na území (Golanských výšin)* a *autonomie na území (Golanských výšin)*, mají nižší preferenci.

V jiné slovosledné modifikaci této věty (IV.6b') (viz následující kapitolu) je možné Pg analyzovat jako neprojektivní doplnění adjektiva.

□

Vraťme se ještě k větě (II.9d) a k některým jejím slovosledným variantám a jejich WOP. Původní věta (zde (IV.7)) obsahovala předložkovou skupinu *na ramínku* rozvíjející jednak sloveso *rovnat*, jednak adjektivum *vystavený* (jde opět o morfematickou homonymii). Zde jsme pro jednoduchost vybrali takové varianty, jejichž slovosled už přímo odpovídá WOP, které chceme ilustrovat. V případě složitějších vět bychom k takovým WOP dospěli po jednom (příp. několika) redukčním kroku.

Příklad: (pokračování)

(IV.7) *Dívka, ..., **rovná na ramínku vystavený kabát**.* (Panevová (1998a))

(IV.7a) *Vystavený kabát **rovná na ramínku***

(IV.7b) ***Na ramínku rovná** vystavený kabát.*

(IV.7c) ***Na ramínku vystavený kabát rovná.***

(IV.7d) ***Kabát vystavený na ramínku** rovná.*

(IV.7e) ***Rovná kabát vystavený na ramínku.***

(IV.7f) ***Rovná kabát na ramínku vystavený.***

(IV.7g) ***Rovná** vystavený **kabát na ramínku**.*

Povrchový slovosled věty (IV.7) odpovídá ‚podezřelému‘ WOP **VPgA**, Pg *na ramínku* stojí mezi slovesem *rovnat* a adjektivem *vystavený*. Oba tyto větné členy jsou tedy na základě slovosledných pravidel vybrány jako řídicí členy Pg (což je ve shodě s ‚úplnou‘ analýzou věty).

Varianty (IV.7a) a (IV.7b) jsou jednoznačné a jsou také jako jednoznačné určeny – v prvním případě jde o typ **VPg** (adjektivum a předložková skupina jsou odděleny slovesem), se závislostí *rovnat na ramínku*, v druhém případě o typ **PgV** (opět je předložková skupina oddělena od adjektiva slovesem), se závislostí *na ramínku rovnat*.

Ve (IV.7c) je naplněn ‚podezřelý‘ WOP **PgAV**, předložková skupina je následována adjektivem a slovesem. Podle definice WOP tedy Pg potenciálně rozvíjí jak sloveso, tak i adjektivum, což je opět ve shodě s ‚úplnou‘ analýzou této věty (morfematická homonymie).

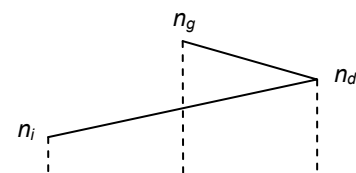
Příklady (IV.7d) a (IV.7e) ilustrují aplikaci WOP s preferovaným čtením – typ **N_iA_iPg**, substantivum *kabát* je následováno shodným adjektivem *vystavený* a předložkovou skupinou *na ramínku*, je preferováno čtení s Pg rozvíjející adjektivum.

Jak ukazuje věta (IV.7f), WOP s preferovaným čtením **N_iPgA_i** zní sice poněkud knižně, ale jistě správně. Pg *na ramínku* rozvíjí následující adjektivum *vystavený*.

Zajímavá je varianta (IV.7g) – omezujeme-li se na projektivní konstrukce, jde o ‚podezřelý‘ typ **VNPg**, Pg *na ramínku* může rozvíjet jednak substantivum *kabát*, jednak sloveso *rovnat*. Ovšem pokud uvažujeme i neprojektivní konstrukce, dostáváme další možnou závislost – adjektivum *vystavený* rozvíjené Pg *na ramínku* (takové konstrukce sice nejsou stylisticky doporučovány, viz např. Uhlířová (1987), nicméně se běžně v textech vyskytují).³⁸ □

IV.1.2. Pg v netypických pozicích a neprojektivní konstrukce

Předpokládáme, že základním typem větné struktury jsou pro češtinu projektivní konstrukce, ve kterých nominální doplnění předchází své řídicí substantivum. Neprojektivní konstrukce je třeba chápat jako takové větné konstrukce, jejichž povrchový slovosled není plně shodný s hloubkovým pořadím jejich prvků. V závislostním formalismu je charakterizuje následující podstrom:



Obr.16: Neprojektivní podstrom

V předchozí kapitole jsme ukázali, že předložkové skupiny primárně tvoří projektivní konstrukce. Nyní se soustředíme na takové typy větných konstrukcí, kde je pro adekvátní určení větných členů, které mohou být rozvíjeny předložkovou skupinou, vhodné uvolnit podmínku projektivnosti i další slovosledná omezení. Následující příklady dokládají existenci takových vět.

Příklady:

(IV.8a) *Od zahrady se ztratil klíč.* (Šmilauer (1966))

(IV.8b) *Z mravů dostal trojku.* (Šmilauer (1966))

³⁸ Pokud bychom však uvolnili podmínku projektivity, museli bychom brát v úvahu i další varianty, vrátíme se k nim v následující kapitole věnované Pg v netypických pozicích.



Obr.17: Neprojektivní analýzy vět (IV.8a) a (IV.8b)

Má-li rozbor těchto vět odpovídat jazykové intuici, musí být předložkové skupiny analyzovány jako neshodné přívlastky substantiv, přestože Pg v obou případech předchází substantiva a je porušena podmínka projektivnosti – takový rozbor je reprezentován neprojektivním stromem. Neshodný přívlastek zde není vázán na sousedství svého řídicího substantiva, může stát před ním nebo za ním – jeho postavení je dáno aktuální stavbou věty (Šmilauer (1966)). □

Popisují-li tedy projektivní konstrukce základní jazykové struktury, je potřeba rozhodnout, jakým způsobem tento projektivní model rozšířit tak, aby pokryl i jevy přesahující toto ‚jádro‘ jazyka.

Základní strategie, akceptovaná v této práci, spočívá v postupném, ‚opatrném‘ uvolňování slovosledných omezení – vždy pouze tak, aby byly k základnímu projektivnímu modelu jazyka přidány jen takové prostředky, které jsou nezbytně nutné k adekvátnímu popisu zkoumané jazykové konstrukce (tj. popisu, který odpovídá jazykové intuici).

Vztáhneme-li takto obecně popsanou strategii na problém homonymie Pg, můžeme formulovat základní principy pro určování závislosti předložkových skupin.

Základní principy pro určování závislosti Pg:

- Primárně tvoří Pg se svými potenciálními řídicími členy projektivní konstrukce. V takových případech je vhodné uplatňovat základní slovosledné vzorce vzhledem k homonymii Pg, WOP, které jsou navrženy v předchozí kapitole.
- Existují ovšem jazykové konstrukce obsahující Pg, pro jejichž adekvátní analýzu je vhodné podmínku projektivnosti a slovosledná omezení uvolnit.
- Podmínka projektivnosti i slovosledná omezení jsou oslabována pouze v definovaných případech. □

Zopakujme zde slovosledná omezení (popsaná v předchozí kapitole), která je pro některé jazykové jevy vhodné uvolnit.

Uvolňovaná slovosledná omezení:

- Předložková skupina chápaná jako nominální doplnění (prototypicky) rozvíjí pouze jména, která v povrchové reprezentaci věty této Pg předcházejí.
- Jakákoli slovesná forma slouží jako ‚zarážka‘ při hledání jména, které může být chápáno jako člen rozvíjený zkoumanou Pg.
- Předložková skupina rozvíjející adjektivum od něj nemůže být oddělena substantivem, které je tímto adjektivem rozvíjeno.

Popíšme nyní jazykové konstrukce obsahující Pg, ve kterých se zdá vhodné (a intuici odpovídající) uvolnit slovosledné podmínky, případně Pg analyzovat tak, že vznikne neprojektivita, a které umíme jednoznačně vymezit. Jak vyplývá ze základní strategie, slovosledná omezení oslabujeme ‚opatrně‘ – tak, aby tato omezení byla uvolněna jen v určených konstrukcích. K ideálnímu popisu jazyka se tedy blížíme ‚zdola‘, zmíněné podmínky uvolníme pouze v opodstatněných případech.

Pro uvolnění slovosledných omezení navrhuje tři jazykové jevy – slovesně jmenné výrazy (pokud jsou zachyceny ve slovníku), substantiva odvozená od sloves (v širokém smyslu) a neprojektivní rozvíjení adjektiv. Tyto tři jevy se mohou navzájem kombinovat.

Slovesně jmenné výrazy

Substantiva, která jsou součástí slovesně jmenného spojení, mohou být rozvíjena Pg v pozici, která není obvyklá pro nominální doplnění. (Slovesně jmenné výrazy jsou výrazy složené ze slovesa, často s velmi obecným významem, a ze substantiva, jež má valenční nebo kvazivalenční požadavek, který může být uspokojen Pg – přiblížíme je v kapitole IV.2.3.)

Následující pozorování (6) a (7) platí pro slovesně jmenné výrazy a Pg, která splňuje valenční nebo kvazivalenční požadavky příslušného substantiva – tato pozorování vedou k uvolnění slovosledných podmínek. Pro volná doplnění vyjádřená Pg ovšem slovosledná omezení platí (pozorování (8)).

Pozorování:

(6) Jmenná část slovesně jmenného výrazu může být rozvíjena (ve shodě s valenčními, příp. kvazivalenčními požadavky) předložkovou skupinou, která tomuto substantivu předchází v povrchovém uspořádání. (Tj. řídicí substantivum může následovat až za Pg, která ho rozvíjí; může být vytvořena projektivní i neprojektivní konstrukce.) Pro slovesně jmenná spojení tedy nemusí platit pozorování (1) z předchozí kapitoly.

(7) Jmenná část slovesně jmenného výrazu může být rozvíjena (ve shodě s valenčními, příp. kvazivalenčními požadavky) také předložkovou skupinou, která je od tohoto substantiva v povrchové realizaci věty oddělena libovolnou formou slovesa (které je součástí tohoto výrazu). (Tj. žádné sloveso ani žádnou část analytického tvaru daného slovesa nelze v případě slovesně jmenného spojení považovat za ‚zarážku‘ při určování závislosti Pg.) Pro slovesně jmenné kolokace tedy neplatí pozorování (2) z předchozí kapitoly. Tímto způsobem může být vytvořena neprojektivní struktura.

(8) Pro fakultativní volná doplnění vyjádřená předložkovou skupinou platí slovosledná omezení popsána v pozorování (1) a (2) – neshodný atribut (Pg) musí následovat za substantivem, které rozvíjí, a nesmí být od něj oddělen slovesem. □

V oddíle III.2. jsme uváděli větu (III.10b) a její neprojektivní variantu (III.10b') se slovesně jmenným vyjádřením *potřebovat peníze*³⁹ a s Pg *na opravu (domů)*, kterou lze chápat jako rozvíjení celé kolokace. Následující příklad ukazuje větu a její slovosledné variace s jiným slovesně jmenným výrazem.

Příklad:

(IV.9) Podle něho **má** „fyzická nebo právnická osoba“ **právo na uveřejnění odpovědi** ... (PDT, bmd23zua.fs #46, část)

(IV.9a) Podle něho **má** „fyzická nebo právnická osoba“ na uveřejnění odpovědi **právo** ...

(IV.9b) Podle něho „fyzická nebo právnická osoba“ na uveřejnění odpovědi **právo má** ...

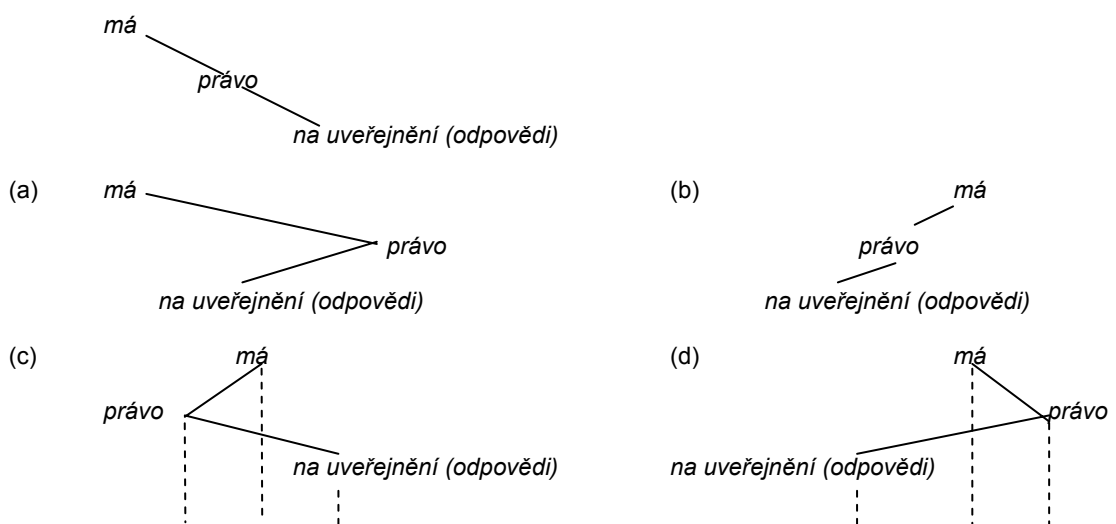
(IV.9c) **Právo má na uveřejnění odpovědi, ale už ne na omluvu.**

(IV.9d) Na uveřejnění odpovědi **má** podle něho „fyzická nebo právnická osoba“ **právo** ...

³⁹ Předpokládáme, že valenční rámec substantiva *peníze* obsahuje patient (Pat (na+Acc). kapitola IV.2.3.) a sloveso *potřebovat* má kvazivalenční doplnění náležící do druhé skupiny, ^Aim (na+Acc) (viz příklad IV.21b).

Substantivum *právo* a sloveso *mít* tvoří sousloví, tj. jedinou pojmenovávací jednotku. Protože sloveso *mít* má velmi obecný význam, navrhuje se ‚zavěšovat‘ Pg *na uveřejnění (odpovědi)* jako rozvíjí substantiva (ve shodě s valencí substantiva, viz kapitola IV.2.3., věnovaná valenci slovesně jmenných výrazů); pokud ovšem nebereme celé sousloví *mít právo* jako jeden predikát (reprezentovaný jediným uzlem).

V původní větě (IV.9) jde tedy o základní, projektivní konstrukci, je naplněn ‚podezřelý‘ WOP **VNPg** (s řetězcem jmen) z předchozí kapitoly.



Obr.18: Závislostní stromy reprezentující slovosledné varianty kolokace *mít právo (na něco)*

Uváděné slovosledné varianty základní věty už nesplňují pozorování z předchozí kapitoly. Varianty (IV.9a) a (IV.9b) ilustrují pozorování (6), Pg *na uveřejnění (odpovědi)* předchází substantivum *právo*, které rozvíjí. Neplatí tedy pozorování (1). Podmínka projektivnosti je zachována.

Ve větě (IV.9c) je porušeno pozorování (2), substantivum a jeho neshodný atribut, vyjádřený Pg, jsou odděleny určitým slovesem, což odpovídá pozorování (7). Vzniká neprojektivní konstrukce.

Varianta (IV.9d) splňuje obě pozorování (6) a (7) – Pg *na uveřejnění (odpovědi)* předchází substantivum *právo*, které rozvíjí, a navíc je odděluje určité sloveso. □

Další příklady jsou uvedeny v kapitole IV.2.2 (věty (IV.26b), *poskytnout slevy na něco*, a (IV.26c), *mít peníze na něco*).

Slovesně jmenné výrazy jsou tedy vhodnými kandidáty na uvolnění slovosledných podmínek (pokud je přímo nechápeme jako jedinou lexikální jednotku). (Automatické) rozpoznání slovesně jmenných výrazů ve větě ovšem vyžaduje, aby takové výrazy byly zachyceny ve slovníku. Zatím však není dostatek podkladů pro jejich systematické zpracování.

Závěr:

Uvolnění slovosledných podmínek pro slovesně jmenné výrazy v rozsahu daném pozorováním (6) a (7) je podmíněno vymezením těchto výrazů ve slovníku. □

Substantiva odvozená od sloves

V kapitole IV.2.2. přiblížíme třídu substantiv odvozených od sloves (a některých substantiv dalších), pro která uvolňujeme slovosledná omezení. Toto vymezení přitom chápeme velmi široce – řadíme sem kromě zřejmých syntaktických derivátů, široce slovesných jmen a lexikálních derivátů i jména, která mají ‚slovesná‘ doplnění (tj. valenční a volná doplnění prototypicky rozvíjející slovesa). Bereme přitom v úvahu i substantiva s kvazivalenčním doplněním (opět IV.2.2.). Taková substantiva – pokud ve větě není sloveso s odpovídajícím valenčním požadavkem – mohou být rozvíjena valenční Pg v pozici netypické pro doplnění substantiva (je vhodné pro ně uvolnit slovosledná omezení).

Pozorování:

(9) Pokud se v klauzi nevyskytuje sloveso s valenčním nebo kvazivalenčním doplněním, které by mohlo být vyjádřeno zkoumanou Pg, může být substantivum odvozené od slovesa (ve shodě se svými valenčními, příp. kvazivalenčními požadavky) rozvíjeno předložkovou skupinou, která tomuto substantivu předchází v povrchovém uspořádání. (Tj. řídicí substantivum může následovat za Pg, která ho rozvíjí; může být vytvořena projektivní i neprojektivní konstrukce.) Pro substantiva odvozená od sloves tedy nemusí platit pozorování (1) z předchozí kapitoly.

(10) Pokud se v klauzi nevyskytuje sloveso s valenčním nebo kvazivalenčním doplněním, které by mohlo být vyjádřeno zkoumanou Pg, může být substantivum odvozené od slovesa (ve shodě se svými valenčními příp. kvazivalenčními požadavky) rozvíjeno také předložkovou skupinou, která je od tohoto substantiva v povrchové realizaci věty oddělena libovolnou formou daného slovesa. (Tj. žádné sloveso ani žádnou část analytického tvaru slovesa nelze v případě takového substantiva považovat za ‚zarážku‘ při určování závislosti valenční Pg.) Pro substantiva odvozená od sloves tedy neplatí pozorování (2) z předchozí kapitoly. Tímto způsobem může být vytvořena neprojektivní struktura.

(11) Fakultativní volná doplnění substantiva vyjádřená Pg musí následovat za tímto substantivem a nesmí být od něj oddělena slovesem. □

Uvolníme-li na základě pozorování (9) a (10) slovosledné podmínky kladené na valenční doplnění substantiv se ‚slovesným‘ rámcem, umožníme adekvátní analýzu Pg ve větách (IV.8a) a (IV.8b) (Pg *od zahrady* lze považovat za fakultativní doplnění původu (Orig) substantiva *klíč*, resp. Pg *z mravů* za kvazivalenční doplnění zřetele (^Regard) substantiva *trojka*), stejně jako v následujících větách .

Příklady:

(IV.10a) *Na takto položenou otázku nenašel odpověď.*

(IV.10b) *(Od Jana žádné zprávy nemáme.) od Mileny jsme si právě na poště vyzvedli dopis.*

(IV.10c) *Od tohoto autora jsme (už) uveřejnili několik článků.* (Oliva)

(IV.10d) *Na některé z jejich patentů získal licenci.* (na základě PDT, bcb21aba.fs #38)

Ve větě (IV.10a) by se sousloví *najít odpověď* případně dalo pokládat za slovesně jmenný výraz,⁴⁰ (jako alternativa slovesa *odpovědět*, příp. s modálním zabarvením *umět odpovědět*), a tedy využít závěru z předchozí kapitoly.

⁴⁰ Valenční rámec substantiva *odpověď* je odvozen v příkladu (IV.22a).

V dalších větách takové řešení už uplatnit nelze, přesto je vhodné v (IV.10b) a (IV.10c) považovat Pg *od*+Gen za valenční doplnění substantiv *dopis*, resp. *článek* (v obou případech fakultativní původ, Orig).

V (IV.10d) lze uvažovat o valenčním doplnění patientu (Pat), které rozvíjí substantivum *licence* a je vyjádřeno Pg *na některé (z jejich patentů)*. □

Závěr:

Navrhujeme dále uvolnit slovosledná omezení pro substantiva odvozená od sloves a další substantiva se „slovesným“ rámcem, v rozsahu daném pozorováními (9) a (10). Uvolnění slovosledných podmínek pro valenční a kvazivalenční doplnění takových substantiv pokrývá i slovesně jmenné výrazy (předpokládáme u nich valenční nebo kvazivalenční doplnění substantiva ve formě Pg). □

Neprojektivní rozvíjení adjektiva

V kapitole věnované základním slovosledným konstrukcím s Pg jsme se dotkli možnosti připustit neprojektivní rozvíjení adjektiva předložkovou skupinou. Vraťme se k poslední zmiňované variantě věty (IV.7) a k jejím dalším variacím.

Příklad: (pokračování)

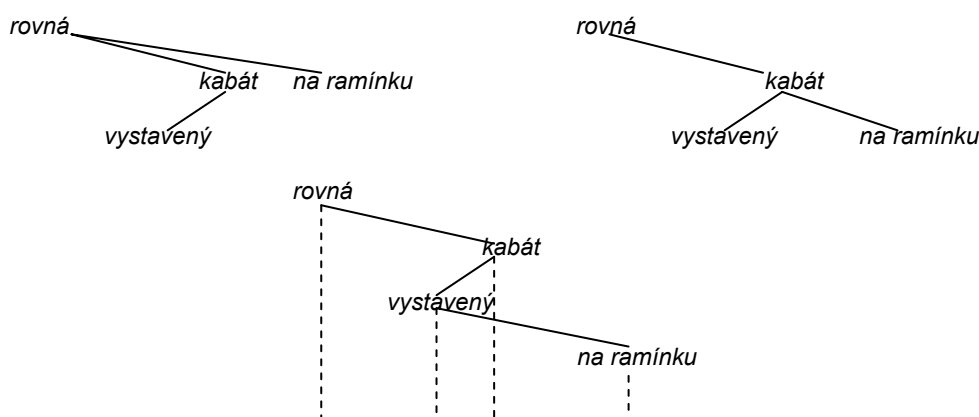
(IV.7g) *Rovná vystavený kabát na ramínku.*

(IV.7h) *?Vystavený kabát na ramínku rovná.*

(IV.7i) *?Na ramínku vystavený rovná kabát.*

(IV.7j) *?Vystavený na ramínku rovná kabát.*

Připustíme-li možnost neprojektivního rozvíjení adjektiva, dostaneme pro (IV.7g) – kromě závislostí už popsaných, *rovnat na ramínku*, *kabát na ramínku* – další možnou závislost, *vystavený na ramínku*.



Obr.19: Možné analýzy věty (IV.7g) □

Pozorování:

(12) Adjektivum (jako shodný atribut substantiva) může být rozvíjeno Pg, která je od tohoto adjektiva oddělena substantivem, v pořadí $A_i N_i P_g$. Uvolňujeme tedy podmínku projektivnosti pro rozvíjení adjektiva předložkovou skupinou.

(13) Možnost zpochybnění slovesa (jakékoli jeho formy) jako ‚zarážky‘ při hledání možného řídicího adjektiva pro Pg zůstává otevřená pro specifické kontexty. □

Podívejme se na modifikaci věty (IV.6b) z předchozí kapitoly.

Příklad:

(IV.6b') *Neschopnost dohodnout se na něm způsobila skluz v realizaci omezené autonomie na území Golanských výšin.* (PDT, blc02zu.fs, pozměněno)

Při projektivní analýze Pg je pro tuto větu (vzhledem k přijaté metodě redukční analýzy) relevantní ‚podezřelý‘ WOP typu **VNPg** s řetězcem jmen *skluz v realizaci (omezené) autonomie*, Pg tedy potenciálně rozvíjí jednak sloveso *způsobit*, jednak zmíněná substantiva *skluz*, *realizace* a *autonomie*. Uvolní-li se podmínka projektivnosti pro rozvíjení adjektiva předložkovou skupinou, je třeba připustit i závislost *omezený na území (Golanských výšin)* – v příslušných stavech RA je naplněn vzorec **A_iN_iPg** z pozorování (12). □

Kombinace jevů s uvolněnými slovoslednými podmínkami

V předcházejících odstavcích jsme se zabývali jazykovými jevy, ve kterých se zdá vhodné uvolnit základní slovosledná omezení daná pozorováními (1) až (5). Zde na několika příkladech ilustrujeme možnost kombinací takových jevů. Je přitom charakteristické, že věty, jejichž slovosled překračuje několik slovosledných omezení, jsou silně stylově zabarvené, hodí se jen do určitých kontextů charakterizovaných kontrastem některých částí výpovědi (a zřejmě též intonace).

Příklady:

(IV.11) *Odsouzení mají (jednou za měsíc) nárok na dopis od rodiny.*

(IV.11a) *Od rodiny na dopis nárok mají (, od přátel už ale ne).*

(IV.11b) *Od rodiny nárok mají na dopis (, ale už ne na balíček).*

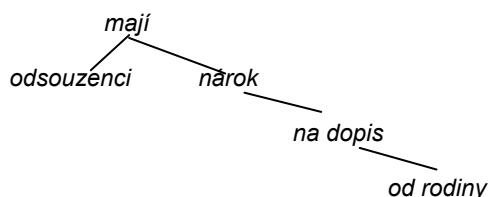
(IV.11c) *?Od rodiny na dopis mají nárok.*

(IV.11d) *Na dopis mají nárok od rodiny (, ne od přátel).*

(IV.11e) *Nárok mají od rodiny na dopis (, ne už na balíček).*

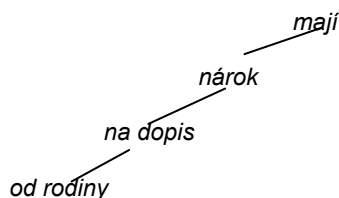
Základní věta (IV.11) obsahuje slovesně jmenný výraz *mít nárok*, který je rozvíjen Pg *na dopis (od rodiny)*; v souladu s valenčními požadavky substantiva je Pg ‚zavěšena‘ na substantivu *nárok*. Navíc podstatné jméno *dopis* má ve svém valenčním rámci fakultativní doplnění původu (Orig) s povrchovou formou *od+Gen* – tento požadavek je zaplněn Pg *od rodiny*. Jde o typickou, projektivní konstrukci.

Zdá se, že (přes slovosledné změny) ve variantách základní věty zůstávají závislosti zachovány, Pg *od rodiny* je doplněním substantiva *dopis*, které (ve formě Pg *na+Acc*) rozvíjí jmennou část slovesně jmenného výrazu *mít nárok*. Takové analýzy ovšem překračují slovosledná omezení:



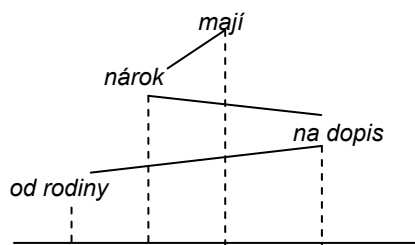
Obr.20: Závislostní strom základní věty (IV.11)

Věta (IV.11a) zůstává projektivní, Pg *od rodiny* jako doplnění substantiva *dopis* předchází své řídicí jméno (pozorování (9)), jmenná část slovesně jmenného výrazu je rozvíjena předcházející Pg (pozorování (6)).



Obr.21: Závislostní strom věty (IV.11a), s neshodnými atributy předcházejícími substantiva

Ve větě (IV.11b) Pg *od rodiny* předchází své řídicí podstatné jméno *dopis* a je od něj oddělena určitým slovesem (pozorování (9) a (10)), navíc je jmenná část slovesně jmenného výrazu oddělena slovesem od svého valenčního doplnění (pozorování (7)).



Obr.22: Neprojektivní analýza věty (IV.11b)

Obdobným způsobem jsou slovosledná omezení překročena i v dalších (neprojektivních) variantách věty.

(IV.12) *Mají právo podat odvolání proti rozsudku. / Mají právo podat proti rozsudku odvolání.*

(IV.12') *Mají právo na podání odvolání proti rozsudku.*

(IV.12a) *Proti rozsudku podat odvolání mají právo.*

(IV.12a') *?Proti rozsudku na podání odvolání mají právo.*

(IV.12b) *Podat proti rozsudku odvolání mají právo.*

(IV.12b') *?Na podání proti rozsudku odvolání mají právo.*

Věta (IV.12) obsahuje dva slovesně jmenné výrazy, *mít právo* a *podat odvolání*, z nichž druhý je rozvíjen Pg *proti rozsudku* (v prototypické pozici nebo předcházející substantivum *odvolání*). Ve větě (IV.12'), která vznikla nominalizací slovesa *podat*, se potom vyskytují dva (primárně) slovesně jmenné výrazy, oba rozvíjené Pg.

Ve větách (IV.12a), (IV.12a') a (IV.12b), (IV.12b') nejsou dodrženy prototypické pozice Pg. To je zřejmě přijatelné ve variantách s infinitivem *podat* – v (IV.12a) i (IV.12b) je slovosled porušen ve vztahu k substantivům *právo* (ovšem doplněním ve formě infinitivu, ne Pg) i *odvolání* –, ale vyvolává to pochybnosti ve variantách s nominalizací *na podání* – v (IV.12a') i (IV.12b') Pg *na podání* předchází své řídicí substantivum *právo* a je od něj oddělena slovesem; substantivum *odvolání* následuje až za svým doplněním *proti rozsudku*.

Následující věta ilustruje další jev, který může vést ke kumulaci neprojektivit.

Příklad:

(IV.13) *Na téma ekologie jsme od tohoto autora letos novinových článků uveřejnili už celou řadu.* (Oliva)

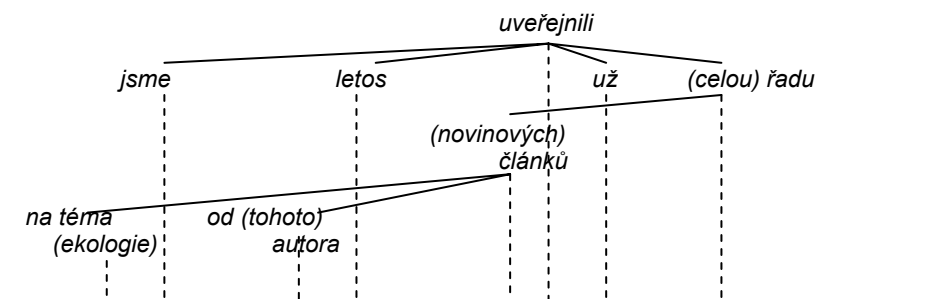
Podstatné jméno *článek* má ve svém rámci fakultativní patient a původ (Pat, forma *o+Loc*, Orig s formou *od+Gen*, viz substantiva se ‚slovesným‘ rámcem v kapitole IV.2.2.).

článek ... Pat (*o+Loc*) Orig (*od+Gen*)

Valenční pozice odpovídající patientu může být kromě typické formy zaplněna také frazeologismy *článek na téma* / *článek k tématu*.

Substantivum *řada* patří k vlastním substantivům, má obligatorní partitiv (Mat) vyjádřený genitivem:

řada ... Mat (Gen)



Obr.23: Kombinace jevů s uvolněnými slovoslednými podmínkami⁴¹

Na základě valenčních požadavků jsou Pg *na téma (ekologie)* a *od (tohoto) autora* určeny jako valenční doplnění podstatného jména *článek*, které je v povrchovém zápisu věty následuje (pozorování (9)), první z nich je od něj navíc odděleno pomocným slovesem (pozorování (10)). Samo substantivum *článek* je určeno jako partitivní doplnění substantiva *řada* (které následuje za svým neshodným atributem v genitivu a je od něj odděleno slovesem). □

Příklad (IV.13) vypovídá o existenci dalších kandidátů na oslabení slovosledných omezení pro předložkové skupiny, konkrétně o třídě vlastních substantiv s obligatorními aktanty. Tuto možnost zatím nepokládáme za dostatečně ověřenou, proto ji zatím (s ohledem na přijatou strategii ‚opatrného‘ uvolňování omezení) pomíjíme.

Shrnutí poznámek k slovosledu

Předložkové skupiny prototypicky tvoří se svými řídicími větnými členy projektivní konstrukce, v nichž doplnění substantiva vyjádřené ve formě Pg následuje za svým řídicím substantivem a není od něj odděleno slovesem. V takových případech lze pro určení Pg ‚podezřelých‘ z morfeatické homonymie použít základních slovosledných vzorců.

Ve snaze přiblížit navrhovanou analýzu možných závislostí předložkových skupin lingvistické intuici jsme pro tři typy jazykových konstrukcí navrhli oslabení slovosledných omezení. Jedná se o slovesně jmenná vyjádření rozvíjená Pg, o substantiva odvozená od sloves

⁴¹ Na RPS jsou analytické tvary sloves reprezentovány jediným uzlem, v tomto stromu je pomocné sloveso *být*, které odděluje substantivum a jeho valenční doplnění, nestandardně zachyceno zvláštním uzlem.

s valenčním nebo kvazivalenčním doplněním realizovaným Pg (pokud není taková analýza zablokována valenčním požadavkem slovesa) a pro rozvíjení adjektiv ve funkci shodných přívlastků substantiv. Ukázali jsme, že tyto tři jevy se mohou kombinovat.

Zachycení slovesně jmenných výrazů ve slovníku zatím není uspokojivě vyřešeno, nicméně uvolnění slovosledných omezení pro substantiva odvozená od sloves pokrývá i takové kolokace.

Slovosledná omezení a podmínku projektivnosti uvolňujeme ‚opatrně‘, pouze v přesně analyzovaných a popsáných případech. Neadekvátní oslabování slovosledných omezení vede v každé automatické analýze jazyka k ‚explozi‘ nepřiměřených (alespoň v naprosté většině) určení závislostí předložkových skupin.

IV.2. VALENČNÍ RÁMCE

Teorie valence, jako jeden z pilířů Funkčního generativního popisu, byla navržena a zpracována nejprve pro slovesa (viz zejména Panevová (1974-1975), Hajičová (1979), Panevová (1980) a (1994)). Byla studována i valence dalších slovních druhů – substantiv (Novotný (1980), Piřha (1981), v poslední době Panevová (2000)) a adjektiv (Piřha (1982), Prouzová (1983), Panevová (1998b)).

Valenční rámce sloves, podstatných jmen a přídavných jmen hrají zásadní roli při adekvátním určování větných členů, které mohou být rozvíjeny předložkovou skupinou. Konceptu valenčních rámců jsme se už dotkli při rozboru několika příkladů, zejména v (I.1) (stanovení závislosti *zájem na (domácím) trhu*), v (I.3) (určení pacientu *ohodnotit (minulý) rok*), v (II.23) (rozbor jmenné skupiny *nároky kladené na demokracii*) a v oddíle (II.2.). V tomto oddíle se pokusíme o soustavnější pohled na problematiku valenčních rámců sloves, substantiv a adjektiv (necháváme stranou valenční rámce adverbií) a na popis jejich přínosu pro zachycení homonymie předložkových skupin.

Z hlediska problému automatické syntaktické analýzy předložkových skupin, adekvátního stanovení jimi rozvíjeného větného členu, případně k určení jejich homonymie se zdá vhodné rozšířit teoreticky zpracovaný pojem valence ve dvou směrech – vhodně jej interpretovat na povrchové rovině a rozšířit o tzv. kvazivalenční doplnění.

(a) FGD ve své klasické verzi rozlišuje dvě roviny popisu syntaxe věty (viz oddíl II.1.) – rovinu hloubkové struktury a rovinu povrchové struktury. Pojmy valence, valenčního rámce a valenčních doplnění se týkají především tektogramatické roviny. Při syntaktické analýze češtiny je potřeba se zabývat jejich vlivem na povrchovou stavbu věty, to znamená jejich vhodnou interpretací na rovině povrchové syntaxe, a zejména na morfemické rovině popisu jazyka.

(b) Pro účely počítačového zpracování jazyka (matematických a elektrotechnických textů) rozlišuje Panevová (1966) tři typy neshodných doplnění substantiv. Valenční doplnění prvního stupně, které je ve vztahu pacientu a vyznačuje se jednoznačným povrchovým vyjádřením (*vliv na něco, srovnání s něčím, odpověď na něco*). Doplnění druhého stupně už není valenčním doplněním v přísně gramatickém smyslu, lze ho charakterizovat jako doplnění vyjadřující terminologická a frazeologická spojení vlastní jednotlivým jazykovým doménám, příp. jednotlivým vědním oborům. Forma rozvíjejícího členu je opět jednoznačná (*zbytek po x-tém členu, souměrnost podle osy*). Doplnění třetího stupně se vyznačuje několika synonymními alternativami morfemického vyjádření (*návod k něčemu/na něco, vztah k něčemu/s něčím*). (Tato hierarchie odpovídá spolehlivosti, s jakou je určen řídicí větný člen pro zkoumané neshodné doplnění, vzhledem k tehdy dostupným technickým prostředkům).

Koncept několika typů valenčních doplnění substantiv, ale i sloves a adjektiv (přínejmenším u adjektiv odvozených od sloves) lze velmi dobře využít při automatickém určování závislostí předložkových skupin.

Rozlišujeme tedy dva typy valenční informace, oba se týkají roviny podkladové reprezentace, oba je potřeba uchovávat ve slovníku: valenci (jak je popisována v teoretických studiích) a kvazivalenci (valenci 2. řádu, 2. stupně), kterou lze parafrázovat jako fakultativní, nicméně ‚obvykle užívané doplnění‘ konkrétního slova.

Poznámka:

Úvahy o valenčních doplnění prezentované v tomto oddíle jsou obecné, neopírají se o konkrétní slovník a gramatiku. □

V následující kapitolách ilustrujeme možnosti takto široce chápané valence pro analýzu závislostí předložkových skupin.

IV.2.1. Valenční rámce sloves

Pokusíme se zde o shrnutí základních pojmů teorie valence, které vypovídají především o podkladové reprezentaci věty; Panevová navíc zdůrazňuje praktickou potřebu zachytit ve valenčních rámcích i jejich možnou povrchovou formu. Budeme se snažit striktně rozlišovat mezi valenčními rámci vztaženými k tektogramatické rovině a jejich interpretací na rovině povrchové syntaxe a v morfematice.

Rovina podkladové reprezentace

Zopakujme zde v krátkosti základní pojmy týkající se valenčních rámců sloves, se kterými budeme dále pracovat.

Slovesný valenční rámec (v užším smyslu) je tvořen aktanty a obligatorními volnými doplněními slovesa (viz dále) – jsou to tzv. valenční doplnění slovesa. Každé sloveso má alespoň jeden valenční rámec (může mít rámců více), jeho charakteristika spolu s možnými povrchovými formami aktantů musí být obsažena v slovníkovém hesle daného slovesa. (Každé sloveso má tolik slovníkových položek, kolik má rámců.)

FGD rozlišuje na tektogramatické rovině pět aktantů a řadu volných doplnění. **Aktanty** jsou vnitřní doplnění slovesa, participanty, jejichž kombinace je charakteristická pro jednotlivá slovesa a která se nemohou vyskytovat více než jedenkrát (bez koordinace nebo apozice) jako rozvítky konkrétního slovesa. Jsou to aktor (konatel, dále Act), patient (Pat), adresát (Addr), původ (Orig) a výsledek (Eff). Naproti tomu **volná doplnění** (např. místa, času, způsobu, prostředku, podmínky atd.) mohou rozvíjet jakékoliv sloveso a mohou se vyskytovat u téhož slovesa vícekrát. (Případná omezení jsou sémantické, nikoli gramatické povahy.)

Aktanty mohou být u jednotlivých řídicích slov **obligatorní** (tj. povinně přítomny v hloubkové struktuře věty) nebo **fakultativní**. Některé obligatorní aktanty ovšem nemusí být realizovány v povrchové reprezentaci věty – říkáme jim (obligatorní) aktanty **vypustitelné**. Aktor je v české větě vypustitelný vždy. Podobně i volná doplnění mohou být pro konkrétní sloveso obligatorní, příp. obligatorní vypustitelná. Kritériem pro obligatornost aktantů a volných doplnění je tzv. **dialogový test** (viz Panevová (1974-1975)). Vypustitelné aktanty je třeba důsledně odlišovat od **všeobecného aktantu** (General Participant)⁴² – všeobecný aktant je jen zvláštním případem lexikálního obsazení (Panevová (1992), Panevová a Řezníčková (2001)).

Pro určení aktoru a patientu FGD užívá **syntaktická kritéria** (v duchu přístupu Tesnière (1959)), pro určení dalších vnitřních doplnění užívá kritéria **sémantická** (Fillmore (1968), (1977), u nás např. Daneš a Hlavsa (1981)). FGD vychází z toho, že se realizuje tzv.

⁴² V rámcích označujeme General.

posouvání aktantů – prvním aktantem je vždy aktor, druhým pacient. U sloves s více aktanty přistupují ohledy sémantické.

Aktanty mají pro každé sloveso pevně danou **morfematickou formu** (často jedinou, může jich být více) uchovávanou ve slovníku, volná doplnění (obligatorní i fakultativní) jsou určena sémanticky (typicky mají různé morfematické vyjádření). Existuje tzv. **prototypické vyjádření** jednotlivých aktantů (v aktivní větě aktor v nominativu, akuzativ pro pacient, dativ pro adresáta).

Valenční rámce se uvádějí pro aktivum, způsob jejich zachycení ve slovníku a pravidla pro **pasivizaci rámce** uvádí např. Skoumalová (1998). Pravidly pro naplnění rámců reflexivními zájmeny, zejména u recipročních vztahů, se zabývá Panevová (1999).

Podívejme se na příklady valenčních rámců s předložkami a jejich aplikaci při určení větných členů rozvíjených Pg. Fakultativní členy valenčního rámce konkrétního slovesa označujeme závorkou. Za názvem aktantu (jeho zkratkou – Act, Pat, Orig, Addr, Eff) obvykle uvádíme v závorce jeho možná morfematická vyjádření, oddělená lomítkem (/) – jde o pád substantiva, specifikaci Pg, vedlejší větu (klauze) (v posledním případě uvádíme možné spojovací výrazy – příslušné spojky, příp. zájmeno (zkratka Pron) nebo adverbium (Adv)). Aktor, tzv. levointenční aktant oddělujeme svislicí (|). Pro obligatorní volná doplnění možné povrchové realizace neuvádíme.

Příklady:

(IV.14a) *Nereaguje-li nájemce na nabídku v průběhu 6 měsíců, ...* (PDT, bcb01aba.fs #20)

Slovosled věty (IV.14a) odpovídá ‚podezřelému‘ WOP **VNPg**. Valenční rámec slovesa *reagovat* obsahuje kromě aktora také pacient vyjádřený zkoumanou Pg:

reagovat ... Act | Pat (na+Acc)

Na základě informace o valenčním rámci slovesa je Pg *na nabídku* určena jako slovesné doplnění, a to pacient.⁴³

(IV.14b) *Kdyby občanům skutečně nezbývalo než čekat celé volební období na možnost vyměnit své poslance, ...* (PDT, bm227zua.fs #29)

Stejný ‚podezřelý‘ WOP **VNPg** má věta (IV.14b). Protože valenční rámec slovesa obsahuje pacient, který může mít formu zkoumané Pg, je Pg analyzována jako objekt odpovídající pacientu.⁴⁴

čekat ... Act | Pat (Acc/na+Acc/klauze-že, zda, Pron, Adv)

(IV.14c) *Navrženo na ni může být české nebo slovenské beletristické dílo ...* (PDT, bmb09eba.fs #20)

Valenční rámec slovesa *navrhnout* se skládá z aktoru, pacientu a fakultativního adresátu:

(i) *navrhnout*₁ ... Act | Pat (Acc/inf/klauze-že, aby, ať, Pron, Adv) (Addr (Dat))

Vydělujeme ještě druhé užití slovesa *navrhnout*, s obligatorním doplněním účelu:

(ii) *navrhnout*₂ ... Act | Pat (Acc) (Addr (Dat)) Aim (na+Acc)

⁴³ Pacient slovesa *reagovat* je obligatorní, může však být zaplněn všeobecným aktantem. Sloveso bývá často rozvíjeno volným doplněním prostředku (Means v instrumentálu, *reagovat smíchem*) nebo způsobu (Manner, *reagovat živě*).

⁴⁴ Sloveso *čekat* je často rozvíjeno volným doplněním zřetle (Regard, s+Ins, *čekat se zahájením*) nebo směru (Dir, od+Gen, *čekat něco od někoho*).

Pasivní věta (IV.14c) naplňuje derivovaný rámec *navrhnout₂* – substantivum *dílo* vyjadřuje patient, Pg *na ni* je valenční účel (Aim).

(IV.14d) *Podle těchto zákonů by se noviny rychle přeměnily z deníků na dementníky!* (PDT, bmd23zua.fs #49)

Sloveso *přeměnit* má následující rámec:

přeměnit ... Act | Pat (Acc) Orig (z+Gen) Eff (na+Acc/v+Acc)

Ve větě (IV.14d) je sloveso *přeměnit se*, což je sekundární reflexivum tantum (viz Panevová (1999)). Dochází k posouvání aktantů Eff → Pat, výsledkem je valenční rámec:

přeměnit se ... Act | Pat (na+Acc/v+Acc) Orig (z+Gen)

Pg *na dementníky* je tedy určena jako patient, Pg *z deníků* jako původ.

(IV.14e) *Převážná většina vývozu směřuje na území bývalého Sovětského svazu.* (PDT, bmd03zua.fs #7)

(IV.14e') *Převážná většina vývozu směřuje do bývalého Sovětského svazu.*

(IV.14e'') *Převážná většina vývozu směřuje pryč.*

Věta (IV.14e) a její modifikace (IV.14e') a (IV.14e'') ilustrují různá povrchová vyjádření obligatorního volného doplnění – Kam?

směřovat ... Act | Direction

Předložkové skupiny *na+Acc* a *do+Gen* i příslovce místa patří k prototypickým formám pro volné doplnění směru. (Pg *na+Acc* je podrobně zkoumána v oddíle IV.3., jde o její primární funkci.) □

V části II. při rozboru příkladů (II.11a) (*válčit s nimi*), (II.11b) (*s nimi prohrával*) a (II.11) (*teroristé se sešli s granáty*) jsme se setkali s pravidly pro naplnění rámců reflexivními zájmeny.

Ilustrovali jsme zde uplatnění valenčních rámců sloves při určování možných závislostí předložkových skupin. Příklady dalších slovesných rámců uvádíme v následujících kapitolách v souvislosti s odvozováním valenčních rámců substantiv a adjektiv.

Rovina povrchové syntaxe

Zkoumání možných závislostí předložkových skupin, tj. analýza povrchového vyjádření vět, nás vede k obecnějším otázkám týkajícím se konceptu vypustitelnosti aktantu v povrchové realizaci věty.

Valence slovesa je teoreticky orientována primárně na tektogramatickou rovinu, ovšem pracujeme-li (např. pro účely implementace) s povrchovou syntaxí, bez valenčních rámců se neobejdeme ani tady. Pro potřeby povrchové syntaxe je nutné valenční rámce nově interpretovat. Zdá se totiž, že neexistuje člen valenčního rámce, který by byl (absolutně) nevypustitelný (alespoň ve specifických kontextech, jako je např. odpověď na otázku). Je tedy vhodné rozšířit vypustitelnost v povrchové realizaci věty (rozdělenou do dvou stupňů) na všechny aktanty i na všechna (obligatorní) volná doplnění.

Na druhé straně zřejmě pro vypouštění platí nějaká omezení; pro konkrétní sloveso realizace některého aktantu na povrchové rovině vyžaduje vyjádření jiného aktantu (alespoň formou zájmena), především patientu. Tato otázka ovšem potřebuje hlubší lingvistické zkoumání.

Zde pouze pro ilustraci několik příkladů, ve kterých označujeme jako ‚vypuštěný‘ i všeobecný aktant, protože na povrchu je vyjádřen nulou.

Hypotéza: (pro patient a ostatní aktanty)

Mají-li aktanty na povrchové rovině svou prototypickou formu, potom platí:

(1) Má-li konkrétní sloveso ve svém rámci obligatorní patient a je-li v povrchové realizaci věty vyjádřen fakultativní adresát (Addr), nebo valenční doplnění původu (Orig) či výsledku (Eff), potom musí být povrchově realizován i patient.

(2) Takové pozorování neplatí pro valenční doplnění výsledku (Eff). □

Tuto hypotézu se pokusíme ilustrovat následujícími větami a jejich modifikacemi.

Příklad:

(IV.15) *Matka předělala dětem loutku z kašpárka na čerta.* (Panevová) Act | Pat Addr Orig Eff

(IV.15a) *Matka předělala loutku z kašpárka na čerta.* Act | Pat Orig Eff

(IV.15b) *Matka předělala dětem loutku na čerta.* Act | Pat Addr Eff

(IV.15c) *Matka předělala dětem loutku z kašpárka.* Act | Pat Addr (Orig → Atr)

(IV.15d) **Matka předělala dětem z kašpárka na čerta.* *Act | Addr Orig Eff

(IV.15e) *Matka předělala dětem loutku.* Act | Pat Addr

(IV.15f) *Matka předělala loutku na čerta.* Act | Pat Eff

(IV.15g) *Matka předělala loutku z kašpárka.* Act | Pat (Orig → Atr)

(IV.15h) **Matka předělala dětem na čerta.* *Act | Addr Eff

(IV.15i) **Matka předělala dětem z kašpárka.* *Act | Addr Orig

(IV.15j) **Matka předělala z kašpárka na čerta.* *Act | Orig Eff

(IV.15k) *Matka předělala loutku.* Act | Pat

(IV.15l) **Matka předělala dětem.* *Act | Addr

(IV.15m) **Matka předělala na čerta.* *Act | Eff

(IV.15n) **Matka předělala z kašpárka.* *Act | Orig

(IV.15o) *Matka předělala.* Act

Valenční rámec slovesa *předělat* obsahuje obligatorní aktor, patient, původ a výsledek a fakultativní adresát:

předělat ... Act | Pat (Acc) (Addr (Dat)) Orig (z+Gen) Eff (na+Acc)

Valenční doplnění výsledku (Eff) je obligatorní, často ovšem zůstává nevyjádřeno na povrchové rovině – v tom případě jde o Eff všeobecný, který se obsahově (referenčně) ‚shoduje‘ s patientem (v naší větě *předělat loutku*, míní se *na jinou loutku*).

V původní větě (IV.15) jsou vyjádřena všechna valenční doplnění.

V modifikacích (IV.15a)-(IV.15d) chybí (v povrchové realizaci) vždy jedno substantivum, resp. jedna Pg. Přitom pouze varianta bez patientu (IV.15d) není správně utvořená česká věta. V (IV.15c) je ovšem třeba preferovat atributivní čtení Pg *z kašpárka*, i když sémanticky není zcela smysluplné.

Ve variantách (IV.15e)-(IV.15j) nejsou vyjádřeny dva členy, přitom věty s pacientem (IV.15e)-(IV.15g) jsou správně utvořené ((IV.15g) opět s preferovaným atributivním čtením *loutka z kašpárka*). Varianty bez pacientu (IV.15h)-(IV.15j) nelze považovat za správné.

Věty (IV.15k)-(IV.15n) mají vyjádřen jediný pravointenční aktant, opět musí být vyjádřen pacient, viz (IV.15k).

Variace (IV.15e), (IV.15k) a (IV.15m) ukazují, že první tvrzení skutečně neplatí pro Eff – ten může být vypuštěn (příp. vyjádřen nulovým lexémem), i když je realizován fakultativní adresát (IV.15e). (IV.15k) a (IV.15m) ilustrují větší ‚těsnost‘ pacientu. □

Příklad:

(IV.16) <i>Babička čte dětem knížku.</i>	Act Pat Addr
(IV.16a) <i>Babička čte knížku.</i>	Act Pat
(IV.16b) <i>Babička čte dětem.</i>	Act Addr
(IV.16c) <i>Babička čte.</i>	Act

Sloveso *číst* má kromě aktoru a pacientu dvě další doplňení, adresát a výsledek (Eff):

číst ... Act | (Pat (o+Loc) (Addr (Dat))⁴⁵ Eff (Acc/klauze-že,Pron,Adv)

Pacient zde není obligatorní, nejsou tedy splněny podmínky hypotézy. Jsou možné všechny varianty vypouštění (viz zejména (IV.16b)). □

Příklad:

(IV.17) <i>Petr dal mamince dárek.</i>	Act Pat Addr
(IV.17a) <i>Petr už dárek dal. / Petr dal malý dárek.</i>	Act Pat
(IV.17b) <i>*Petr dal mamince.</i>	*Act Addr
(IV.17c) <i>Petr dal.</i>	Act

Valenční rámec slovesa *dát* (v tomto užití) sestává ze tří valenčních doplňení v prototypických formách:

dát ... Act | Pat (Acc) Addr (Dat)

Přestože adresát je obligatorní, platí tvrzení hypotézy – pokud je na povrchové rovině vyjádřen adresát, musí být vyjádřen i pacient, viz (IV.17b). □

Příklad:

(IV.18) <i>Petr informoval rodiče o nových poznatcích.</i>	Act Pat Addr
(IV.18a) <i>Tisk už informoval o nových poznatcích.</i>	Act Pat
(IV.18b) <i>Petr mě už informoval.</i>	Act Addr
(IV.18c) <i>Petr informoval.</i>	Act

Sloveso *informovat* má tři obligatorní valenční doplňení – aktor, pacient a adresát:

informovat ... Act | Pat (o+Loc/klauze-že, zda, Pron, Adv) Addr (Acc)

Pacient a adresát tu nemají prototypickou formu – adresát je vyjádřen jménem v akuzativu, což je forma typická pro pacient. Adresát není fakultativní. (IV.18b) dokládá, že pacient může být na

⁴⁵ Zůstává otázkou, zda adresát pokládat za fakultativní, nebo zda je vždy přítomen (někdy jako všeobecný aktant).

povrchové rovině vypuštěn, i když je realizován adresát. (Nejsou splněny podmínky hypotézy.)⁴⁶ □

Rozbor věty (IV.18) napovídá, že podstatný vliv na možnost vypouštění může mít povrchová forma jednotlivých aktantů.

Zdá se, že podobná ‚souvýskytová‘ omezení, jako byla popsána v hypotéze (1), platí i pro aktanty a (alespoň některá) volná doplnění. Pro ilustraci formulujeme jednu z možných hypotéz.

Hypotéza: (pro patient a volné doplnění prospěchu)

(3) Obsahuje-li valenční rámec konkrétního slovesa obligatorní patient a je-li v povrchové realizaci věty vyjádřeno volné doplnění prospěchu (benefactive, Ben), musí být povrchově realizován i patient. □

Příklad:

(IV.19) *Od magistrátu se podařilo získat pro všechny postižené náhradní byty na sídlišti.* (PDT, bl101js.fs #39) Pat Orig Ben

(IV.19c) **Od magistrátu se podařilo získat pro všechny postižené.* *Orig Ben

(IV.19d) *Podařilo se získat náhradní byty na sídlišti.* Pat

(IV.19e) **Podařilo se získat pro všechny postižené.* *Ben

(IV.19f) **Od magistrátu se podařilo získat.* *Orig

Slovesu *získat* přináleží obligatorní doplnění aktor a patient a fakultativní doplnění původ.

získat ... Act | Pat (Acc) (Orig (od+Gen))

Je-li ve větě vyjádřen původ (zde *od magistrátu*), musí být na povrchové rovině přítomen i patient (*byty*) ((IV.19c) a (IV.19f), hypotéza (1)).

Volné doplnění prospěchu může být vyjádřeno substantivem v dativu nebo Pg *pro+Acc*,⁴⁷ tvrzení hypotézy (3) je podporováno nepřijatelností vět (IV.19c) a (IV.19e). □

Cenným zdrojem dat pro zkoumání problému vypustitelnosti obligatorních aktantů v povrchové realizaci věty je Pražský závislostní korpus (PDT), zejména anotace na tektogramatické rovině. Ale i anotace na analytické rovině, která aproximuje povrchovou syntax a je v současné době ve fázi dokončování, poskytuje cenné údaje.

Kvazivalenční doplnění sloves

Kvazivalenční doplnění slovesa zachycuje ustálené, často používané spojení slovesa a jazykové jednotky, která ho rozvíjí, v našem případě předložkové skupiny. Může se jednat o přenesené (někdy až idiomatičké) užití slovesa. Charakteristika kvazivalenčních doplnění včetně možných povrchových forem musí být, stejně jako u standardních valenčních doplnění, součástí slovníkového hesla. Kvazivalenční doplnění budeme značit symbolem ^ před jménem doplnění.

⁴⁶ V (IV.18a) je pozice adresátu obsazena všeobecným aktantem, v (IV.18b) je všeobecný patient; nicméně navržená hypotéza se týká povrchového vyjádření, takže ji to neovlivní.

⁴⁷ Kromě benefaktoru bývá sloveso *získat* rozvíjeno volným doplněním účelu (Aim, s formou *na+Acc/za+Acc*, *získat (mládež) na brigádu / za přítele*), substituce (Subst, *za+Acc*, *získat něco za něco*) a způsobu nebo prostředku (Means, *získat koupí, darem*).

Slovesná doplnění, která je možno popsat pomocí kvazivalence, lze v zásadě rozdělit do dvou skupin. Do **první skupiny** patří specifická doplnění sloves, která jsou užitá v posunutém nebo přeneseném významu, a také sloves tvořících se svým doplněním realizovaným jako Pg idiomatická spojení.

Příklady:

(IV.20a) *A Slováci na tento fakt hřeší.* (PDT, blc03zu.fs #10, zkráceno)

Sloveso *hřešit* může být použito třemi způsoby:

(i) *někdo*_{Nom} *hřeší*₁ (*něčím*_{Ins}) (*proti něčemu*_{Dat})

*hřešit*₁ ... Act

(Tedy jako jednovalenční sloveso, které se často pojí se dvěma volnými doplněními s prototypickými formami, prostředku (Means, v instrumentálu) a prospěchu (Benefactor, s povrchovým vyjádřením *proti*+Dat.)).

(ii) *někdo*_{Nom} *hřeší*₂ *na něco*_{Acc}

*hřešit*₂ ... Act | ^Pat (*na*+Acc)

(Tedy s aktorem v nominativu a pacientem vyjádřeným jako Pg *na*+Acc; v tomto případě jde o idiomatické užití slovesa, které je zachyceno pomocí kvazivalenčního rámce.)

(iii) *někdo*_{Nom} *hřeší*₃ *na někom*_{Loc}

*hřešit*₃ ... Act | ^Pat (*na*+Loc)

(Také s aktorem a pacientem, tentokrát s povrchovou formou *na*+Loc; opět jde o idiomatické užití slovesa a jeho kvazivalenční rámec.)⁴⁸

Díky morfologické charakteristice Pg *na (tento) fakt (na*+Acc) je pro analýzu použit kvazivalenční rámec *hřešit*₂.

(IV.20b) *Snažíme se tlačit na vládu, aby ...* (PDT, bl102js.fs #9)

Valenční rámec slovesa *tlačit* je tvořen aktorem a pacientem (v akuzativu), sloveso je často rozvíjeno fakultativním volným doplněním směru (Direction, jako fakultativní volné doplnění do rámce nepatří).

(i) *tlačit*₁ ... Act | Pat (Acc)

Kromě tohoto ‚základního‘ užití lze sloveso *tlačit* použít ve spojení:

(ii) *někdo*_{Nom} *tlačí*₂ *na někoho*_{Acc} / *na něco*_{Acc}

*tlačit*₂ ... Act | ^Pat (*na*+Acc)

Toto spojení má jak primární význam *působit tlakem (tlačit na tužku)*, tak význam přenesený, *činit nátlak (zájmové skupiny tlačí na vládu)*. V obou případech jde o ‚obvyklé‘ spojení zachycené kvazivalenčním rámcem, často rozvíjené volným doplněním prostředku v instrumentálu (Means).

Dále je možné sloveso *tlačit* užit jako idiom (*dítě sedí na nočníku a tlačí*).

Při analýze věty (IV.20b) je využit kvazivalenční rámec *tlačit*₂, Pg *na vládu* je určena jako ‚obvyklé‘, tj. kvazivalenční doplnění slovesa.

(IV.20c) *Japonský lék nepomohl na kanadské problémy.* (na základě PDT, bcc05zua.fs #2)

⁴⁸ Poznamenejme již zde, že pro výběr patřičného kvazivalenčního rámce jsou podstatné sémantické rysy doplnění – zatímco pro *hřešit*₂ je kvazivalenční pacient neživotný, *hřešit*₃ vyžaduje ^Pat s rysem ‚human‘, příp. ‚anim‘.

Sloveso *pomoci* má ve svém ‚standardním‘ rámci aktor, fakultativní patient a obligatorní adresát (často vyjádřen všeobecným aktantem):

(i) *pomoci*₁ ... Act | (Pat (s+Ins/inf/klauze) Addr (Dat)

Kromě toho může být rozvíjeno volným doplněním prostředku (Means, v instrumentálu) a účelu (Aim, prototypická forma *k*+Dat), resp. ‚negativního‘ účelu (s formou *od*+Gen), jako ve větách *Lékař pomohl (svou radou) pacientovi ke zdraví.*, resp. *Lékař pomohl (svou radou) pacientovi od bolesti.*

Sloveso *pomoci* může být užito také druhým způsobem:

(ii) *něco*_{Nom} *pomáhá*₂ *na něco*_{Acc}
*pomoci*₂ ... Act | ^Pat (*na*+Acc)

(V tomto případě jde o idiomatičtější užití slovesa a jeho rámec, který sestává z aktoru a kvazivalenčního patientu, tentokrát s povrchovou formou *na*+Acc. V tomto užití může být sloveso rozvíjeno volným doplněním prospěchu – *pomohl mu na jeho problémy.*)

Na základě kvazivalenčního požadavku je tedy Pg *na (jeho) problémy* určena jako doplnění slovesa. K větě (IV.20c) se ještě vrátíme v souvislosti s valenčním rámcem substantiv. Její slovosledné varianty jsou rozebírány v příkladu (IV.32). □

Poznámka:

Slovesa a jejich doplnění, která vyjadřují posunutý nebo přenesený význam slovesa, se svými vlastnostmi blíží ‚klasické‘ valenci. To má vliv zejména na stanovení spolehlivosti jednotlivých závislostí při křížení valenčních požadavků. Pro účely automatického zpracování předložkových skupin klasifikujeme první typ kvazivalenčních doplnění sloves stejně jako jejich doplnění valenční (viz kapitola IV.2.5., příklad (IV.32)). □

Druhou skupinu doplnění, kterou lze popsat pomocí kvazivalence, tvoří volná doplnění, která nejsou obligatorní (a tedy nepatří do standardního rámce), nicméně jejich častý výskyt u daného slovesa ukazuje na to, že k němu jistým způsobem přináležejí. V tomto případě se dá mluvit o obvyklosti doplnění (*jít* – ^Kam, Direction *bydlet* – ^Kde?, Locative). Často jde o doplnění zřetele (^Regard) a účelu (^Aim), která potom mají jednoznačně určenou povrchovou formu. Zachycení těchto doplnění má nesporně význam pro automatickou syntaktickou analýzu.⁴⁹

Příklad:

(IV.21a) *Majitele domu, ... , by obec natolik zvýhodnila na daních, že ...* (PDT, bl102js.fs #14)

Valenční rámec slovesa *zvýhodnit* sestává z aktoru a patientu (v akuzativu):

zvýhodnit ... Act | Pat (Acc) ^Regard (v+Loc/na+Loc)

Pg *na daních* (*na*+Acc) je kvazivalenčním doplněním zřetele (^Regard).

(IV.21b) *Příští rok by na nutné opravy potřebovali navíc nejméně 300 000 korun.* (PDT, bmc07aba.fs #5)

Pro rozbor věty (IV.21b) je podstatný rámec slovesa *potřebovat*, který je tvořen valenčním aktorem a patientem a kvazivalenčním účelem (^Aim):

⁴⁹ Vzhledem k tomu, že pro účely automatické syntaktické analýzy Pg přiřazujeme kvazivalenční doplnění první skupiny k doplněním valenčním, zůstávají ‚obvyklá doplnění‘ jediným typem popsaným pomocí kvazivalence sloves.

potřebovat ... Act | Pat (Acc) ^Aim (na+Acc)

Pg *na (nutné) opravy* je tedy kvazivalenčním doplněním účelu (^Aim).⁵⁰ □

Zvláštní případ představují slovesa, která tvoří se svým doplněním realizovaným Pg frazeologická spojení.

Příklad:

(IV.21c) *Zdůraznil, že měl na mysli zbytečnost vzhledem k dohodám dvou vítězných parlamentních stran.* (PDT, bm207zua.fs #16, zkráceno)

Slovní spojení *mít na mysli* tvoří idiomatické spojení a jako takové může být jeho Pg zachyceno jako kvazivalenční doplnění:

mít ... Act | Pat (Acc/klauze-že, ?aby, Pron, Adv) ^Locative (na *mysli*)

Druhou možností je pokládat tuto kolokaci za frazém, který se dále syntakticky nerozebírá. V tom případě by se zdálo vhodné stanovit kvazivalenční rámec celého spojení:

mít na mysli ... Act | ^Pat (Acc/klauze-že, ?aby, Pron, Adv)

Myšlenka zavedení rámců pro slovní spojení ovšem přesahuje záměry této práce, kloníme se k principu podle možnosti syntakticky analyzovat i frazémy – velká část z nich se svou syntaktickou skladbou neliší od běžných větných celků. □

Závěr:

Valenční i kvazivalenční informace přispívají zásadním způsobem ke zpracování předložkových skupin; při automatických aplikacích, zpracovávajících psané věty (tj. jejich povrchovou reprezentaci), je však lze využít pouze ‚jednosměrně‘ – pokud je ve větě identifikován člen splňující valenční požadavky slovesa, je vhodné ho analyzovat jako valenční doplnění. Skutečnost, že se nepodařilo zaplnit valenční pozice, ovšem nic nevyovídá o syntaktické úplnosti (správnosti) věty. □

IV.2.2. Valenční rámce substantiv

Valence substantiv byla studována již v osmdesátých letech, v poslední době se její problematika opět dostává do popředí v souvislosti s vytvářením Pražského závislostního korpusu, konkrétně s anotací na úrovni lingvistického významu (odpovídající rovinně hloubkové struktury, TGR).

V následujících poznámkách se zaměříme na **podstatná jména odvozená od sloves**, neboť právě tato substantiva mohou mít ve svém ‚zdeděném‘ valenčním rámci zachována doplnění vyjádřená předložkovými skupinami. Třidu takových substantiv chápeme široce – počítáme sem i substantiva, která se vyskytují jako součást predikátu, a také substantiva, která nejsou morfologicky odvozena od sloves, nicméně mají ‚slovesná‘ valenční doplnění (tj. valenční a volná doplnění prototypicky rozvíjející slovesa). Nebudeme se zde zabývat valencí primárních substantiv, která jsou typicky rozvíjena doplněním v genitivu, a nejsou tedy zajímavá z hlediska zachycení homonymie předložkových skupin.

⁵⁰ Zajímavé je srovnání se slovesně jmenným výrazem, *potřebovat peníze/prostředky (na něco)*, viz kapitola IV.2.3.

Tektogramatická rovina

Pro podrobnou analýzu valenčních rámců substantiv je potřeba rozlišovat různé typy derivátů – jde o deriváty syntaktické, o široce dějová jména a o deriváty lexikální (Kuryłowicz (1936)). Obecně lze říci, že substantiva odvozená od sloves určitým způsobem dědí valenční rámec zdrojových sloves (viz Panevová (2000)). Může ovšem docházet k určitému stupni abstrakce – potom jméno některá doplnění ztrácí. Je tedy vhodné uvažovat stejný repertoár aktantů a volných doplnění jako u sloves, který u lexikálních derivátů rozšiřujeme o speciální nominální doplnění.

Syntaktické deriváty slovesa (slovesná, příp. verbální substantiva) sdílejí rámec s originálním slovesem (Piřha (1981): zůstávají sémanticky slovesy), dochází k regulárním změnám v povrchovém vyjádření, viz dále. Příkladem syntaktické derivace je nominalizace.

Také při derivaci tzv. **široce dějových jmen**, která jsou na hranici derivace syntaktické a lexikální, se přenáší valenční rámec původního slovesa v úplnosti. Ten se u některých substantiv zachovává i formálně, jindy dochází ke změnám v povrchovém vyjádření, viz níže.

Pod pojmem **lexikální deriváty** sloves se skrývá široká škála podstatných jmen, která už mají vlastnosti sémantických substantiv, nicméně v jejich rámci se stále ještě mohou zachovávat doplnění podkladového slovesa. Proto se u nich předpokládá možnost slovesných i speciálních nominálních doplnění. Nominální doplnění jsou partitiv (Part, někdy též material, Mat, v širokém smyslu kvantitativní určení), přináležitost (App, appurtenance), identita, restriktivní a deskriptivní přívlastek. Partitiv a identitu lze považovat za aktanty (partitiv u některých jmen obligatorní – *skupina, hektar*, u jiných fakultativní – *sklenice, koš*), ostatní jsou volná doplnění (přináležitost je však např. u relačních jmen obligatorní – *bratr, přítel*). Při derivaci často dochází k posunu významu (*náklady*) nebo ke konkretizaci (*půllitr*).

Zvláštní pozornost věnujeme valenci substantiv, která mohou být **součástí predikátu** a která tvoří spolu se slovesem **slovesně jmenné výrazy**.

Valenční rámce substantiv, resp. pravidla, podle kterých se odvozují (včetně nesystémových změn u jednotlivých lexikálních položek), spolu s možnými povrchovými formami jednotlivých doplnění, je nutné uchovávat ve slovníku.

Povrchová rovina

Ačkoliv se u slovesných derivátů může zachovávat sémantická obligatornost valenčního doplnění (tj. obligatornost doplnění na TGR), neexistuje u substantiv povrchová obligatornost valenčních doplnění. Všechny členy valenčních rámců substantiv jsou v povrchové realizaci věty **vypustitelné**. Valenční informace lze tedy při automatických aplikacích využít pouze jedním směrem – pokud je nalezen člen splňující valenční požadavky substantiva, je vhodné ho analyzovat jako nominální valenční doplnění (stejně jako u sloves).

Při odvozování podstatných jmen od sloves se určitým způsobem zachovává rámec na TGR. U některých typů derivace zůstává formálně nezměněn (některá dějová jména). Při syntaktické derivaci dochází k **regulárním změnám** v povrchové realizaci valenčních členů (např. změna pádu), resp. doplnění mohou nabývat několika forem. Uplatňuje se hierarchie forem (nepřevádějí se jednotlivé aktanty, ale celé rámce (Panevová (2000))).

Obvykle se mění doplnění v nominativu (subjekt) na genitiv (subjektový) nebo přívlastňovací adjektivum (Adj_{pos-subj}), případně je vyjádřen instrumentálem nebo Pg *od+Gen*. Slovesná doplnění vyjádřená akuzativem často mění formu na genitiv objektový. Naproti tomu

doplnění v genitivu, v dativu a v instrumentálu typicky zůstávají zachována, stejně jako předložkové skupiny.

Nom	→	Gen _{subj} / Adj _{pos-subj} (příp. Ins / <i>od</i> +Gen)
Acc	→	Gen _{obj}
Gen	→	Gen _{obj}
Dat	→	Dat
Ins	→	Ins
<i>na</i> +Acc	→	<i>na</i> +Acc

V této práci nevyznačujeme závorkami fakultativnost jednotlivých doplňení substantiv (nepovažujeme to za potřebné pro automatické zpracování doplňení předložkovou skupinou).

Syntaktické deriváty a široce dějová jména

Příklady:

(IV.22a) *Podle něho má „fyzická nebo právnická osoba“ právo na uveřejnění odpovědi na uveřejněné stanovisko či názor, který se jí přímo dotýká.* (PDT, bmd23zua.fs #46)

Deverbativum *odpověď* v (IV.22a) je odvozeno od slovesa *odpovídat*, které má následující valenční rámec:

odpovídat ... Act | (Pat (*na*+Acc)) (Addr (Dat)) Eff (klauze-že, Pron, Adv / General)

Rámec slovesa *odpovídat* je tedy tvořen obligatorním aktorem a výsledkem (Eff) a fakultativním patientem a adresátem.⁵¹ Valenční doplnění výsledku má často povahu všeobecného participantu, ve smyslu *odpovídat odpověď*; může být vyjádřen také jako závislá klauze (*Odpořídá, co mu slina na jazyk přinesla.*)

Odvozené substantivum *odpověď* dědí rámec původního slovesa:

odpověď ... Act Pat Addr Eff

Morfematické vyjádření aktoru závisí na formě ostatních aktantů (především patientu):

*někdo*_{Nom} *odpovídá* (*někomu*_{Dat}) *na něco*_{Acc}, *že* ...

→ *někoho*_{Gen-subj} / *něčí*_{pos-subj} *odpověď* (*někomu*_{Dat}) *na něco*_{Acc}, *že* ...

Ve větě (IV.22a) je tedy Pg *na (uveřejněné) (stanovisko či názor)* v souladu s valenčními požadavky substantiva *odpověď* analyzována jako jeho patient.

(IV.22b) *K nominaci Jiřého V. Kotase na čs. prezidenta předseda Strany zelených Aleš Mucha poznamenal, že ...* (PDT, bm122zua.fs #2)

Ve větě (IV.22b) je podstatné jméno *nominace* odvozeno od slovesa *nominovat*, jehož rámec sestává ze tří valenční pozic – pro aktor, adresát a patient:

nominovat ... Act | Pat (*na*+Acc) Addr (Acc)

nominace ... Act Pat Addr

Při derivaci substantiva je zachován valenční rámec, může mít tři různá morfematická vyjádření – (i) aktor (genitiv subjektový nebo přivlastňovací adjektivum), adresát (v genitivu) a patient (*na*+Acc), (ii) homonymní přivlastňovací adjektivum, resp. genitiv subjektový / objektový, (odpovídající buď aktoru, nebo adresátu) a patient (*na*+Acc) nebo (iii) adresát (genitiv objektový nebo přivlastňovací adjektivum), patient (*na*+Acc) a případně aktor v instrumentálu.

⁵¹ Opět zůstává otázkou, zda adresát pokládat za fakultativní, nebo zda je vždy přítomen (někdy jako všeobecný aktant), stejně ve větě (IV.16).

*někdo*_{Nom} *nominuje někoho*_{Acc} *na něco*_{Acc3}

→ *někoho*_{Gen-subj} / *něčí*_{pos-subj} *nominace někoho*_{Gen-obj} *na něco*_{Acc}

→ *někoho*_{Gen-subj/obj} / *něčí*_{pos-sub/obj} *nominace na něco*_{Acc}

→ *někoho*_{Gen-obj} / *něčí*_{pos-obj} *nominace na něco*_{Acc} (*někým*_{Ins})

Protože Pg *na (čs.) prezidenta* splňuje valenční požadavky substantiva *nominace*, je přednostně analyzována jako atributivní doplnění, patient.

(IV.22c) *Ani poslední jednání mezi manažerem pražské Slavie V. Ježkem a fotbalovými funkcionáři Olomouce nerozhodlo o budoucím angažmá brankáře Luboše Přibyla.* (PDT, bl123jsa.fs #40)

Pro sloveso *jednat* rozlišujeme tři lexikální jednotky – kromě *jednat*₁ také jednovalenční *jednat*₂ a dále *jednat*₃ (= *zacházet*) s aktorem a obligatorním patientem (*s+Ins*) a obligatorním doplněním způsobu (*Manner*).

(i) *jednat*₁ ... Act^R | (Pat (o+Loc)) Addr^R (*s+Ins*)

(ii) *jednat*₂ ... Act

(iii) *jednat*₃ ... Act | Pat (*s+Ins*) Mann

Ve větě (IV.22c) substantivum *jednání* (jako proces) vzniklo nominalizací slovesa *jednat*₁ (ve smyslu *vyjednávat*), kromě toho *jednání* označuje událost (‘vzdálenější’ od původního slovesa).

Sloveso *jednat*₁ je trojvalenční, kromě obligatorního aktoru má fakultativní patient a adresát; aktor a adresát mohou být použity recipročně (index R, viz Panevová (1999)) (*Přátelé jednají (spolu).*) To se projevuje i při derivaci substantiva:

jednání ... Act Pat Addr

Možné vyjádření jednotlivých aktantů je opět ovlivněno hierarchií forem:

*někdo*_{Nom} *jedná*₁ (*s někým*_{Ins}) (*o něčem*_{Loc})

→ *jednání něčí*_{pos-subj}/*někoho*_{Gen-subj} (*s někým*_{Ins}) (*o něčem*_{Loc})

*někdo*_{Nom}^R *jedná*₁ (*o něčem*_{Loc}) → *jednání někoho*_{Gen-subj}^R (*o něčem*_{Loc})

→ *jednání (někoho*_{Gen-subj} *a někoho*_{Gen-subj}) (*o něčem*_{Loc})

→ *jednání (mezi někým*_{Ins} *a někým*_{Ins}) (*o něčem*_{Loc})

Předložková skupina *mezi (manažerem (pražské Slavie V. Ježkem) a (fotbalovými) funkcionáři (Olomouce))* je na základě valenčních požadavků substantiva *jednání* určena jako jeho valenční doplnění (patient). □

Při derivaci jmen od sloves může docházet k ‘terminologizaci’ určitých vazeb, k posunu významu či k osamostatňování substantiv. Piřha (1981) mluví o různém stupni abstrakce.

Příklady:

(IV.23a) *Návrh na zachování vlastnického práva obce na pozemek, na němž stojí příslušný činžovní dům, Sněmovna zamítla.* (PDT, bcb01aba.fs #15)

Věta (IV.23a) obsahuje deverbativum *návrh*. V rozboru věty (IV.14c) jsme specifikovali dva valenční rámce slovesa *navrhnout*:

(i) *navrhnout*₁ ... Act | Pat (Acc/inf/klauze-že, aby, ať, Pron, Adv) (Addr (Dat))

(Ve významu *vláda navrhuje (parlamentu) rozpočet / zachovat rozpočet / že zachová rozpočet.*)

(ii) *navrhnout*₂ ... Act | Pat (Acc) (Addr (Dat)) Aim (*na*+Acc)

(Ve významu *kritik navrhuje (porotě) film na ocenění.*)

Od obou sloves je možno derivovat substantiva, která dědí příslušné valenční rámce (není zachována sémantická obligatornost patientu):

(i) *návrh*₁ ... Act Pat Addr

*někdo*_{Nom} *navrhuje*₁ (*někomu*_{Dat}) *něco*_{Acc} / inf / *že* ...

→ *něčí*_{pos-subj} *návrh*₁ *něčeho*_{Gen-obj} / *na něco*_{Acc} / inf / *že* ...

(Ve významu *ministrův návrh rozpočtu / na zachování rozpočtu / zachovat rozpočet / že bude zachován rozpočet.*)

→ *něčí*_{pos-subj} *návrh*₁ *někomu*_{Dat}

(ve významu *ministrův návrh parlamentu*)

→ *něčí*_{pos-subj}/*někoho*_{Gen-subj} *návrh*₁ *na něco*_{Acc} / inf / *že* ...

(ve významu *ministrův návrh na zachování rozpočtu / zachovat rozpočet, resp. návrh vlády na zachování rozpočtu / zachovat rozpočet*)

→ *něčí*_{pos-subj}/*někoho*_{Gen-subj} *návrh*₁ *někomu*_{Dat}

(ve významu *ministrův_{pos-subj} návrh parlamentu_{Dat} / návrh vlády_{Gen-subj} parlamentu_{Dat}*)

Obdobně pro sloveso *navrhnou*₂:

(ii) *návrh*₂ ... Act Pat Addr Aim

*někdo*_{Nom} *navrhne*₂ (*někomu*_{Dat}) *něco*_{Acc} *na něco*_{Acc}

→ *něčí*_{pos-subj} *návrh*₂ *něčeho*_{Gen-obj} *na něco*_{Acc}

(ve významu *kritikův_{subj} návrh filmu_{obj} na ocenění*)

→ *něčí*_{pos-subj/obj} *návrh*₂ *na něco*_{Acc}

(ve významu *kritikův_{subj} návrh na ocenění, resp. ?umělcův_{obj} návrh na ocenění*)

→ *návrh*₂ *někoho*_{Gen-subj/obj} *na něco*_{Acc}

(ve významu *návrh kritika_{subj} na ocenění, resp. návrh umělce_{obj} na ocenění*)

→ *něčí*_{pos-subj}/*někoho*_{Gen-subj} *návrh*₂ *někomu*_{Dat}

(ve významu *kritikův_{subj} návrh porotě, resp. návrh kritika_{subj} porotě*)

Ve větě (IV.23a) jde o deverbativum *návrh*₁ (derivát slovesa *navrhnout*₁), Pg *na zachování* je patient (Pat). Forma Pg *na* povrchové rovině ovšem splývá s valenčním doplněním účelu (Aim) substantiva *návrh*₂.

(IV.23b) *Roční náklady na provoz speciální školy J.A.Komenského činí 6 milionů korun.* (PDT, bm102zua.fs #21, zkráceno)

Ve větě (IV.23b) je substantivum *náklad* (pl *náklady*), které je derivátem slovesa *nakládat*. Rozlišujeme dva základní a jeden přenesený význam slovesa *nakládat*, které jsou specifikovány třemi valenčními rámci:

(i) *nakládat*₁ ... Act | Pat (Acc) Direction

(Ve významu *Petr nakládá písek na auto, resp. Petr nakládá písek* (valenční pozice pro obligatorní volné doplnění směru je zaplněn všeobecným participantem); ale i s posunutým významem *Petr nakládá okurky do octa.*)

(ii) *nakládat*₂ ... Act | Pat (Acc)

(Ve významu *Petr nakládá auto*, pacient typicky vyjádřen substantivem s rysem ‚kontejner‘; často rozvíjeno volným doplněním prostředku (Means), *Petr nakládá auto lopatou*, ale i *nakládat auto pískem*.)

(iii) *nakládat*₃ ... Act | Pat (s+Ins) Mann

(Ve významu *Petr špatně nakládá s manželkou*.)

Primárně je substantivum *náklad* derivátem slovesa *nakládat*₁ nebo *nakládat*₂:

(i) *náklad*_i ... Act Pat Direction

*někdo*_{Nom} *nakládá*₁ *něco*_{Acc} *někam* (*na něco*_{Acc})

→ (*něčí*_{pos-subj}/*někoho*_{Gen-subj}) *náklad*_i *něčeho*_{Gen} (*někam* / *na něco*_{Acc} / ...)

(Ve významu *náklad písku* (*na auto*), aktor obvykle nevyjádřen.)

(ii) *náklad*_j ... Act Pat

Means: *někdo*_{Nom} *nakládá*₂ *něco*_{Acc}

→ (*něčí*_{pos-subj}/*někoho*_{Gen-subj}) *náklad*_j *něčeho*_{Gen}

(Ve významu *náklad auta*, aktor obvykle nevyjádřen.)

Sekundárně se substantivum *náklad* používá v posunutém významu, s pozměněným valenčním rámcem (obvykle v plurálu) – se ‚zabudovaným‘ pacientem (viz Panevová (2000)).⁵²

(iii) *náklad*_k ... Act Eff (*na*+Acc)

(Ve významu *něčí/někoho náklady na živobytí*, zabudovaný pacient,⁵³ který zastupuje *peníze, sumu, energii, ...*)

Podstatné jméno *náklad* se užívá ještě v jednom přeneseném významu, kde ‚mizí‘ aktor (vlivem abstrakce):

(iv) *náklad*_m ... Ø

(Ve významu *náklad knihy*, kde se jedná o nominální doplnění, partitiv; v tomto významu také tvoří idiomy *vyšlo nákladem x výtisků*, *nákladem vlastním*.)

Ve větě (IV.23b) je ve spojení *náklady na provoz* (...) užito substantiva *náklad*_k, je naplněn jeho valenční rámec – proto považujeme Pg *na provoz* (...) za valenční doplnění výsledku (Eff).

□

V této kapitole se zabýváme substantivy odvozenými od sloves. Derivace ovšem může probíhat i obráceně, ‚původním‘ výrazem může být substantivum (tj. slovotvorně základním, ať už je významově základní substantivum nebo sloveso). I v takových případech je zajímavé porovnat jejich valenční rámce – následující příklad ilustruje postup, kdy je sloveso odvozeno od substantiva.

Příklad:

(IV.23c) *Japonský lék na kanadské problémy*. (PDT, bcc05zua.fs #2)

Substantivum *lék* je primárním substantivem, jeho valenční rámec obsahuje jedině doplnění, fakultativní pacient:

⁵² Zdá se, že jde o vliv slovesa *vynakládat*, jehož rámec je tvořen aktorem, pacientem v akuzativu a efektem vyjádřeným Pg *na*+Acc:

vynakládat ... Act | Pat (Acc) Eff (*na*+Acc)

⁵³ Pojem zabudovaného aktantu zavádí např. Panevová (2000).

lék ... Pat (na+Acc/proti+Dat)

Ve spojeních *lékařův_{subj}/Pavlinin_{obj} lék na kašel/proti kašli*, resp. *lékaře_{subj}/Pavliny_{obj} lék na kašel/proti kašli* vyjadřují formy Gen/Adj_{pos} volné doplnění přináležitosti (App).

Srovnajme tento rámec substantiva s valenčními rámci odvozeného slovesa *léčit*:

(i) *léčit₁* ... Act | Pat (Acc)

(ve významu *lékař léčí Pavlinu*)

(ii) *léčit₂* ... Act | Pat (Acc)

(ve významu *lékař léčí angínu*)⁵⁴

Obě slovesa *léčit* jsou často rozvíjena volným doplněním prostředku (Means) v instrumentálu (*léčit bylinkami*), *léčit₁* navíc volným doplněním zřetele (Regard) (např. ve větě *Léčil pacienta na chřipku, on však měl angínu.*)⁵⁵

Větší ‚vzdálenost‘ mezi původním substantivem a derivovaným slovesem než v případě odvozování substantiva dokládá možnost společného užití, což u slovesných derivátů obvykle nelze – lze říci *léčit lékem*, ale už ne **navrhovat návrh*. □

Některá další substantiva se ‚slovesným‘ rámcem

Také lexikální deriváty sloves mohou ve svém rámci zachovávat některá slovesná doplnění. Třída lexikálních derivátů je ovšem velmi široká, zde se (na základě zkoumaného materiálu) soustředíme pouze na taková substantiva, která nejsou morfologicky odvozena od sloves, nicméně mají ve svém valenčním rámci ‚slovesná‘ doplnění. Může jít o supletiva (*láska – milovat*), řadíme sem například názvy činností, výsledky činnosti nebo názvy artefaktů (*psát – dopis*).

Příklady:

láska (← *milovat*) ... Pat (k+Dat)

sen (← *zdát se*, ale i *snít*) ... Pat (o+Loc)

dopis (← *psát*) ... Pat (o+Loc) Orig (od+Gen)

článek (← *psát*) ... Pat (o+Loc) Orig (od+Gen)

certifikát na potraviny ... Pat (na+Acc)

zákon o zdraví lidu ... Pat (o+Loc)

zákon na ochranu spotřebitele ... ^Aim (na+Acc)

Doplnění s formou Gen_{subj}/Adj_{pos-subj}, odpovídající na otázku *Čí?*, pokládáme za volná doplnění přináležitosti (App) (viz i Panevová (2001)).

Alternativně, při dějovém chápání těchto substantiv, jsou Gen_{subj} a Adj_{pos-subj} homonymní formy pro aktor, patient nebo přináležitost (App) (viz Hajičová et al (2000)).

Zajímavý jev nastává u substantiva *zákon*, kde je zřejmě možnost kvazivalenčního doplnění účelu spojena s lexikálním obsazením. □

⁵⁴ Rozlišení dvou lexikálních jednotek je motivováno rozdílnými požadavky na prototypické vyjádření pacientu, na jeho sémantické rysy. Takové rozlišení je podpořeno i valenčním rámcem primárního substantiva, sémantické rysy jeho pacientu jsou shodné s požadavky pacientu slovesa *léčit₂*.

⁵⁵ Od slovesa *léčit* je nominalizací odvozeno substantivum *léčení₁* a *léčení₂*:

léčení₁ ... Act Pat (Gen)

léčení₂ ... Act Pat (Gen)

Kvazivalenční doplnění substantiv

Kvazivalenční doplnění substantiv odpovídá druhému a třetímu stupni doplnění substantiv, které charakterizuje Panevová (1966) – zachycuje především **terminologická a frazeologická** spojení vlastní odbornému názvosloví (*tečna v bodě x, souměrnost podle osy x*).

Při zpracování žurnalistických textů se zdá vhodné přistupovat stejným způsobem také k **obvykle užívaným spojením**. Spadají sem například ustálená spojení typická pro ekonomické či právní texty (které přitom nelze označit za odbornou terminologii), je možné pracovat zde se slangem (ve sportovních komentářích) – podle typu textů, které mají být zpracovány.

Dalším typem kvazivalenčních doplnění jsou některá **zdeděná doplnění** – obligatorní volné doplnění slovesa, případně jeho kvazivalenční doplnění se může transformovat v **obvyklé** doplnění substantiva odvozeného od tohoto slovesa.

Všechny tři typy kvazivalenčních doplnění, včetně možného povrchového vyjádření, musí být vyznačeny (a rozlišeny) ve slovníku. Předpokládáme přitom, že kvazivalenční doplnění substantiva může mít několik víceméně synonymních forem (*daň na něco / daň za něco / daň z něčeho*). Značíme je (stejně jako u sloves) symbolem ^ před jménem doplnění.

Příklady:

(IV.24a) *Výběrové řízení na historickou budovu na Brusnici* (PDT, bm102zua.fs #18, zkráceno)

Ve větě (IV.24a) má substantivum *řízení* kvazivalenční doplnění zřetele:

řízení ... ^Regard (na+Acc)

Jde o užití obvyklé v určitých typech textů (ekonomických, právních), které ovšem nelze považovat za terminologii. Ke vzniku takového doplnění viz následující hypotézu (např. ze slovesně jmenného výrazu *vypsat řízení*).

(IV.24b) *Nechtěný pobyt na čerstvém vzduchu nabízí svým žákům Svobodná speciální škola J. A. Komenského v Praze na Kampě.* (PDT, bm102zua.fs #15)

Ve větě (IV.24b) má substantivum *pobyt*, odvozené od slovesa *pobývat*, kvazivalenční doplnění místa (Locative) transformované z obligatorního volného doplnění slovesa.

pobývat ... Act | Kde

pobyt ... Act ^Kde

Takové doplnění se může ztrácet při **terminologizaci** (*je tu hlášen k trvalému pobytu*).

(IV.24c) *Odkud vyšel impuls, abyste poněkud polevili v tlaku na vymáhání částek a hledali, jak je rozložit na delší období?* (PDT, bcc25zua.fs #22)

Věta (IV.24c) ilustruje možnost **dedění** kvazivalenčního doplnění slovesa. Ve větě (IV.20b) jsme rozebírali možné valenční rámce slovesa *tlačit*, druhý rámec zachycoval jeho idiomatické užití:

*někdo*_{Nom} *tlačí*₂ *na někoho*_{Acc} / *na něco*_{Acc}

*tlačit*₂ ... Act | ^Pat (na+Acc)

Tento idiom může být základem pro odvození substantiva:

*někdo*_{Nom} *tlačí*₂ *na někoho*_{Acc} / *na něco*_{Acc}

→ (*něčí*_{Adj-pos/někoho}_{Gen-subj}) *tlak na někoho*_{Acc} / *na něco*_{Acc}

tlak ... Act ^Pat

Pg *na vymáhání (částek)* bude tedy analyzována na základě kvazivalenčního požadavku podstatného jména *tlak* (^Pat). □

Zdá se, že je možné vyslovit následující hypotézu o vzniku kvazivalenčních doplňení substantiv.

Hypotéza:

Kvazivalenční doplnění substantiva může mít původ v roztržení slovesně jmenného výrazu (viz níž). Sloveso, které udává vztah mezi substantivem a daným doplněním (v našem případě předložkovou skupinou), může být eliminováno, dojde k osamostatnění vztahu substantivum – doplnění (Pg). Takové doplnění je potom zachyceno ve slovníku jako kvazivalenční rozvíť substantiva. □

Často jde o doplnění účelu (^Aim) a zřetele (^Regard), transformované z volného doplnění slovesa.

Příklady:

(IV.25a) *V těchto případech na ten přistavený majetek náš fond **vydá akcie**, ty prodá, a ...*
(PDT, bcc25zua.fs #16)

(IV.25b) *Na vybudování nákupního střediska na tomto pozemku, ..., dostal 450 milionový úvěr.*
(PDT, blc13zu.fs #32)

Ve větě (IV.25a) sousloví *vydat akcie na majetek* tvoří slovesně jmenný výraz *vydat akcie*, rozvíjený Pg *na majetek*. Tuto Pg lze chápat jako volné slovesné doplnění vyjadřující účel. Dochází zde k posunu, k ustavení samostatného vztahu substantiva a doplnění – *akcie na majetek*, který opět vyjadřuje účel (a jako takový je zachycen jako kvazivalenční doplnění):

vydat akcie na majetek → *akcie na majetek*

akcie ... ^Aim

Obdobně ve větě ((IV.25b):

dostat úvěr na něco → *úvěr na něco*

úvěr ... ^Aim

Obě dvě spojení, *akcie na něco* a *úvěr na něco* jsou příkladem ustálených, „obvyklých“ spojení typických pro určité druhy textů. □

Závěr:

Valenční a kvazivalenční rámce substantiv lze s úspěchem použít při automatickém zpracování předložkových skupin, pro zachycení případné homonymie. Stejně jako u sloves se tyto rámce používají „jednosměrně“, tj. je-li nalezena Pg (v přípustné slovosledné pozici), která splňuje valenční, příp. kvazivalenční požadavky substantiva, je tato Pg analyzována jako takové doplnění. Skutečnost, že se nepodařilo zaplnit (kvazi)valenční pozice ovšem nic nevypovídá o syntaktické úplnosti (správnosti) věty. □

IV.2.3. Slovesně jmenné výrazy

V kapitole IV.1.2. jsme se odvolávali na termín slovesně jmenný výraz. Identifikace slovesně jmenných výrazů a zachycení jejich valence se zdá být dalším efektivním nástrojem pro určení možných závislostí předložkových skupin. Slovesně jmenný výraz je spojení (primárně) slovesa a substantiva, které tvoří jedinou pojmenovávací jednotku a jehož jmenná část obvykle vyžaduje valenční nebo kvazivalenční doplnění ve formě předložkové skupiny. Tato Pg se potom významově vztahuje jak k substantivu, tak i ke slovesu (sloveso udává vztah mezi substantivem a Pg).

Skutečnost, že předložková skupina rozvíjí celé slovesně jmenné spojení a nemůže být považována za čistě jmenné doplnění, se odráží ve změnách syntaktické struktury takovýchto skupin při derivaci adjektiva (ukážeme na příkladech v kapitole IV.2.4.).

Na TGR jsou u zřejmých idiomů slovesně jmenné výrazy zachyceny jako jediný uzel (což závisí na zpracování těchto výrazů ve slovníku); z čistě syntaktického hlediska (na RPS, kde je sloveso i substantivum reprezentováno samostatným uzlem) lze předložkovou skupinu rozvíjející slovesně jmenný výraz analyzovat dvojím způsobem:

- jako doplnění substantiva, jeho neshodný atribut (potom ovšem pro taková spojení neplatí pozorování (1) a (2) týkající se slovosledných omezení (viz výše, IV.1.1.) – jmenná část takového spojení může stát až za svým doplněním realizovaným jako Pg, substantivum a jeho doplnění mohou být odděleny slovesem (řešeno v IV.1.2.));
- jako slovesné doplnění, a to buď valenční (naplňující příslušné požadavky lexikální jednotky zachycené ve slovníku), nebo fakultativní volné doplnění.

Tyto dvě struktury mají obvykle stejný význam (mají stejné pravdivostní podmínky), ačkoli mají různé reprezentace jak na povrchové rovině, tak na rovině hloubkové (TGR) (pokud zde nejsou u zřejmých idiomů zachyceny jako jediný uzel).

Vysvětlení pojmu (slovesně jmenný výraz):

V kontextu naší práce pokládáme slovesně jmenný výraz za kolokaci sestávající ze slovesa (obvykle s velmi obecným významem) a ze substantiva, které má ‚slovesný‘ valenční, příp. kvazivalenční rámec (tj. má valenční nebo kvazivalenční doplnění, která prototypicky rozvíjejí slovesa). Zabýváme se zde těmi případy, kdy má toto substantivum valenční nebo kvazivalenční doplnění ve formě Pg, sloveso může a nemusí mít takové valenční či kvazivalenční doplnění. Sloveso a substantivum přitom tvoří jedinou pojmenovávací jednotku, substantivum v takovém spojení bývá označováno jako **součást predikátu**. Substantivum má funkci patientu daného slovesa. □

Takováto spojení mohou být alternativou k existujícímu slovesu (*mít zájem o něco / na něčem X zajímat se o něco*), jindy sloveso nebylo vytvořeno (*mít názor na něco, mít šanci na něco*).

Zdá se, že neexistuje ostrá hranice mezi široce dějovými substantivy a substantivy označovanými jako součást predikátu – alternativy ‚sloveso – slovesně jmenný výraz‘ tvoří široké přechodové pásmo.

Příklady:

mít názor / nárok / právo / šanci na něco

názor / nárok / právo / šance ... Pat (na+Acc)

mít peníze / prostředky na něco

peníze / prostředky ... Pat (*na+Acc*)

mít návrh na něco (ale i *navrhovat něco* (viz výš))

návrh ... Pat (*na+Loc*)

mít zájem o něco / na něčem (ale i *zajímat se o něco*)

zájem ... Pat (*o+Acc / na+Loc*)

podat odvolání proti něčemu (ale i *odvolat se proti něčemu*)

odvolání ... Pat (*proti+Dat*)

podat stížnost na něco / proti něčemu (ale i *stěžovat si na něco*)

stížnost ... Pat (*na+Acc / proti+Dat*)

poskytnout / potřebovat / shánět / získat peníze / prostředky na něco

peníze / prostředky ... Pat (*na+Acc*)

stanovit sazbu na něco

sazba ... Pat (*na+Acc*)

poskytnout úvěr na něco

úvěr ... ^Aim (*na+Acc*)

vydat akcie na něco

akcie ... ^Aim (*na+Acc*)

Případné doplnění substantiv s formou Gen_{subj} / Adj_{pos-subj}, které odpovídá na otázku *Čí?*, je zde volným doplněním přináležitosti (App).

Zvláštní zmínku zasluhují kolokace obsahující sloveso *mít* a substantivum s valenčním doplněním ve formě *na+Acc*. Valenční rámec slovesa *mít* nemá pozici pro danou Pg (proto tato spojení řadíme mezi slovesně jmenné výrazy, existuje ovšem frazeologické spojení *mít na něco* (ve smyslu *mít peníze/prostředky/energii/schopnosti ... na něco*, substantivum je vypuštěno, jde o systémovou elipsu, viz Štícha (1997)).⁵⁶ To může při analýze založené na redukci vést ke vzniku neadekvátní závislosti – jako ve větě (IV.2) se spojením *mít nárok na příspěvek*.

Níže navrhované řešení blokuje vznik takových závislostí. □

V následující poznámce přiblížíme, proč v rámci substantiv odvozených od sloves (jako *zájem* (viz příklad (I.1)), *návrh* (viz příklad (IV.23a)) *odvolání* a *stížnost*) neuvažujeme jiná doplnění než pacient a kvazivalenční účel (^Aim) vyjádřená předložkovou skupinou.

Poznámka:

Některá substantiva, která mohou být součástí slovesně jmenného výrazu, jsou deriváty sloves a jako taková dědí jejich valenční rámce. Jsou-li tato podstatná jména použita ve slovesně jmenném výrazu, dochází ke změnám v jejich rámci:

- Tvoří-li substantivum slovesně jmenný výraz, je v jeho valenčním rámci eliminován aktor – ten se ‚přesouvá‘ ke slovesu.
- V rámci substantiva tvořícího slovesně jmenný výraz může ‚zmizet‘ i jiný aktant (např. adresát u substantiva: substantivum *návrh*₁ má v rámci aktor, pacient a adresát (viz (IV.23a)), ale

⁵⁶ Za zaznamenání stojí skutečnost, že zatímco substantiva *peníze*, *prostředky* se vážou s Pg *na+Acc* (ve shodě s vazbou spojení *mít na něco*), podstatné jméno *schopnost* má vazbu jinou – pojí se s infinitivem slovesa nebo jeho nominalizací (*schopnost jednat*, *schopnost jednání*), případně s Pg *k+Dat* (*schopnost k něčemu*), možná vlivem spojení *schopnosti potřebné k něčemu*.

ve slovesně jmenném výrazu *mít návrh* může být rozvíjeno jen patientem vyjádřeným Pg (*na+Acc*).

Změny ve valenčních rámcích substantiv ve slovesně jmenných výrazech zatím nejsou dostatečně zpracovány a popsány, v této práci se jimi dále zabývat nebudeme – na základě dostupného materiálu se zdá, že se tyto změny netýkají doplnění vyjádřených předložkovými skupinami. □

Věty obsahující slovesně jmenný výraz rozvíjený Pg, které se liší pouze ‚zavěšením‘ Pg na substantivum nebo sloveso, označují stejnou situaci. Proto považujeme za vhodné vybrat pouze jednu závislostní strukturu. Preferujeme analýzu Pg jako doplnění substantiva, a tedy valenční požadavky před požadavky slovoslednými.

Návrh řešení:

V případě slovesně jmenných výrazů (která jsou rozvíjena valenčním, příp. kvazivalenčním doplněním ve formě zkoumané Pg), bude vytvořena pouze jedna struktura, se zkoumanou Pg analyzovanou jako doplnění substantiva – ve shodě s jeho (kvazi)valenčními požadavky. Podle pozorování (6) a (7) z kapitoly IV.1.2. může Pg předcházet (v uspořádání daném povrchovým slovosledem) své řídicí jméno a může od něj být oddělena slovesem.

Případný valenční nebo kvazivalenční požadavek slovesa bude zablokován. □

Následující věty ilustrují navržené řešení.

Příklady:

(IV.26a) *Na poslední místo v týmu měl největší šanci některý z mladíků.* (PDT, bl123jsa.fs #20)

Věta (IV.26a) obsahuje slovesně jmenný výraz *mít šanci (na místo)*, rozvíjený Pg *na (poslední místo (v týmu))*.

mít ... Act | Pat (Acc) (Orig(od+Gen / z+Gen))

šance ... Pat (*na+Acc*)

Jmenná část spojení, podstatné jméno *šance*, je patientem slovesa *mít*, přitom má valenční doplnění ve formě Pg. Sloveso *mít* nemá valenční pozici pro danou Pg. Podle navrženého řešení bude Pg analyzována jako neshodný atribut substantiva odpovídající hloubkovému patientu. Vznikne přitom neprojektivní konstrukce.

(IV.26b) *Stát neposkytne na nákup bytů žádné slevy.* (PDT, bcb01aba.fs #16, zkráceno)

Sloveso *poskytnout* tvoří spolu s podstatným jménem *sleva* (pl *slevy*) (patient) slovesně jmenný výraz rozvíjený Pg (*ne*)*poskytnout slevu na nákup*.

poskytnout ... Act | Pat (Acc) Addr (Dat)

sleva ... Pat (*na+Acc*)

Substantivum *sleva* má valenční doplnění (patient), které je vyjádřeno Pg. Sloveso *poskytnout* může být také rozvíjeno zkoumanou Pg, v takovém případě by ovšem šlo o volné doplněním účelu. Pg tedy bude podle navrženého řešení určena jako doplnění substantiva. Opět vznikne neprojektivní konstrukce. K sousloví (*ne*)*poskytnout slevu (na nákup)* se vrátíme v kapitole IV.2.4. v souvislosti s derivací adjektiva *poskytnutý*, příklad (IV.29a).

(IV.26c) *O to nové většinou nemá ani zájem a taky na to většinou nemá peníze.* (PDT, bcc25zua.fs #15)

Spojení (*ne*)*mít peníze (na to)* je slovesně jmenný výraz, jde o sloveso a jeho patient s valenčními rámci:

mít ... Act | Pat (Acc) (Orig(od+Gen / z+Gen))

peníze ... Pat (na+Acc)

Podle navrhovaného řešení bude Pg *na to* určena na základě valenčních požadavků substantiva jako atribut slova *peníze* odpovídající pacientu. Bude vytvořen neprojektivní strom.

(IV.2) *Tyto rodiny mají nárok na státní vyrovnávací příspěvek.* (na základě PDT, b1103zu.fs #42)

Větou (IV.2) jsme ilustrovali potíže s verifikací zachování správnosti v jednotlivých krocích analýzy pracující na principu redukce. Navrhované řešení těmto nesnázím předchází.

Substantivum a sloveso tvořící slovesně jmenný výraz *mít nárok (na příspěvek)* mají následující valenční rámce:

mít ... Act | Pat (Acc) (Orig(od+Gen / z+Gen))

nárok ... Pat (na+Acc)

Pg bude určena pouze jako doplnění substantiva *nárok*. Protože Pg nebude interpretována jako slovesné doplnění (nebude ‚zavěšena‘ na slovesu *mít*), nebudou při rozboru věty podle zásad redukční analýzy realizovány ty větve analýzy, ve kterých by byla vytvořena ‚falešné větná dvojice‘ *mít na příspěvek*. □

Existují slovesně jmenné výrazy, ve kterých substantivum i sloveso vyžadují valenční doplnění (ve svém primárním či přeneseném užití) ve formě zkoumané Pg. Obě lexikální jednotky tedy mají valenční, příp. kvazivalenční požadavky, který lze zaplnit stejnou Pg.

Příklady:

klást důraz na perspektivu podniku (X *zdůrazňovat* ← *důraz*)

klást ... Act | Pat (Acc) Direction

důraz ... Pat (na+Acc)

klást nároky na něco

klást ... Act | Pat (Acc) Direction

nárok ... Pat (na+Acc)

uvalit embargo na něco

uvalit ... Act | Pat (Acc) Direction

embargo ... Pat (na+Acc) □

I v takových případech se tedy zdá vhodné postupovat podle navrženého řešení (předpokládáme, že na RPS je slovesně jmenný výraz zachycen dvěma uzly) – vytváří se jediná závislost Pg na substantivu (v souladu s jeho (kvazi)valenčními požadavky), příslušné požadavky slovesa jsou zablokovány. Jsou uvolněna slovosledná omezení (pozorování (6) a (7)). Následující příklad ukazuje, že uvolňování slovosledných omezení (opodstatněné u slovesně jmenných výrazů), není adekvátní v případech, kdy sloveso i substantivum sice mají (kvazi)valenční požadavky, které lze uspokojit zkoumanou Pg, ale netvoří slovesně jmenný výraz (varianta (IV.27b)).

Povšimněme si další zajímavé vlastnosti takových slovesně jmenných výrazů – daná Pg se (alespoň v některých případech) nemůže vyskytnout v klauzi dvakrát (varianta (IV.27c)).

Příklad:

(IV.27) *Kladl (příliš vysoké) nároky na demokracii / na svého syna / na bezpečnost* (na základě PDT, bm227zua.fs #22, zde (II.23))

(IV.27a) *Na demokracii / na svého syna / na bezpečnost kladl (příliš vysoké) nároky.*

(IV.27b) *Na stůl kladl své návrhy.*

(IV.27c) **Na stůl kladl své nároky na vyšší plat ve formě dopisů.*

Sloveso *klást* tvoří spolu s podstatným jménem *nároky* slovesně jmenný výraz. Sloveso i substantivum mají ve svém valenčním rámci pozici, která může být zaplněna zkoumanou Pg (jednou z prototypických forem volného doplnění místa – směru (Direction) je předložková skupina *na+Acc*):

klást ... Act | Pat (Acc) Direction

nárok ... Pat (*na+Acc*)

Podle navrženého řešení budou ve větě (IV.27) a její variantě (IV.27a) Pg *na demokracii / na svého syna / na bezpečnost* analyzovány jako valenční doplnění substantiva *nárok* (Patient), bez ohledu na jejich slovoslednou pozici.

Ve větě (IV.27b) netvoří sloveso *klást* a substantivum *návrh* slovesně jmenný výraz, budou respektována slovosledná omezení a vytvoří se jediná závislostní struktura, s Pg *na stůl* analyzovanou jako valenční doplnění slovesa (ačkoliv má substantivum *nárok* valenční doplnění, které může být vyjádřené Pg *na+Acc*, viz příklad (IV.23a)).

Nepřijatelnost varianty (IV.27c) lze vysvětlit pozorováním, že slovesně jmenný výraz zachovává už jen valenci substantiva. □

Závěr:

Pokud by byly slovesně jmenné výrazy uspokojivě zachyceny ve slovníku, představovaly by důležitý nástroj pro vyloučení nadbytečných závislostí předložkových skupin, které rozvíjejí takový výraz jako celek (případná adnominální i adverbální analýza Pg mají stejný význam). □

IV.2.4. Valenční rámce adjektiv

V této kapitole se zaměříme především na valenci **adjektiv odvozených od sloves**. Tato adjektiva jsou běžně rozvíjena předložkovou skupinou. Shrneme zde poznatky, které lze využít při analýze závislostí předložkových skupin (viz zejména Panevová (1998b)).

Při odvozování adjektiv od sloves se určitým způsobem zachovává rámec na **tektogramatické rovině** (TGR). Proto je vhodné předpokládat stejný repertoár aktantů a volných doplnění, jako mají slovesa. Dochází ke dvěma regulárním změnám:

- Adjektivum přichází o jedno doplnění z ‚úplného‘ valenčního rámce slovesa – to je zaplněno, ‚obsazeno‘ slovem rozvíjeným zkoumaným adjektivem (obvykle jde o aktor, *zpívající dívka*, nebo patient, případně adresát, *oslovené obecenstvo*).

- Na **povrchové rovině** (RPS) není žádné z obligatorních doplnění povinné, všechna doplnění jsou vypustitelná (se stejnými důsledky jako u vypustitelnosti slovesných a jmenných doplnění – pokud je v přípustné slovosledné pozici nalezen člen splňující valenční požadavky adjektiva,

je vhodné ho analyzovat jako adjektivní doplnění; nepřítomnost obligatorního doplnění neznamená automaticky syntaktickou neúplnost věty).

Následující příklady ilustrují odvození adjektiva a jeho rámce obsahujícího valenční doplnění vyjádřené předložkovou skupinou. Vrátime se k větě (IV.6b), ve které adjektivum sdílí s originálním slovesem valenční i některá volná doplnění. V další větě je zkoumané adjektivum rozvíjeno obligatorními a fakultativními volnými Pg (podobně jako v podrobně rozebírané větě (IV.7) (*rovná na ramínku vystavený kabát*)). Zvláštní pozornost věnujeme zachovávání kvazivalenčních doplnění.

Příklad:

(IV.6b) *Neschopnost způsobila skluz v realizaci autonomie omezené na území Golanských výšin.* (PDT, blc02zu.fs, slovosledná varianta)

Ve větě (IV.6b) byla na základě WOP typu $N_i A_i P_g$ stanovena preferenční závislost *omezený na území (Golanských výšin)* (viz kapitola IV.1.1.) Tato závislost je podporována i valenčním požadavkem deverbativního adjektiva, jehož rámec je odvozen následujícím způsobem.

Adjektivum *omezený* je odvozeno od slovesa *omezit*, které má následující valenční rámec:

omezit ... Act | Pat (Acc) (Eff (na+Acc))

Sloveso *omezit* je často rozvíjeno volným doplněním prostředku (Means, v instrumentálu) a zřetele (Regard, s formou v+Loc).

Adjektivum *omezený* dědí valenční doplnění, 'obsazen' je pacient:

omezený ... Act (Ins) Eff (na+Acc)

$někdo_{Nom}$ *omezuje něco_{Acc}* na $něco_{Acc}$

→ $něco_{Nom}$ *omezené (někým_{Ins})* na $něco_{Acc}$

(Ve významu *Petr omezuje kouření na minimum*, odtud *kouření omezené (Petrem) na minimum*.)

Při derivaci se mohou zachovávat i volná doplnění (opět 'obsazen' pacient):

Means: $někdo_{Nom}$ *omezuje něco_{Acc}* ($něčím_{Ins}$) (na $něco_{Acc}$)

→ $něco_{Nom}$ *omezené někým_{Ins}/něčím_{Ins}* (na $něco_{Acc}$)

(Ve významu *Petr omezuje kouření (žvýkačkami) (na minimum)*, odtud *kouření omezené Petrem (na minimum)*, resp. *kouření omezené žvýkačkami (na minimum)*.)

Regard+Means: $někdo_{Nom}$ *omezuje někoho_{Acc}* (v $něčem_{Loc}$) ($něčím_{Ins}$) (na $něco_{Acc}$)

→ $někdo_{Nom}$ *omezený* (v $něčem_{Loc}$) $někým_{Ins}/něčím_{Ins}$ (na $něco_{Acc}$)

(Ve významu *Petr omezuje Pavla v kouření na minimum (žvýkačkami)*, odtud *Pavel omezený v kouření Petrem.Act / žvýkačkami.Means na minimum*.)

Pro analýzu Pg *na území (Golanských výšin)* ve větě (IV.6b) je podstatný její WOP s preferenčním čtením pro adjektivum – je stanovena závislost *omezený na území (Golanských výšin)*. Adjektivum *omezený* má navíc ve svém valenčním rámci doplnění výsledku (Eff) ve formě Pg *na+Acc*. Zkoumaná Pg *na území (Golanských výšin)* tedy bude s vysokou preferencí určena jako valenční doplnění tohoto adjektiva. (Ostatní možné analýzy budou mít nižší 'váhu' – vrátíme se k této větě a ke stanovení všech přípustných analýz Pg v části V., v příkladu (V.2').)

(IV.28a) *Šiiští fundamentalisté v jižním Iráku a Kurdové žijící převážně na severu země povstali* (PDT, bl104js.fs #47)

Slovosledná pozice Pg (typ **N_iA_iPg**) vede k určení její závislosti na adjektivu, *žijící na severu*. Tuto analýzu podporuje i valenční požadavek deverbativního adjektiva:

Adjektivum *žijící* je odvozeno od slovesa *žít*, které má dva valenční rámce.

(i) $žít_1 \dots \text{Act} | (\text{Pat} (\text{Acc}))^{57}$

(ii) $žít_2 \dots \text{Act} | \text{Locative}$

Od obou sloves lze derivovat adjektivum, od přechodného *žít₁* adjektiva dvě (rozvíjející členy, které vyjadřovali jednak aktant, jednak patient původního slovesa):

(i) $žijící_1 \dots \text{Pat} (\text{Acc})$

$někdo_{\text{Nom}} žije_1 (něco_{\text{Acc}}) \rightarrow někdo_{\text{Nom}} žijící_1 (něco_{\text{Acc}})$

(Ve významu *člověk žije (život)*, odtud *člověk žijící (život)*, často s volným doplněním místa, času nebo způsobu.)

$žitý_1 \dots \text{Act} (\text{Ins})$

$někdo_{\text{Nom}} žije_1 něco_{\text{Acc}} \rightarrow něco_{\text{Acc}} žité_1 (někým_{\text{Ins}})$

(Ve významu *člověk žije (život)*, odtud *život žitý (člověkem)*, často s volným doplněním místa, času nebo způsobu.)

Od slovesa *žít₂* s obligatorním určením místa je odvozeno adjektivum, též s obligatorním volným doplněním místa (,obsazen' aktor):

(ii) $žijící \dots \text{Locative}$

$někdo_{\text{Nom}} žije_2 někde \rightarrow někdo_{\text{Nom}} žijící_2 někde$

Jednou z běžných forem volného určení místa, odpovídajícího na otázku *Kde?* (Locative), je Pg *na+Loc*. Tato forma se při derivaci nemění.

Na základě WOP s preferenčním čtením pro adjektivum je Pg *na severu (země)* ve větě (IV.28a) identifikováno jako doplnění adjektiva. Tato Pg navíc splňuje valenční požadavky tohoto adjektiva. Tato závislost má vysokou preferenci (viz dále, část V.)

(IV.28b) *Ani v jejich případě však není žádným tajemstvím, že prostředky **shromážděné na nadační činnost** jsou ve většině nadací ... používány na* (PDT, bce17zua.fs #50)

Sloveso *shromáždit* tvoří spolu se substantivem *prostředky* slovesně jmenný výraz, který je rozvíjen Pg *na (nadační) činnost*. Tato Pg je podle navrženého řešení ,zavěšena' na podstatné jméno.

Při rozboru klauze *prostředky shromážděné na nadační činnost jsou ve většině nadací ... používány na ...* je nutno respektovat slovoslednou pozici Pg vyjádřenou WOP s preferenčním čtením, Pg *na (nadační) činnost* je tedy analyzována jako doplnění adjektiva. □

Také následující příklady ukazují změny syntaktické struktury, ke kterým dochází při derivaci adjektiva od příslušného slovesa (nezávisle na tom, zda má toto sloveso valenční, příp. kvazivalenční požadavek, který je možné zaplnit Pg).

Příklady:

(IV.29a) *poskytnout slevy_{Acc} na nákup_{Acc} → slevy_{Nom} poskytnuté na nákup_{Acc}*

Ve větě (IV.26b) byl zdůvodněn rozbor Pg jako nominálního doplnění. Odlišná situace nastává u konstrukce s adjektivem odvozeným od slovesa: *poskytnutý* (pl *poskytnuté*) je shodným přívlastkem, který následuje své řídicí substantivum *slevy* (patient původního slovesa); podle

⁵⁷ Tento rámec zachycuje např. řídice se vyskytující spojení *žít život*.

WOP s preferovaným čtením (kapitola IV.1.1.) je Pg *na+Acc* primárně analyzována jako doplnění adjektiva.



Obr.24: Deverbativní adjektivum *poskytnutý* – změny v syntaktické struktuře

(IV.29b) *klást nároky_{Acc} na demokracii_{Acc}* → *nároky kladené na demokracii_{Acc}*



Obr.25: Deverbativní adjektivum *kladený* – změny v syntaktické struktuře

Obdobně v (IV.29b) – při rozboru slovesně jmenného výrazu je Pg ‚zavěšena‘ na substantivum *nárok*, pro adjektivum *kladený* však platí preference daná WOP: pokud je substantivum následováno shodným adjektivem, po kterém následuje Pg, je tato Pg analyzována jako doplnění adjektiva (typ N_iA_iPg). □

Předložkové skupiny mohou rozvíjet i **adjektiva, která nejsou odvozena od sloves** (příklad (II.24)). Valenci primárních adjektiv se zabývá zejména Prouzová (1983) a Panevová (1998b). Lze shrnout:

- Adjektiva, která nejsou odvozena od sloves, mají stejný repertoár aktantů a volných doplnění jako slovesa.
- Slovo, které je rozvíjeno adjektivem, váže/zaplňuje podle typu adjektiva (obvykle) aktor (participium aktivní) nebo pacient (participium pasivní).
- Navíc existují speciální doplnění, která rozvíjejí druhý a třetí stupeň adjektiv (pokud je lze vytvořit).

Kvazivalenční doplnění adjektiv

Zatím jsme se zabývali valenčními rámci adjektiv odvozených od sloves – taková adjektiva určitým způsobem dědí valenční rámec původního slovesa. Zde na několika příkladech ukážeme, že adjektiva odvozená od sloves mohou s původním slovesem sdílet i kvazivalenční doplnění. To není překvapující i vzhledem k tomu, že adjektiva odvozená od sloves mohou s původním slovesem sdílet jeho fakultativní volná doplnění, viz (IV.28b).

Příklady:

(IV.30a) *Vláda schválila na svém zasedání zásady pro odškodňování pozůstalých ... a občanů postižených v této souvislosti na zdraví.* (PDT, bl103js.fs #49)

Adjektivum *postižený* je odvozeno od slovesa *postihnout*, které má následující rámce:

- (i) *postihnout*₁ ... Act | Pat (Acc) ^Regard (na+Loc)
- (ii) *postihnout*₂ ... Act | Pat (Acc)

Můžeme rozlišit dva významy slovesa *postihnout*, v jednom významu může být rozvíjeno kvazivalenčním doplněním zřetele (^Regard) ve formě Pg *na+Loc* (*postihnout na zdraví*), druhé má často volné doplnění prostředku v instrumentálu (Means) (*postihnout (pouhým) okem*).

Možnost odvození adjektiva (a jeho doplnění) reflektuje oba tyto významy:

(i) *postižený*₁ ... Act (Ins) ^Regard (*na+Loc*)

*něco*_{Nom} *postihne*₁ *někoho*_{Acc} *na něčem*_{Loc} → *někdo*_{Nom} *postižený*₁ (*něčím*_{Ans}) *na něčem*_{Loc}

(Ve významu *Petra postihla nemoc (trvale) na zdraví*, odtud *Petr postižený (nemoci) na zdraví*; ,obsazen' patient.)

(ii) *postižený*₂ ... Act (Ins)

*někdo*_{Nom} *postihne*₂ *něco*_{Acc} *něčím*_{Ins} → *něco*_{Nom} *postižený*₂ *někým*_{Ins}

→ *něco*_{Nom} *postižený*₂ *něčím*_{Ins}

(Ve významu *Petr postihl význam (jedinou) větou*, odtud *význam postižený Petrem*, resp. *význam postižený (jedinou) větou*; opět ,obsazen' patient.)

Ve větě (IV.30a) je předložková skupina *na zdraví*_{Acc/Loc} v pozici, která (po redukci Pg v (*této souvislosti*)) odpovídá WOP typu **N_iA_iPg**, kdy je preferován rozbor Pg jako doplnění adjektiva. Původní sloveso *postihnout*₁ má kvazivalenční doplnění zřetele s možným vyjádřením *na+Loc*, které se přenáší i na odvozené adjektivum – Pg tedy rozvíjí adjektivum, *postižený na zdraví*. (Protože jde o kvazivalenční doplnění slovesa, které je uvedeno ve slovníku, objeví se tato informace – v rámci odvozovacích pravidel pro adjektivum – i u adjektiva).

(IV.30b) *majitel zvýhodněný na daních* X *zvýhodněná sazba*

Ve větě (IV.21a) jsme uváděli rámeček slovesa *zvýhodnit*, který obsahuje valenční patient a kvazivalenční zřetel (Regard).

zvýhodnit ... Act | Pat(Acc) ^Regard(*na+Loc*)

Od slovesa *zvýhodnit* je možné odvodit adjektivum, ,obsazen' je patient:

zvýhodněný ... Act (Ins) ^Regard(*na+Loc*)

*někdo*_{Nom} *zvýhodňuje někoho*_{Acc}/*něco*_{Acc} *na něčem*_{Loc}

→ *někdo*_{Acc} *zvýhodněný* (*někým*_{Ins}) (*na něčem*_{Loc})

V případě spojení *majitel zvýhodněný na daních* je realizováno kvazivalenční doplnění zřetele adjektiva (tato analýza je podpořena i WOP typu **N_iA_iPg**, ve kterém je preferována závislost Pg na adjektivu). Ve spojení *zvýhodněná sazba* je aktor všeobecný, zřetel vyjádřen není; patient je ,obsazen' jménem *sazba*, které je adjektivem rozvíjeno. □

Závěr:

Adjektiva odvozená od sloves dědí valenční i kvazivalenční rámeček původního slovesa (s jistými systémovými změnami). Tyto valenční rámce lze, stejně jako u sloves a substantiv, použít pro rozbor možných závislostí předložkových skupin. □

IV.2.5. Vyhodnocování navržených závislostí

Na příkladu slovesně jmenných výrazů jsme viděli, že může docházet ke křížení valenčních požadavků (viz třeba varianty věty (IV.27), *klást nároky na demokracii*), sloveso i substantivum mohou vyžadovat doplnění se stejnou formou (v našem případě Pg). Tato situace

je ovšem obecnější, není omezena na slovesně jmenná vyjádření – křížení valenčních a kvazivalenčních požadavků se obecně může týkat všech sloves, substantiv i adjektiv a jejich kombinací.

Věty, které se liší analýzou předložkové skupiny (ve kterých Pg rozvíjí různé větné členy), jsou obecně víceznačné. Fakt, že takové věty mohou mít stejný význam (tj. mohou se ve všech kontextech shodovat v pravdivostních podmínkách), je dán jejich kognitivním obsahem, ne jejich jazykovou strukturací. To vede k nutnosti vytvořit všechny závislostní struktury, které (s ohledem na slovosledná omezení) vyhovují všem valenčním požadavkům. Některé z těchto struktur se dají vyloučit na základě doplňujících údajů, především údajů o slovesně jmenných výrazech.

Nicméně typy valenčních požadavků, které vedou k vytvoření možných závislostí (v našem případě) předložkových skupin, mají různou ‚váhu‘, umožňují tedy určit spolehlivost jednotlivých kroků analýzy založené na principu redukce. Na základě klasifikace valenčních doplnění lze stanovit preference pro vytvořené závislostní reprezentace zkoumané věty, které odrážejí ‚pravděpodobnost‘ jednotlivých čtení pro danou Pg.

Klasifikace jednotlivých typů valenčních požadavků

Tato klasifikace valenčních požadavků slouží pro ohodnocení výsledných syntaktických struktur zkoumané věty, pro stanovení spolehlivosti možných valenčních závislostí Pg.⁵⁸

1. ‚Klasické‘ valenční rámce sloves (včetně rámců zachycujících jejich přenesené a posunuté užití, tj. kvazivalenční doplnění prvního typu) mají nejvyšší prioritu při určování možné závislosti konkrétního doplnění (v našem případě Pg).
2. Následují valenční požadavky adjektiv (pokud jsou splněna slovosledná omezení).
3. Valenční rámce substantiv (s jednoznačně určenou formou) má nejnižší prioritu ze standardních valenčních doplnění; jejich uplatnění pro netypickou a neprojektivní pozici Pg je podmíněno nepřítomností valenčních požadavků slovesa.
4. Následují doplnění daná kvazivalenčními požadavky jednotlivých slovních druhů:
 - kvazivalenční doplnění substantiv,
 - kvazivalenční doplnění sloves (druhý typ) a
 - kvazivalenční doplnění adjektiv.
5. Nejnižší preferenci mají volná doplnění.

Bereme ovšem v úvahu vypustitelnost valenčních doplnění na povrchové rovině. □

Následující věty ilustrují křížení valenčních požadavků s různým stupněm priority. S obdobou věty (IV.31a) jsme už pracovali (je modifikací věty (IV.1)), zde uvádíme slovosledné varianty a modifikace, ve kterých se projevuje navržená klasifikace.

Příklady:

(IV.31a) (*Domy jsou vraceny často v žalostném stavu*), ***zákon dbá na ochranu nájemníků snad až příliš.*** (PDT, bl101js.fs #18, zkráceno)

⁵⁸ Navržená klasifikace valenčních doplnění různých slovních druhů v této fázi nereflexuje kritéria založená na povrchovém slovosledu. Výsledné uspořádání, které bere v úvahu všechny typy zkoumaných kritérií, je uvedeno v V. části práce, věnované uplatnění a evaluaci navržených kritérií.

(IV.31a') (Podle toho, co jsme se dozvěděli), **dbá zákon na ochranu nájemníků** snad až příliš.

(IV.31a'') (Dbají na tato omezení zákony?) **Zákon na ochranu nájemníků dbá** snad až příliš.

Budeme zkoumat druhou část věty (IV.31a) a její varianty (IV.31a') a (IV.31a''). Sloveso *dbát* a substantivum *zákon* vyžadují valenční, resp. kvazivalenční doplnění ve formě Pg:

*dbát*₁ ... Act | Pat (na+Acc / o+Acc)

*dbát*₂ ... Act | Pat (Gen)

zákon ... Pat (o+Loc)

zákon ... ^Aim (na+Acc)

Předložková skupina *na ochranu (nájemníků)* tedy splňuje valenční požadavky slovesa *dbát*₁ (Patient, nejvyšší priorita) a požadavek podstatného jména *zákon* na ‚obvyklé‘ doplnění (^Aim, nižší priorita).

V původní větě (IV.31a) slovosledná omezení blokují analýzu Pg jako nominálního doplnění (pozorování (2) z kapitoly IV.1.1. – Pg nemůže rozvíjet substantivum, které je odděleno slovesem; valenční požadavek slovesa nedovoluje uvolnit podmínku projektivnosti, nejsou splněny předpoklady pozorování (10) v IV.1.2.). Pg *na ochranu (nájemníků)* je tedy analyzována jako patient slovesa *dbát*.

Ve větách (IV.31a') a (IV.31a'') slovosled dovoluje adverbální i adnominální analýzu Pg (‚podezřelé‘ WOP typu **VNPg** a **NPgV**). Na základě valenčních, resp. kvazivalenčních požadavků budou vytvořeny obě závislosti, *dbá na ochranu (nájemníků)* (Patient) a *zákon na ochranu (nájemníků)* (^Aim), s prioritou valenčního, resp. kvazivalenčního doplnění. (Kvůli obligatornosti slovesného doplnění na TGR je analýza Pg jako doplnění substantiva ve většině kontextů nepřijatelná. Nicméně např. užití v dialogu (IV.31a'') vytvoření takové závislosti ‚ospravedlňuje‘.)

(IV.31b) *Byly na obce převedeny i nějaké akcie?* (PDT, bcc25zua.fs)

(IV.31b') *Byly převedeny nějaké akcie na obce?*

(IV.31b'') *Byly převedeny nějaké akcie na majetek?*

Sloveso *převést* a substantivum *akcie* vyžadují valenční, resp. kvazivalenční doplnění, které může mít formu *na+Acc*:

převést ... Act | Pat Direction

akcie ... ^Aim (na+Acc)

V (IV.31b) je uplatněn valenční požadavek slovesa na valenční doplnění místa – směru (Direction), je vytvořena závislost *na obce převést*. Vznik neprojektivní struktury, ve které byl uspokojen kvazivalenční požadavek substantiva *akcie*, je blokován valenčním požadavkem slovesa (nelze uplatnit pozorování (9) a (10)).

V (IV.31b') a (IV.31b'') slovosled dovoluje adverbální i adnominální analýzu Pg (WOP typu **VNPg**), budou tedy vytvořeny vždy dvě závislostní struktury – *převést na obce*, resp. *převést na majetek* (s vyšší preferencí) a *akcie na obce*, resp. *akcie na majetek* (s nižší preferencí). Skutečnost, že v případě (IV.31b') je nutno dát přednost adverbální analýze Pg (*převedeny na obce*), zatímco v (IV.31b'') jde spíše o doplnění substantiva (*akcie na majetek*), je dána konkrétním lexikálním obsazením zkoumaných vět.⁵⁹

⁵⁹Viz i obsahová nepřijatelnost syntakticky správně utvořené věty **Byly na majetek převedeny i nějaké akcie?*

(IV.31c) **Rozšíření silniční daně na všechna vozidla** - to je otázka, o níž podnikatelé nejčastěji hovoří. (PDT bcb01aba.fs #37, shortened)

Deverbativum (syntaktický derivát) *rozšíření* dědí valenční rámec originálního slovesa *rozšířit*; substantivum *daň* vyžaduje kvazivalenční doplnění. Valenční požadavky mohou být naplněny Pg *na+Acc*:

rozšířit ... Act | Pat (Acc) Eff (*na+Acc*)

rozšíření ... Act (Adj_{pos}/Gen) Pat (Gen) Eff (*na+Acc*)

daň ... ^Pat (*na+Acc/za+Acc/z+Gen*)

Pg *na* (*všechna*) *vozidla* bude určena jednak jako doplnění deverbativa *rozšíření* (s vyšší preferencí, naplňuje valenční požadavky), jednak jako doplnění podstatného jména *daň* (s nižší mírou spolehlivosti, která je dána kvazivalencí substantiva). □

Poznámka:

Na tomto místě je vhodné znovu upozornit na různorodost doplnění, která jsou zachycena pomocí kvazivalence. Spolehlivost rámců sloves s posunutým nebo přeneseným užitím (první typ kvazivalenčních doplnění sloves) odpovídá spolehlivosti valenčních doplnění. Naopak, druhému typu kvazivalenčních doplnění sloves jsou nadřazena kvazivalenční doplnění substantiv (v souladu s řešením navrženým pro slovesně jmenné výrazy). □

„Sílu“ rámce slovesa v přeneseném užití dokládá i následující příklad (IV.32), který obměňuje věty (IV.23c) a (IV.20c).

Příklad:

(IV.32) *Japonský lék nepomohl na kanadské problémy.* (jde o větu (IV.20c))

(IV.32') *Japonský lék na kanadské problémy nepomohl.*

Substantivum *lék* má ve svém valenčním rámci patient s formou *na+Acc* (viz (IV.23c)), sloveso *pomoci* může být rozvíjeno kvazivalenčním patientem (první typ kvazivalenčního doplnění sloves, viz (IV.20c)):

lék ... Pat (*na+Acc/proti+Dat*)

*pomoci*₂ ... Act | ^Pat (*na+Acc*)

V případě slovesa *pomoci* jde o kvazivalenční doplnění, není možné uvolnit slovosledná omezení (kvazivalenční požadavek slovesa stačí k zablokování neprojektivní adnominální analýzy). Na základě WOP typu **VPg** je vytvořena jediná struktura, *(ne)pomoci na (kanadské) problémy*.

V (IV.32') se uplatní obě analýzy Pg, WOP typu **NPgV** – *(ne)pomoci na (kanadské) problémy a lék na (kanadské) problémy*. Vytvoření obou závislostních struktur je přitom důležité, neboť jde o větu homonymní, různé pravdivostní podmínky jednotlivých čtení lze zachytit pomocí možných pokračování:

Japonský lék na kanadské problémy nepomohl (Kanadě, pomohl však na problémy Británie.)

(s jiným výběrem lexikálních jednotek – *Lék na kašel nepomohl, teplotu však snížil.*)

Japonský lék na kanadské problémy nepomohl (Británii, přestože Kanadě pomohl.)

(s jiným výběrem lexikálních jednotek – *Lék na kašel nepomohl, spíš pomohlo, že zůstal v posteli.*) □

Typickým případem křížení různých typů valenčních požadavků jsou slovesně jmenné výrazy – podrobněji jsme se jim věnovali v kapitole IV.2.3. Protože slovesně jmenné výrazy tvoří jedinou pojmenovávací jednotku, stanovujeme (jedinou) výjimku z požadavku vytváření všech struktur, které reflektují valenční a kvazivalenční požadavky.

Výjimka:

Slovesně jmenné výrazy (pokud jsou uvedeny ve slovníku) tvoří jedinou výjimku v navržené klasifikaci – protože význam vět obsahujících takové výrazy nezávisí na analýze Pg (nominální / slovesné doplnění), navrhuje vytvořit jedinou strukturu s Pg ‚zavěšenou‘ jako valenční, resp. kvazivalenční doplnění substantiva. Případný valenční nebo kvazivalenční požadavek slovesa je zablokovan. □

Tato výjimka ovšem platí pouze pro slovesně jmenné výrazy – pokud sloveso a substantivum netvoří jedinou pojmenovávací jednotku, je potřeba vytvořit obě závislostní struktury, neboť mohou mít různé pravdivostní podmínky (jak bylo vidět ve větě (IV.32), s dvojicí *lék na (kanadské) problémy X na problémy (ne)pomoci*).

Závěr:

Věty, ve kterých dochází ke křížení valenčních požadavků, jsou (pokud slovosledná omezení neblokují některá čtení) vždy ‚podezřelé‘ z hlediska víceznačnosti jejich struktur. Klasifikace valenčních doplnění umožňuje ohodnotit získané závislostní struktury lišící se analýzou předložkové skupiny, na jejím základě je možné určit spolehlivost jednotlivých rozborů. K podmínkám úplnějšího řešení ovšem patří i co nejsoustavnější zpracování frazeologických spojení s různým stupněm těsnosti. □

Shrnutí poznámek k valenci

Valenční požadavky sloves, substantiv a adjektiv hrají klíčovou roli při určování závislostí předložkových skupin a pro zachycení jejich morfemtické homonymie. Snažíme se přitom o striktní rozlišování valenčních rámců na rovině podkladové syntaxe a jejich povrchového (resp. morfemtického) vyjádření.

Pro účely automatického zpracování Pg rozšiřujeme koncept valenčních rámců o terminologická a ‚obvykle užívaná‘ spojení, tj. o tzv. kvazivalenční doplnění sloves, substantiv i adjektiv.

Zvláštní roli hrají slovesně jmenné výrazy, které (pokud jsou i s požadovaným valenčním doplněním zachyceny ve slovníku) vedou k odstranění některých redundantních závislostí předložkových skupin.

Ukázalo se, že valenční požadavky jednotlivých slovních druhů mají různou ‚váhu‘, což v případě křížení různých požadavků umožnilo stanovit preference pro jednotlivá čtení.

Pro adjektiva jsou důležité též preference dané slovosledem – pro WOP s preferovaným čtením pro adjektiva je vhodné upřednostnit závislosti vytvořené na základě WOP před požadavky valenčními.

IV.3. SÉMANTIKA

Zabývá-li se člověk jako adresát promluvy syntaktickou analýzou vět, využívá své znalosti jazyka, tedy jeho slovní zásoby, morfologie a syntaxe. Kromě toho se opírá o široký kontext promluvy (textu), a také o informace mimojazykové, vyplývající ze znalosti sémantiky a pragmatiky. To se v plné míře týká i rozboru předložkových skupin a stanovení jejich případné morfematické homonymie. Uvažujeme-li proto o možnosti automatického zpracování Pg, je potřeba určit, jakým způsobem a v jaké míře zahrnout sémantické údaje.

Byly navrženy systémy pro automatickou syntaktickou analýzu přirozeného jazyka, které se vyhýbají využívání sémantické informace. Pro češtinu jde např. o výše zmíněný Robustní parser,⁶⁰ kde je tato skutečnost dána konkrétním úkolem (rozpoznáváním syntaktické správnosti věty), pro který je parser koncipován. Jiné systémy předpokládají využití určitého typu sémantických údajů, především údajů uchovávaných jako sémantické rysy jednotlivých lexikálních jednotek; pro češtinu jsou to např. už jmenované parsery pro projekty RUSLAN (Oliva (1989)) a LATESLAV (Oliva (1996)), podrobnější popis sémantických rysů však nebyl publikován.

Protože pro účely jednotlivých projektů postupně vzniká rozsáhlý slovník obsahující rovněž sémantické rysy jednotlivých lexikálních jednotek (návrh hesel uvádí Skoumalová (1997)), zdá se účelné pokusit se těchto rysů využít pro určování závislosti předložkových skupin.

Sémantické rysy všech potenciálních řídicích členů zkoumané Pg, sémantické vlastnosti předložky a sémantické rysy substantiva tvořícího (spolu s předložkou) jádro Pg slouží jako pomocné kritérium při vyhodnocování možných závislostí Pg, resp. při stanovení preferencí pro jednotlivá čtení vět s homonymní Pg. Na jejich základě lze navýšit ‚váhu‘ jednotlivých analýz. Těchto sémantických rysů se využívá především pro identifikaci volných doplnění, a to obligatorních (a tedy valenčních) i fakultativních.

Metoda pro návrh sémantických rysů substantiv

Základem pro návrh sémantických rysů podstatných jmen se staly výsledky popsané Buráňovou (1980). E. Buráňová navrhuje metodu pro sémantickou klasifikaci substantiv a pro určování jejich sémantických rysů založenou na zkoumání, jaké funkce na TGR (vzhledem k predikátu) mohou zastávat substantiva s daným sémantickým rysem.⁶¹ Podstatným znakem této metody je fakt, že pracuje s čistě lingvistickými kritérii, nespolehá se na kritéria obsahová. Substantiva s jednotlivými sémantickými rysy byla použita (jako slovesná doplnění) v různých pádech a s různými předložkami. Určila se nejpravděpodobnější ‚primární funkce‘ jednotlivých forem (viz Kuryłowicz (1936); v úvahu se bral pouze základní význam slova). Na základě dalšího zpřesňování rysů byla sestavena tabulka ‚primárních‘ tektogramatických funkcí substantiv s navrženými rysy v některých předložkových pádech. Jako podkladový materiál byly použity elektrotechnické texty.

⁶⁰ Více viz oddíl VI.1.

⁶¹ Vychází přitom z výzkumu, který sloužil pro vytváření tzv. instrukčních sítí pro určování významových funkcí jednotlivých pádů (tyto algoritmy byly navrženy v letech 1974-76 pod vedením P. Sgalla a J. Panevové).

Zdá se, že tento způsob klasifikace sémantických rysů, navržený pro určování funkcí předložkových skupin na TGR, lze (s mírnými modifikacemi) použít i pro návrh automatického zpracování morfemické homonymie předložkových skupin, tj. pro práci na rovině povrchové syntaxe, pro analýzu povrchové realizace věty.

IV.3.1. Návrh sémantických rysů pro substantiva

Pro určení sémantických tříd substantiv využíváme kritéria navržená Buráňovou (1980), s jednou podstatnou změnou – popsaná metoda předpokládá, že je dána syntaktická struktura věty (zkoumané substantivum je přitom uvažováno jen jako doplnění slovesa), zatímco při vyšetřování závislosti Pg je nutno syntaktickou strukturu věty určit spolu s typem doplnění, které Pg vyjadřuje. Další změny s sebou nese také výběr textů – metoda byla navržena pro zkoumání elektrotechnických textů, my její kritéria aplikujeme na texty žurnalistické.

Jak vyplývá z popisu metody, pracuje se s jednotlivými předložkovými skupinami, na jejichž základě se stanoví soubor sémantických rysů pro substantiva. V této kapitole se pro ilustraci soustředíme na dvě předložkové skupiny, *na+Acc* a *na+Loc*, na určení relevantních sémantických rysů a na ‚primární funkce‘ substantiv s navrženými rysy (využíváme přitom substantiv z vět vybraných z PDT, viz část III.).

Předložkové skupiny *na+Acc* a *na+Loc* mohou mít (kromě vnitřních valenčních doplnění)⁶² následující funkce (viz Šmilauer (1966)) :

<i>na+Acc</i>		
určení místa:	Kam?	<i>Tonča hodila tašku <u>na lavici</u>.</i>
určení času:	Kdy? Na jak dlouho?	<i><u>na druhý den</u>, <u>přesně na den</u> <u>na tři dni</u></i>
určení způsobu:	přímo účinkem výsledkem míra přímo míra omezením zřetel	<i><u>zavřít na klíč</u> <u>nabarvit na růžovo</u> <i>Sklo se rozsypalo <u>na padrt'</u>.</i> <i>(vrátit) peníze <u>na groš</u>, bylo nás <u>na sto</u></i> <i>Všechno zapomněla až <u>na slovo</u> <u>kišasonka</u>.</i> <i><u>Na doktora</u> píše čitelně.</i></i>
určení příčiny:	pomůcka podmínka účel příčina	<i><u>přisahat na kříž</u> <u>vydat na potvrzení</u> <u>na zahřátí</u> <u>zemřít na tuberkulózu</u></i>
parenthese		
<i><u>na mou duši</u></i>		
<i>na+Loc</i>		
určení místa:	Kde?	<i>Taška ležela <u>na lavici</u>.</i>
určení času:	Kdy?	<i><u>na úsvitě</u>, <u>na konci</u></i>
určení způsobu:	pomůcka zřetel	<i><u>Chystali se odjet na žebříňáku</u>. <u>prospívat na duchu</u>.</i>
určení příčiny:	účel	<i><u>být na mřenkách</u>, <u>na jahodách</u>, <u>na lovu</u></i>

⁶² Např. *patient v myslit na někoho/na něco*.

Primární funkce Pg podle sémantických rysů substantiv

Na podkladovém materiálu vět obsahujících Pg *na+Acc* a *na+Loc* jsme vyšetřovali následující soubor sémantických rysů (vzhledem k daným předložkám). Substantiva zastávala ve zkoumaných větách uvedené tektogramatické funkce vyjádřené Pg *na+Acc/Loc*. U jednotlivých substantiv zároveň uvádíme, zda byla zkoumaná Pg určena jako nominální nebo verbální doplnění (obsah závorcky odpovídá vždy jednomu výskytu Pg; lomítko označuje možnost ‚zavěšení‘ na různé větné členy, tedy morfematickou homonymii).

‚human‘

prezident (Pat, N);

žák (na koho – distribut., N)

‚instituce‘ → kde (Loc), kam (Acc)

ministerstvo (kde, N / N) (kde, N) (kam, V);

gymnázium (kde, N);

odbor (kde, N);

úřad (kam, V) (kde, V/?N);

obce (kam/na koho, N);

magistrát (kam/na koho, V)

‚instituce+abstr‘ → kde (Loc), kam (Acc)

generace (kde/na kom, V);

veřejnost (kam, V)

‚konkr‘ → kde, kam, resp. nástroj

budova (zřetel, V / kde, N);

prostředek (=dopr.) (kde, N) (kde, N / N) (kam, V);

vozidlo (kam, V) (kde, N / zřetel, N);

počítač (nástroj, N);

lyže (nástroj, N);

stůl (kde, N);

videokazeta (kde, V);

farma (kam, V);

sedátko (kde, V)

‚místo‘ → kde (Loc), kam (Acc)

ulice (kde, V);

trávník (kde, V);

post (kde, N);

místo (kde, V) (Pat, N / zřetel, V – frazém);

území (kde, N / N / N / V / kam, V);

předměstí (kde, V);

zastávka (kde, V / N / N);

chodba (kde, V)

‚čas‘ → kdy (na kdy), na jak dlouho

neděle (frazém-kdy, N);

rok (na jak dlouho, V);

období (kdy, V);

doba (na jak dlouho, V);

okamžik (na jak dlouho, V)

pátek (kdy, V)

poledne (kdy, V)

„událost“ → kam (Acc), kde (Loc)

MS (=mistrovství světa) (kam, N / kde, N / V);
zasedání (kde, V);
konference (kde, V);
kurz (kde, V / N / NaN);
brífing (kde, V);
závod (kde, V);
jednání₁ (kam, V)

„děj“ (činnost, výsl. činnosti) → účel, zřetel, resp. kde

nákup (účel, N / V);
výkyv (zřetel, N);
uveřejnění (Pat, N / zřetel, V – frazém);
vývoj (účel, V);
ochrana (účel, N);
slunění (účel, N);
náhrada (účel, V);
oprava (účel, N / V);
hostování (???, V);
program (kde, V);
činnost (účel, V);
nákup (Pat, N / účel, V);
hostování (kde, V)
programu jsou (kde, V - frazém)

„přír. jevy“ (skupenství) → kde (Loc), kam (Acc)

vzduch (kde, N);
moře (kde, V)

„part“ (část) → podle dalšího rysu (konkr., místo, čas): kde, kam, resp. kdy, na kdy

povrch (kam, V) (kam, V);
stránka (kde, V);
pozadí (kde, V);
výčnělky (kam/na co, V);
hrot (kde, N / V);
začátek (kdy, V);
přelom (na kdy, V) (kdy, V)

„jméno“ → podle dalšího rysu (human, instit., událost, čas, místo)

Dukla (kam, V);
Collège Bart (kde, V / N / N);
Turné čtyř můstků (účel, N);
Intersportturné (Pat, N) (kde, V);
Nový Rok (kdy, V);
Kampa (kde, N / V);
Brusnice (kde, N / V);
Slovensko (kde, V) (kde, N / N / V);
Rus (na Rusi) (kde, N);
Prostějovsko (kde, N / N / V);
Štrbské Pleso (kde, N);
Masarykova ulice (kde, V / N / N);
Jírova ulice (kde, V / N / ?N / N);
Riviéra (kde, V);
Oxford Street (kde, V);
Morava (kde, V);

‚kvant‘ → distribuce k V, resp. míra, přibližnost

vše (Pat, N);
desítky (distribut., V);
stovky (distribut., V);
150 (přibl., V);
sedmdesát (přibl., V);

‚abstr‘ → kde (Loc), kam (Acc), resp. jak, zřetel

trh (kde, N / N / V);
svoboda (kde, N);
patent (zřete, NI);
akcie (kam, N);
struktura (kam, N);
liga (kam, V);
bezvýslednost (Pat, V);
majetek (zřetel, N / V);
úroveň (kde, V);
světlo (kam, V);
dálka (jak/kam, V - frazém);
fronta (kde, V)

‚modalita‘

možnost (Pat, V)

Navržená klasifikace substantiv se víceméně přidržuje sémantických tříd navržených Buráňovou, dochází jen k několika změnám:

- třída substantiv s rysem **‚konkr‘** je užší než třída předmětů (pouze předměty, „na které lze sáhnout“);
- třída substantiv s rysem **‚místo‘** zahrnuje kromě prostorových objektů i např. zeměpisná určení;
- zavádí se třída substantiv s rysem **‚instituce+abstr‘** (abstraktní pojmy, které označují skupiny lidí), **‚událost‘** (tato podstatná jména mají často též rys ‚děj‘), **‚part‘** (část), **‚jméno‘** a **‚modalita‘**;
- z hlediska určování závislostí na povrchové rovině se nezdá nutné dále rozlišovat následující skupiny substantiv:
 - třída substantiv s rysem **‚děj‘** zahrnuje substantiva označující činnost a výsledek činnosti; syntakticky jde o jména odvozená od sloves; z hlediska určování závislostí se nezdá nutné dále je rozlišovat;
 - třída substantiv s rysem **‚abstr‘** je velmi široká, Buráňová ji dále dělí na substantiva označující fyzikální a intelektuální kategorie, vlastnosti a společenská ustanovení; pro určování závislostí na povrchové rovině se nezdá nutné dále je rozlišovat;
- substantiva mohou mít v jednom významu několik rysů (to se týká zejména rysů part a jméno).

Závěr:

Na tomto místě je nutné poznamenat, že na základě zkoumání jediné předložky tvořící dvě Pg není možné prokázat účelnost všech navržených rysů substantiv. Navíc se v podkladovém

souboru vět substantiva s některými rysy ve zkoumaných Pg nevyskytla buď vůbec, nebo pouze v jednom nebo dvou případech.

Nicméně jsme v této kapitole ověřili možnost využití metody klasifikace sémantických rysů pomocí funkčních kritérií, a to i bez předpokladu znalosti syntaktické struktury věty. □

IV.3.2. Návrh pravidel využívajících sémantické rysy substantiv

Jak už bylo řečeno výše, pokud některý z valenčních, příp. kvazivalenčních rámců obsahuje volné doplnění, zejména času a místa, je nasnadě využití sémantických kritérií pro určení takového doplnění.

Navíc, vyskytuje-li se v analyzované větě předložková skupina (která není určena jako valenční, případně kvazivalenční doplnění některého slovesa, substantiva nebo adjektiva) na ‚vhodné‘ slovosledné pozici (viz slovosledné vzorce, oddíl IV.1. a formální kritéria, IV.4.), je možné aplikovat pravidla využívající sémantických rysů potenciálních řídicích členů zkoumané Pg, sémantických vlastností předložky a sémantických rysů substantiva tvořícího (spolu s předložkou) jádro Pg také pro určení této Pg jako volného doplnění.

V následujících odstavcích formulujeme (a na příkladech ilustrujeme) některá pozorování, která pracují se sémantickými rysy substantiv a umožňují rozhodnout o závislosti Pg na předcházejícím substantivu. Na základě takových pravidel lze stanovit preferované čtení Pg. Omezíme se na předložky *na*, *v* a *do*, které mají převážnou většinou místní (a časový) význam. První dvě se pojí s akuzativem a lokálem, tvoří tedy předložkové skupiny *na+Acc*, *na+Loc*, *v+Acc* a *v+Loc*,⁶³ třetí se pojí s genitivem, *do+Gen*. Soustředíme se především na substantiva označující reálné předměty (s rysem ‚**konkr**‘) a na substantiva s rysem ‚**místo**‘ a ‚**čas**‘ (vzhledem k primárním funkcím předložek), případně i na substantiva s dalšími rysy, jejichž primární funkcí je (pro předložky *na*, *v* a *do*) určení místa a času.

Uveďme předem, že ani v těchto případech, ve kterých se využití sémantických rysů přímo nabízí, není analýza Pg na základě uvedených pozorování zcela spolehlivá.

Místní určení vyjádřené vlastním jménem

Následující pozorování odráží předpoklad, že podstatná jména s rysem ‚**konkr**‘ jsou obvykle prostorově omezená, a místní určení vyjádřené vlastním jménem je situují ve světě. Jde tedy o ‚užší‘ určení místa následované určením ‚širším‘. (Na tomto rozlišení jsou založena i pozorování v dalším odstavci.)

Pozorování: (pro substantiva s rysy ‚**konkr**‘ a ‚**místo**‘ + ‚**jméno**‘)

(14) Následuje-li po substantivu s rysem ‚**konkr**‘ (v prostém nebo předložkovém pádu) jedna z předložkových skupin *na+Acc*, *na+Loc*, *v+Loc* nebo *do+Gen*, ve které substantivum nese sémantické rysy ‚**místo**‘ a ‚**jméno**‘, potom zkoumaná Pg obvykle rozvíjí bezprostředně předcházející podstatné jméno. Pg je v takovém případě jeho neshodným atributem. □

⁶³ Nezabývali jsme se zde podrobně předložkovými skupinami *v+Acc*, *v+Loc* a *do+Gen*. Vzhledem k podobným syntaktickým i sémantickým vlastnostem předložek *na*, *v* a *do* lze předpokládat (viz i Šmilauer (1966), Buráňová (1980)), že bychom dostali analogické výsledky jako pro Pg *na+Acc* a *na+Loc*; Pg *v+Acc* udává především čas (Kdy?).

Příklad:

(IV.33) *Ministerstvo školství vypsalo výběrové řízení na historickou budovu na Brusnici.* (PDT, bm102zua.fs #18, zkráceno)

Pg *na (historickou) budovu (na+Acc)* je na základě kvazivalenčního požadavku substantiva *řízení* analyzována jako adnominální doplnění (zřetel, ^Regard). Substantivum *budova* má sémantický rys ‚konkr‘, *Brusnice* má rysy ‚místo‘ a ‚jméno‘. Podle pozorování (14) je preferována analýza Pg *na Brusnici (na+Acc/na+Loc*, fonologická homonymie) jako volného doplnění substantiva *budova* (Kde? pro *na+Loc*; Kam, příp. také účel, Aim pro *na+Acc*). Další možné analýzy, *řízení na Brusnici* a *vypsat na Brusnici*, budou klasifikovány jako méně spolehlivé (v obou případech jde o volná doplnění, jejichž pravděpodobnost, ‚váha‘, není navýšena sémantickými preferencemi). □

Pozorování: (pro substantiva s rysy ‚místo‘ a ‚místo‘ + ‚jméno‘)

(15) Následuje-li po substantivu s rysem ‚místo‘ (v prostém nebo předložkovém pádu) jedna z předložkových skupin *na+Acc*, *na+Loc*, *v+Loc* nebo *do+Gen*, ve které substantivum nese sémantické rysy ‚místo‘ + ‚jméno‘, potom zkoumaná Pg obvykle rozvíjí předcházející podstatné jméno. Pg je v takovém případě jeho neshodným atributem. □

Příklad:

(IV.34) *Nový školní rok musí zahajovat na trávníku parku na Kampě.* (PDT, bm102zua.fs #6, zkráceno)

Substantivum *Kampa*, které tvoří s předložkou *na* Pg *na+Loc*, má sémantické rysy ‚místo‘ a ‚jméno‘. Bezprostředně předcházející podstatné jméno *park* je označeno jako ‚místo‘ – podle pozorování (15) bude preferována atributivní analýza *park na Kampě*. Analýzy *trávník na Kampě* a *zahajovat na Kampě* budou upozaděny. □

Pozorování: (pro dvě substantiva s rysy ‚místo‘ + ‚jméno‘)

(16) Vyskytnou-li se bezprostředně za sebou dvě substantiva se sémantickými rysy ‚místo‘ + ‚jméno‘, přičemž druhé z nich je částí jedné z předložkových skupin *na+Acc*, *na+Loc*, *v+Loc* nebo *do+Gen*, potom je jejich vzájemný vztah dán kognitivním obsahem věty. Tato mimojazyková znalost může být zachycena pomocí velmi jemné a podrobné sémantické reprezentace daného úseku světa:

- následuje-li po širším určení místa určení bližší, jde obvykle o dvě jednotky, které rozvíjejí stejný větný člen, tj. jsou vzájemně syntakticky nezávislé;

- následuje-li po bližším určení místa určení širší, potom druhé jméno obvykle rozvíjí předcházející jméno, tvoří jeho neshodný atribut;

- pokud tato dvě určení místa nelze porovnat vzhledem k inkluzi, jde o dvě nezávislé jednotky, obvykle se společným řídicím členem. □

Příklad:

(IV.35) *Nechtěný pobyt na čerstvém vzduchu nabízí svým žákům Svobodná speciální škola J.A. Komenského v Praze na Kampě.* (PDT, bm102zua.fs #15, druhá část)

Substantiva *Praha* i *Kampa* mají rysy ‚místo‘ a ‚jméno‘. Protože *Kampa* je částí *Prahy*, a tedy po širším určení místa následuje určení bližší, jsou podle prvního bodu pozorování (16) Pg *v Praze* a Pg *na Kampě* určeny jako dvě nezávislé jednotky.

Pro identifikaci jejich společného řídicího substantiva lze (po redukci názvu) využít rozšíření pozorování (15). Substantivum *škola*, které předchází Pg *v Praze*, nese rys ‚instituce‘, a protože

substantiva s tímto rysem často představují místní určení, je Pg *v Praze* (a tedy i její sesterský uzel *na Kampě*) přednostně identifikována jako člen rozvíjející substantivum *škola*. (Adverbální analýza *nabízet v Praze* je upozaděna, analýza *žáci v Praze* je vyloučena na základě ‚principu separace‘.)

(IV.35') ... *Svobodná speciální škola J.A. Komenského na Kampě v Praze*.

Na druhou stranu, pokud Pg *na Kampě* předchází před Pg *v Praze*, a tedy užší určení místa předchází určení širšímu, je podle druhého bodu pozorování (16) Pg *v Praze* určena jako atribut první Pg, *na Kampě*. □

Předcházející příklady jasně ukazují, že pozorování (16) předpokládá velmi jemnou sémantickou charakteristiku jednotlivých objektů. Takto podrobný model světa ovšem nemůže být zahrnut do obecného automatického systému, může popisovat nejvýše omezenou doménu. Obecně proto musí automatická analýza zachovat dvě výše zmíněné syntaktické struktury:

- zkoumaná Pg rozvíjí předcházející substantivum;
- substantivum a následující Pg jsou syntakticky nezávislé, rozvíjejí však stejný větný člen.

Důsledek: (pozorování (16))

Na základě pozorování (16) tedy lze snížit preference analýzám, ve kterých není Pg ‚zavěšena‘ na předchozím substantivu, ani toto substantivum a Pg nemají společný řídicí větný člen (nejde o ‚sestry‘ ve stromové reprezentaci věty). □

Zobecnění pro jiná místní a časová určení

Pozorování (14)-(16) popisovala situaci, kdy je ve větě místní určení vyjádřené vlastním jménem. V některých takových případech lze s přihlédnutím k sémantickým rysům upřednostnit některé závislosti předložkových skupin. V následujících odstavcích navrhne, jak tato pozorování zobecnit pro místní určení, která nejsou vyjádřena vlastním jménem, případně i pro doplnění časová.

Příklady:

(IV.36a) *Po mnoha urgencích a zdržení (výměny na postech na českém ministerstvu školství) nakonec komise rozhodla, že projekt této školy neuspěl.* (PDT, bm102zua.fs #19)

(IV.36b) *To může zvládnout odbor třeba na ministerstvu financí.* (PDT, bcc25zua.fs #48)

V (IV.36a) má substantivum *ministerstvo* rys ‚institute‘ a jako takové může být místním určením, stejně jako substantivum *post*. Pg *na (českém) ministerstvu* je místní doplnění, které jednak rozvíjí bezprostředně předcházející substantivum *post*, jednak deverbativum *výměna* (stejně jako substantivum *post*).

V tomto případě mají obě analýzy stejný význam, proto stačí jediný rozbor – preferujeme závislost *post na (českém) ministerstvu*, tj. atributivní závislost Pg na bezprostředně předcházejícím substantivu (tento rozbor je podporován pozorováním (16), po bližším určení místa následuje určení širší – pokud ovšem lze takový vztah automaticky určit).

Obdobně i ve větě (IV.36b), kde substantiva *odbor* i *ministerstvo* mají rys ‚institute‘ a prvé z nich navíc nese rys ‚part‘ (opět jde o dvě určení místa, po ‚užším‘ označení následuje ‚širší‘) – je upřednostněna adnominální analýza *odbor na ministerstvu*.

(IV.36c) *Když ... ministr hospodářství ČR Karel Dyba sdělil na shromáždění na brífinku ekonomických ministrů, že návrh zásad zákona o neziskových organizacích byl „schválen“ s připomínkami, dal najevo, že ...* (PDT, bce17zua.fs #42, pozměněno)

Substantiva *shromáždění* i *brífink* nesou rys **„událost“**, obě jsou v tomto případě určením místa. Druhá Pg, *na brífinku*, je homonymní, může rozvíjet jednak předcházející substantivum *shromáždění*, jednak sloveso *sdělit*. Tato dvě substantiva není možné porovnat na základě sémantických rysů vzhledem k inkluzi, budou tedy zachovány obě analýzy se stejnou preferencí. □

Podívejme se, zda se dají vysledovat nějaké zákonitosti i pro volná určení časová.

Příklady:

(IV.37a) *Volební právo se rozhodně nedá ztotožnit s povinností vzdát se vždy v den voleb na několik let správy svých záležitostí.* (PDT, bm227zua.fs #28)

Substantiva *den* i *rok* mají rys **„čas“**. Pg *v den (voleb)* (v+Acc) je doplnění času – Kdy?, Pg *na (několik) let (na+Acc)* je rovněž doplnění času, ale odpovídající na otázku – Na jak dlouho? Protože jde o různá doplnění času, bude vytvořena jediná syntaktická struktura – Pg *v den (voleb)* a Pg *na (několik) let* budou určeny jako vzájemně nezávislé větné členy, které rozvíjejí sloveso *vzdát se*.

(IV.37b) *Když minulý týden v úterý ministr hospodářství ČR Karel Dyba ... sdělil, že ...* (PDT, bce17zua.fs #42)

Substantiva *týden* i *úterý* jsou určení **času**. Pokud by se dalo na základě sémantického modelu rozpoznat, že *týden* je širší určení než *den*, byl by to argument pro preferenci vzájemné syntaktické nezávislosti substantiva *týden* a za ním následující Pg *v úterý*. Potom by byly oba tyto členy určeny adekvátně jako časová doplnění slovesa *sdělit* (obdobně jako v pozorování (16)).

(IV. 37c) *V noci na neděli střežil nad Mločím převisem oheň.* (PDT, bm122zua.fs #47, pozměněno)

(IV. 37c') *?Na neděli v noci střežil nad Mločím převisem oheň.*

Substantiva *noc* a *neděle* nesou sémantický rys **„čas“**, není je možno porovnat vzhledem k inkluzi. Skutečnost, že spojení *noc na neděli* lze klasifikovat jako frazém (a jako takový je obsažen ve slovníku), vede k preferenci analýzy Pg *na neděli* jako neshodného atributu předcházející Pg.

Naproti tomu v (IV. 37c') – pokud ji budeme pokládat za správně utvořenou větu – je potřeba obě Pg s časovým určením analyzovat jako vzájemně nezávislé větné členy rozvíjející sloveso *střežit*. □

Shrňme zákonitosti ilustrované v předchozích příkladech do pozorování (17) a (18).

Pozorování: (pro dvě substantiva se stejnými sémantickými rysy)

(17) Vyskytnou-li se bezprostředně za sebou dvě doplnění stejného typu, která jsou vyjádřena **substantivy** se sémantickými rysy s primární funkcí **místa** (vzhledem k předložkám *na*, *v* a *do*), přičemž druhé z nich je částí jedné z předložkových skupin *na+Acc*, *na+Loc*, *v+Acc*, *v+Loc* nebo *do+Gen*, potom platí následující pozorování:

• pokud je možné obě substantiva porovnat na základě sémantických rysů vzhledem k inkluzi (např. má-li jedno z nich ‚part‘), potom nastává situace analogická pozorování (16):

◦ následuje-li po širším určení místa určení užší (např. druhé substantivum má rys ‚part‘), jde obvykle o dvě jednotky, které rozvíjejí stejný větný člen, tj. jsou vzájemně syntakticky nezávislé;

◦ následuje-li po bližším určení místa určení širší (např. první substantivum má rys ‚part‘), potom druhé jméno obvykle rozvíjí předcházející jméno, jde o jeho neshodný atribut;

• pokud tato dvě určení místa nelze porovnat vzhledem k inkluzi, jde o dvě nezávislé jednotky, obvykle se společným řídicím členem.

(18) Vyskytnou-li se bezprostředně za sebou dvě substantiva se sémantickým rysem ‚čas‘, přičemž druhé z nich je částí jedné z předložkových skupin *na+Acc*, *na+Loc*, *v+Acc*, *v+Loc* nebo *do+Gen*, jde obvykle o dvě jednotky, které rozvíjejí stejný větný člen, tj. jsou vzájemně syntakticky nezávislé. □

Poznamenejme, že zkoumání jiných kombinací substantiv s místním, resp. časovým určením nedává přesvědčivé doklady pro stanovení dalších závěrů použitelných pro zamýšlené ohodnocování závislostí předložkových skupin.

Kombinace místních určení vyjadřujících směr

Místní určení směru – kam je primární funkcí Pg *na+Acc* a *do+Gen*. Pokud takovému určení předchází, resp. je následuje místní určení vyjadřující směr – odkud (např. Pg *z+Gen*, nebo jinou Pg s primární funkcí místního určení směru – odkud), vyjadřují takové větné členy obvykle vztah odněkud–někam. Tuto heuristiku zachycuje následující pozorování.

Pozorování: (pro dvě substantiva vyjadřující směr odkud – kam)

(19) Vyskytnou-li se bezprostředně za sebou dvě doplnění stejného typu, která jsou vyjádřena **substantivy** se sémantickými rysy s primární funkcí **místa** (vzhledem k předložkám *na*, *v* a *do*), platí:

• pokud první substantivum vyjadřuje směr – odkud a druhé z nich je částí předložkové skupiny *na+Acc* nebo *do+Gen*, potom zkoumaná Pg obvykle rozvíjí stejný větný člen jako předcházející jméno, jsou syntakticky nezávislé. □

Kombinace místních a časových určení

Povšimněme si možnosti využít pro určení ‚spolehlivosti‘ jednotlivých analýz Pg tzv. systémové uspořádání slovesných doplnění, které udává jejich primární slovosledné uspořádání v rématu (viz Sgall et al (1986b)). Systémové uspořádání (system ordering, SO), které je vlastní rovině podkladové reprezentace, stanoví, že časové určení v bezpříznakové větě (v rématu) předchází určení místnímu. Tato skutečnost se odráží i v obvyklém povrchovém slovosledu, dá se proto využít k určení preference některých analýz Pg (tedy při práci na rovině povrchové reprezentace).

Pozorování: (pro substantiva s rysy ‚čas‘ a ‚konkr‘, ‚místo‘, ‚místo‘ + ‚jméno‘)

(20) Následuje-li po určitém slovese substantivum s rysem ‚čas‘ (v prostém nebo předložkovém pádu) a jedna z předložkových skupin *na+Acc*, *na+Loc*, *v+Loc* nebo *do+Gen*, ve které substantivum nese sémantické rysy ‚konkr‘, ‚místo‘, nebo ‚místo‘ + ‚jméno‘, potom jde obvykle o dva ‚sesterské‘ uzly – zkoumaná Pg obvykle rozvíjí stejný větný člen jako předcházející jméno, jsou syntakticky nezávislé. □

Poznamenejme zde explicitě, že právě proto, že jde o odraz podkladového SO na povrchové rovině, lze je využít jen pozitivně – objeví-li se v povrchové reprezentaci věty napřed časové určení a potom určení místní, jde pravděpodobně o dvě doplnění rozvíjející předcházející sloveso. Z opačného pořadí doplnění nebo u doplnění stojících před slovesem se nic vyvozovat nedá.

Příklad:

(IV.38) ... , pokračuje na Nový rok v Garmisch-Partenkirchenu, (PDT, bl123jsa.fs #9, zkráceno)

(IV.38') ... , pokračuje v Garmisch-Partenkirchenu na Nový rok,

Ve větě (IV.38) povrchový slovosled souhlasí se systémovým uspořádáním. Substantivum (Nový) rok a Pg v Garmisch-Partenkirchenu jsou v souladu s pozorováním (20) vzájemně nezávislé větné členy, oba rozvíjejí sloveso pokračovat.

Naopak, slovosledná varianta (IV.38') ilustruje skutečnost, že pozorování (20) nelze automaticky otočit – v její povrchové realizaci předchází místní určení určení časovému, přesto oba větné členy rozvíjejí sloveso pokračovat. (To je dáno zřejmě skutečností, že není splněn předpoklad bezpříznakovosti, tj. že nanejvýš jen druhá z obou Pg je v rématu.) □

Vyskytne-li se ve větě více volných doplnění, může výsledný počet syntaktických analýz neúměrně narůstat. Následující příklad ukazuje možnosti, které jsou dány pravidly využívajícími sémantické rysy substantiv. Právě jejich aplikace na řetězce volných doplnění přispívá k výraznému potlačení nadbytečných struktur.

Příklad:

(IV.39) Osmnáctiletý mladý muž s nožem v ruce **přinutil** v pátek v poledne na zastávce na Jírově ulici v Brně řidiče MHD, aby ho odvezl do Opavy a zpět. (PDT, bl209pp.fs #32, pozměněno)

Pokud by nebyla k dispozici pravidla využívající sémantické rysy, byla by věta (IV.39) (bez technických pravidel typu ‚last attachment principle‘, viz I.2.) silně homonymní (morfematická homonymie) – kumuluje se v ní pět volných doplnění, jejichž závislost nelze blíže určit na základě slovosledu ani valenčních požadavků, pouze některé analýzy jsou zablokovány ‚principem separace‘ (byla by stanovena závislost každého doplnění na slovese a na předcházejících substantivech (s přihlédnutím k ‚principu separace‘), což by znamenalo desítky analýz řetězce V – Pg_i – Pg_j – Pg_k – Pg_l – Pg_m).

Sémantická pozorování umožní stanovit vysokou preferenci jediné analýzy celého řetězce.

Podle pozorování (18) jsou Pg v pátek a Pg v poledne obvykle syntakticky nezávislé, obě rozvíjejí sloveso přinutit.

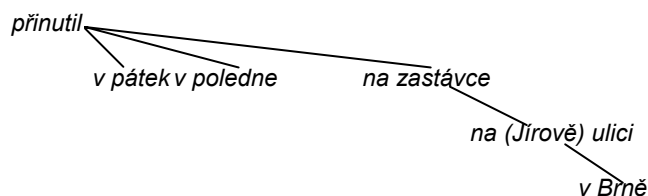
Bereme-li v úvahu pozorování (20), jsou časová a místní určení (v tomto pořadí, za slovesem) nezávislá – Pg na zastávce bude přednostně určena jako doplnění slovesa přinutit. (Bez tohoto pozorování přistupuje ještě atributivní závislost, Pg na zastávce bude určeno též jako rozvíjení substantiva poledne; analýza * pátek na zastávce je blokována ‚separačním principem‘.)

Podle pozorování (15) rozvíjí Pg s místním určením vyjádřeným vlastním jménem, na (Jírově) ulici,⁶⁴ přednostně substantivum s rysem ‚místo‘, zastávka.

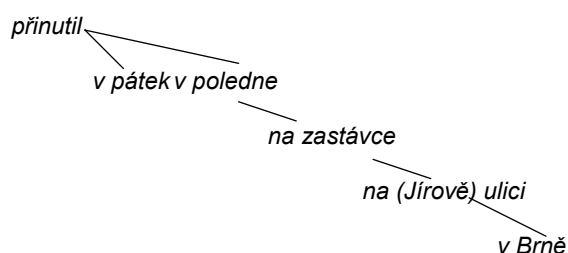
⁶⁴ Předpokládáme, že v předcházejících krocích RA bude Pg redukována na jádro, nicméně zůstane zachována informace, že neredukovaná Pg je tvořena předložkou na a vlastním jménem v lokálu.

Víme-li, že se *Jírova ulice* nachází *v Brně* (model světa), je na základě pozorování (16) Pg *v Brně* analyzována jako neshodný atribut Pg *na (Jírově) ulici*.

Sémantická pozorování tedy umožňují vysoko ohodnotit jednu z mnoha možných analýz řetězce sestávajícího ze slovesa a pěti volných doplňků času a místa. (Neuplatníme-li pozorování (20), budou výsledkem analýzy dvě preferované syntaktické struktury.)



Obr.26: Jediná preferovaná struktura věty (IV.39) určená na základě sémantických pozorování



Obr.27: Druhá syntaktická struktura (IV.39) (bez uplatnění pozorování (20))

Ostatní analýzy volných určení budou potlačeny. □

Závěr:

V této kapitole jsme popsali pozorování, která pracují se sémantickými rysy substantiv s primární funkcí času a místa (vzhledem k Pg *na+Acc*, *na+Loc*, *v+Acc*, *v+Loc* a *do+Gen*). Na základě takových pozorování lze v některých případech stanovit adekvátní závislost předložkových skupin a zejména potlačit ‚málo pravděpodobné‘ analýzy. □

Shrnutí

Byla ověřena možnost využití metody klasifikace sémantických rysů pomocí funkčních kritérií, a to i bez předpokladu znalosti syntaktické struktury věty – použili jsme ji pro určení rysů substantiv (vzhledem k Pg *na+Acc*, *na+Loc*, *v+Acc*, *v+Loc* a *do+Gen*).

Zkoumali jsme řetězce substantiv s rysy, které v testovaných Pg vyjadřovaly primárně čas a místo, formulovali jsme pozorování o pravděpodobných závislostech v takových řetězcích.

Na základě těchto pozorování, která pracují se sémantickými rysy jednotlivých lexikálních položek, lze říci, že – ačkoli v jednotlivých případech mohou být jejich výsledky povzbuzující – je jejich využití problematické. Ani v případech, kdy je využití sémantických rysů nasnadě (jako právě u určení místa a času), není analýza Pg na základě uvedených pozorování zcela spolehlivá.

Podle našeho mínění jsou sémantická kritéria vhodná především pro identifikaci jednotlivých typů volných doplňků (valenčních i fakultativních), sémantické vlastnosti jednotlivých předložek slouží k další specifikaci jednotlivých volných doplňků.

V případě morfematické homonymie předložkových skupin mohou být pravidla založená na sémantických rysech použita jako pomocná kritéria pro ohodnocení existujících syntaktických struktur. Mohou sloužit k ohodnocení jednotlivých analýz a k jejich uspořádání (podle ‚spolehlivosti‘).

IV.4. SLOVOSLED A FORMÁLNÍ KRITÉRIA

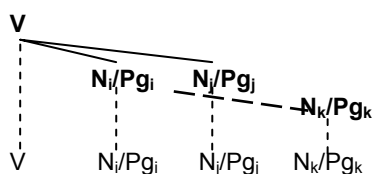
V oddíle IV.1. jsme se zabývali slovoslednými podmínkami, na jejichž základě lze určit pozice potenciálně homonymních předložkových skupin v povrchovém vyjádření věty (byly navrženy základní ‚podezřelé‘ slovosledné vzorce vzhledem k homonymii Pg). Zároveň byly vymezeny jazykové jevy, ve kterých se zdá opodstatněné tato slovosledná omezení uvolnit, kdy řídicí člen zkoumané Pg leží mimo množinu slov vymezenou slovoslednými omezeními. Oddíl IV.2. byl věnován valenčním rámcům jednotlivých slov, případně jejich kolokací. Vraťme se nyní znovu ke slovosledu a jeho konfiguracím. Formulujeme dva principy, které vylučují struktury s nepřipustnými závislostmi předložkové skupiny – ‚princip separace‘ a pravidlo založené na povrchové pozici příklonky.

IV.4.1. ‚Princip separace‘

‚Princip separace‘ lze uplatnit v případě, kdy už jsou (na základě jiných pravidel) navrženy některé závislosti. Popisuje možné závislosti v situaci, kdy je ve větě sloveso následované třemi podstatnými jmény (v prostém (N) nebo předložkovém pádu (Pg)), z nichž druhé je určeno jako slovesné doplnění – vylučuje jeden typ neprojektivní struktury.

Pozorování: (‚princip separace‘)

(21) Mějme povrchovou reprezentaci věty, ve které se vyskytuje řetězec složený ze slovesa bezprostředně následovaného třemi (redukovánými) jmennými nebo předložkovými skupinami⁶⁵ ($V - N_i/Pg_i - N_j/Pg_j - N_k/Pg_k$). Pokud je druhá skupina, N_j/Pg_j (resp. její řídicí substantivum) již určena jako doplnění tohoto slovesa, potom třetí skupina, N_k/Pg_k nemůže záviset na řídicím substantivu první skupiny, N_i/Pg_i . To znamená, že substantivum (v prostém i předložkovém pádu) analyzované jako doplnění slovesa nemůže oddělovat předcházející podstatné jméno a jeho následující (neshodný) atribut. □

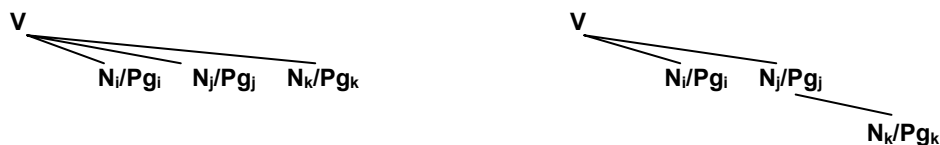


Obr.28: Nepřipustná analýza řetězce $V - N_i/Pg_i - N_j/Pg_j - N_k/Pg_k$

Důsledek:

V situaci, kterou popisuje pozorování (21), může být substantivum, příp. předložková skupina N_k/Pg_k analyzována jednak jako slovesné doplnění, jednak jako atribut druhého jmenného členu řetězce, N_j/Pg_j (viz následující obrázek). □

⁶⁵ Termín redukováná jmenná skupina používáme v obdobném významu jako redukováná Pg – typicky jde o řídicí substantivum jmenné skupiny. (RA umožňuje pomocí přepisování zpracovat i případy, kdy pád řídicího substantiva v neredukované jmenné skupině neodpovídá distribuci celé skupiny, např. ve výrazech s číslovkami.)



Obr.29: Dvě možné analýzy řetězce V – Ni/Pgi – Nj/Pgj – Nk/Pgk

Díky zvolenému přístupu, analýze založené na principu redukce, lze substantiva a redukované Pg v pozorování (21) nahradit jmennými skupinami a (neredukovanými) Pg.

Následující věty ilustrují možnosti ‚principu separace‘.

Příklad:

(IV.40a) *Vysoký činitel ... se snaží přimět Palestince k ústupkům na poslední chvíli.* (PDT, blc02zu.fs #36)

(IV.40a') *Vysoký činitel ... se snaží přimět Palestince k odchodu z Jeruzaléma.*

(IV.40a'') *Vysoký činitel ... se snaží přimět k odchodu Palestince z Jeruzaléma.*

(ostatní slovosledné variace ve zkoumaném řetězci nesplňují podmínku pozorování (21))

Pro analýzu věty (IV.40a) a jejích variací je podstatný valenční rámec slovesa *přimět*:

přimět ... Act | Pat (k+Dat/inf/klauze-aby) Addr (Acc)

Na jeho základě je ve všech větách (IV.40a) substantivum *Palestinci* určeno jako adresát (přesněji vyjádření hloubkového Addr na povrchové rovině), Pg *k ústupkům*, resp. *k odchodu* jako patient (jeho povrchové vyjádření). Je splněna podmínka pozorování (21), vždy druhý jmenný člen zkoumaného řetězce závisí na slovese.

Původní věta je víceznačná, Pg *na (poslední) chvíli* může rozvíjet jednak sloveso *přimět*, jednak Pg *k ústupkům*, což je zcela v souladu s důsledkem pozorování (21).

Ve větě (IV.40a') je vyloučena závislost *Palestinci z Jeruzaléma*, podle důsledku pozorování (21) může být Pg *z Jeruzaléma* analyzována jednak jako rozvíjení slovesa *přimět*, jednak jako atribut substantiva *odchod*. Na základě valenčního požadavku deverbativa *odchod* (směr, Direction – odkud) je preferována adnominální analýza.

V (IV.40a'') pozorování (21) blokuje analýzu *odchod z Jeruzaléma* (přestože Pg splňuje valenční požadavky substantiva stejně jako v předchozí větě), což také odpovídá významu věty. Pg *z Jeruzaléma* může rozvíjet substantivum *Palestince* a sloveso *přimět*. Stanovení preference adekvátní analýzy, *Palestinci z Jeruzaléma*, je záležitostí obsahovou.⁶⁶ □

Skutečnost, že už určené závislosti jsou dány valenčními požadavky slovesa, není pro platnost ‚principu separace‘ důležitá – v následujících větách jsou jako ‚separační členy‘ využita volná doplnění slovesa.

Příklady:

(IV.40b) *Toto vše dělali po celý rok pro Palestince bez omezení.*

Pg *pro Palestince* je volným doplněním prospěchu (Ben, Benefactive), přesto znemožňuje vytvoření závislostního páru *po (celý) rok bez omezení*

(IV.41) *Přijel vlakem z Budapešti do Prahy.*

⁶⁶ Tuto závislost můžou podporovat i vhodně stanovená pravidla využívající sémantické rysy substantiv, jimž byl věnován oddíl IV.3. (Podrobněji jsme se zabývali Pg *na+Acc*, *na+Loc*, *v+Acc*, *v+Loc* a *do+Gen.*)

(IV.41') *Přijel vlakem z Budapešti do Prahy, kterou chtěl už dlouho navštívit.*

(IV.41'') *Do Prahy přijel vlakem z Budapešti do Berlína.*

Obě Pg ve větě (IV.41), z *Budapešti* a *do Prahy*, jsou homonymní – mohou rozvíjet jednak sloveso *přijet*, jednak substantivum *vlak*, s různými pravdivostními podmínkami; rozdíly lze ilustrovat příklady (IV.41') a (IV.41''). Přitom v (IV.41) analýza druhé Pg závisí na analýze první Pg (na základě významu).

Pokud je Pg z *Budapešti* chápána jako volné doplnění slovesa *přijet* (směr, Direction – odkud), nemůže být Pg *do Prahy* určena jako doplnění substantiva *vlak*. Tuto skutečnost odráží ‚princip separace‘, který při automatickém zpracování blokuje tuto nepřipustnou analýzu. Valenční požadavky slovesa *přijet*, které vyžaduje kromě aktoru i obligatorní (vypustitelné) doplnění místa (směr, Direction – Kam), vedou k adverbální analýze Pg *do Prahy*. (Tato analýza je podporována i sémantikou obou předložek. Pg z+Gen typicky vyjadřuje směr – odkud, Pg do+Gen směr – kam, je tedy vhodné upřednostnit analýzu obou Pg jako sester rozvíjejících stejný větný člen.)

Pokud je Pg z *Budapešti* určena jako neshodný atribut substantiva *vlak*, zůstává Pg *do Prahy* homonymní. Může být analyzována buď jako obligatorní doplnění slovesa *přijet*, nebo jako doplnění podstatného jména *vlak* (větné členy vyjadřující vztah odněkud-někam ‚zavěšeny‘ jako sesterské uzly); ve druhém případě je sémanticky obligatorní doplnění slovesa vypuštěno na povrchové rovině. (Adnominální rozbor s Pg z *Budapešti do Prahy* může být vyloučen na základě sémantiky Pg, viz oddíl IV.3.)

(IV.36c) *Když ... ministr hospodářství ČR Karel Dyba sdělil na shromáždění na brífinku ekonomických ministrů, že návrh zásad zákona o neziskových organizacích byl „schválen“ s připomínkami, dal najevo, že ...* (PDT, bce17zua.fs #42, pozměněno)

Při tomto povrchovém slovosledu je vyloučena závislost *shromáždění ministrů*, předložková skupina *na brífinku* odděluje Pg *na shromáždění* a substantivum *ministrů*. Pokud je Pg *na brífinku* určena jako doplnění slovesa,⁶⁷ lze ji považovat za ‚separační člen‘ a neadekvátní závislostní pár *shromáždění ministrů* vyloučit na základě pozorování (21). (Zdá se ovšem případnější nepřipustit takovou analýzu pomocí méně komplexních podmínek pro pozici neshodného atributu v genitivu, které pracují pouze s povrchovým slovosledem a lze je uplatnit bez ohledu na rozbor Pg *na brífinku*.) □

Zdá se, že obdobné tvrzení jako tvrzení v pozorování (21) platí i pro řetězec jmenných nebo předložkových skupin, po kterých následuje sloveso.

Zobecnění pozorování: (zobecněný princip separace‘)

(21') Mějme povrchovou reprezentaci věty, ve které se vyskytuje řetězec složený ze třech (redukovaných) jmenných nebo předložkových skupin, před nimiž nebo po nichž následuje sloveso ($V - N_i/Pg_i - N_j/Pg_j - N_k/Pg_k$, resp. $N_i/Pg_i - N_j/Pg_j - N_k/Pg_k - V$). Pokud je druhá skupina, N_j/Pg_j (resp. její řídicí substantivum) již určena jako doplnění tohoto slovesa, potom třetí skupina, N_k/Pg_k nemůže záviset na řídicím substantivu první skupiny, N_i/Pg_i . To znamená, že substantivum (v prostém i předložkovém pádu) analyzované jako doplnění slovesa nemůže oddělovat předcházející podstatné jméno a jeho následující (neshodný) atribut. □

Pozměníme-li slovosled věty z příkladu (IV.40a), dostaneme situaci splňující zobecněné pozorování (21') (tato slovosledná varianta není bezpříznaková, nicméně na ní lze ilustrovat toto

⁶⁷ Pg *na brífinku* je homonymní, může rozvíjet jednak předcházející substantivum *shromáždění*, jednak sloveso *sdělit*. K možnostem upřednostnit jednu z těchto analýz na základě porovnání sémantických rysů substantiv viz kapitolu (IV.3.2.).

zobecnění). Opět nezávisí na tom, zda jde o valenční nebo volná doplnění (viz (IV.36c')). Navíc se zřejmě zkoumaný řetězec může skládat z více jmenných, resp. předložkových skupin – podstatná je skutečnost, že ‚separační člen‘ nerozvíjí žádnou z předcházejících jmenných, resp. předložkových skupin ((IV.36c'), (IV.40b')), vyžaduje ovšem další ověření).

Příklad:

(IV.40a) *Vysoký činitel Palestince k ústupkům na poslední chvíli přiměl.* (PDT, blc02zu.fs #36, pozměněno)

(IV.40a') *Vysoký činitel Palestince k odchodu z Jeruzaléma přiměl.*

(IV.40a'') *Vysoký činitel k odchodu Palestince z Jeruzaléma přiměl.*

Opět je ve všech větách (IV.40a) základě valenčních požadavků slovesa určeno substantivum *Palestinci* jako adresát (vyjádření hloubkového Addr na povrchové rovině), Pg *k ústupkům*, resp. *k odchodu* jako patient (jeho povrchové vyjádření). Je splněna podmínka pozorování (21'), vždy druhý jmenný člen zkoumaného řetězce (tj. Pg *k+Dat* v (IV.40a) a (IV.40a'), resp. substantivum *Palestinci* v (IV.40a'')) závisí na slovese. Třetí jmenný člen řetězce (tj. Pg *na (poslední) chvíli*, resp. *z Jeruzaléma*) nelze analyzovat (v souladu s pozorováním (21')) jako atribut první jmenné, resp. předložkové skupiny, ačkoliv by byl takový rozbor podporován sémantickými / pragmatickými zřetely.

(IV.36c') *Když ... ministr hospodářství ČR Karel Dyba na shromáždění na brífinku ekonomických ministrů sdělil, že návrh zásad zákona o neziskových organizacích byl „schválen“ s připomínkami, dal najevo, že*

(IV.36c'') *Návrh zásad zákona o neziskových organizacích byl „schválen“ s připomínkami, sdělil ministr hospodářství ČR Karel Dyba na shromáždění na brífinku ekonomických ministrů.* □

(IV.40b') *Dělali to po celý rok pro Palestince bez omezení.*

Závěr:

‚Princip separace‘ představuje velmi silné negativní pravidlo, které lze aplikovat na valenční i volná doplnění. Nepodařilo se nám (zatím) nalézt nebo zkonstruovat větu, ve které by byl tento princip porušen.

Tento typ kritérií, který vylučuje jisté typy neprojektivních konstrukcí, slibuje další zajímavé výsledky; v příkladech byl naznačen možný směr dalšího výzkumu. □

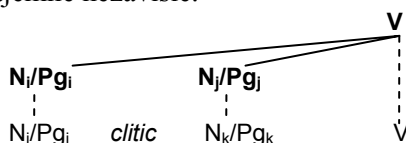
IV.4.2. Pravidlo o postavení příklonek ve větě

Dalším pravidlem, kterého lze využít pro určení možných závislostí předložkových skupin, je pravidlo o postavení příklonek, klitik v české větě. Příklonky obvykle zaujmají ‚druhou pozici‘ (Wackernagelovu pozici) ve větě, tedy pozici za prvním přízvučným úsekem věty (viz Uhlířová (1987)). Tímto prvním přízvučným úsekem může být skupina slov tvořící rozvitý větný člen. Proto může pozice příklonky přispět k zřetelnějšímu syntaktickému členění věty – příklonka může zabránit vzniku ‚falešné větné dvojice‘.⁶⁸

⁶⁸ Pokud by byly všechny příklonky na RPS reprezentovány vlastním uzlem, jako je tomu na analytické rovině, bylo by pozorování (22) důsledkem ‚zobecněného principu separace‘ – klitika by v takovém případě sloužila jako separační člen.

Pozorování:

(22) Příklonka může oddělovat jednu jmennou, příp. předložkovou skupinu od jiné jmenné, příp. předložkové skupiny. Pokud jsou dvě (redukované) jmenné, resp. předložkové skupiny ve zkoumané větě odděleny příklonkou (N_i/Pg_i – příklonka – N_j/Pg_j), potom druhou skupinu N_j/Pg_j nelze analyzovat jako atribut první skupiny N_i/Pg_i – členy těchto dvou skupin jsou syntakticky vzájemně nezávislé. □



Obr.30: Možný rozbor řetězce N_i/Pg_i – příklonka – N_k/Pg_k

Následující příklady ilustrují přínos pozice reflexivního zájmena *se* pro určení větného členu, který je rozvíjen předložkovou skupinou.

Příklady:

(IV.42) ... vláda a následně Parlament ČR se po dlouhých odkladech pokusí stanovit pro podnikání seriózní a legislativně jasná pravidla hry. (PDT, bce17zua.fs #42, zkráceno)

Ve větě (IV.42) jsou dva rozvíjené větné členy, koordinovaná jmenná skupina [*vláda a Parlament ČR*] a Pg *po (dlouhých) odkladech*, které jsou odděleny reflexivním zájmenem *se*. Podle pozorování (22) jsou tyto větné členy syntakticky nezávislé. Skutečně, oba modifikují sloveso *pokusit* [*vláda a Parlament ČR*] vyjadřuje hloubkový aktor, Pg je volné doplnění času (Twhen – kdy)).

(IV.43) Pro vytváření generátorů impulsů se v číslicových systémech používá oscilátorů. (Uhlířová (1987))

(IV.43') Pro vytváření generátorů impulsů v číslicových systémech se používá oscilátorů.

V (IV.43) jsou dvě Pg (*pro vytváření generátorů impulsů* a *v číslicových systémech*), které jsou odděleny příklonkou *se*. Obě Pg závisí na slovese *používat*, v souladu s pozorováním (22).

Naproti tomu v (IV.43') je první Pg (*pro vytváření generátorů impulsů*) rozvíjena bezprostředně následující Pg *v číslicových systémech*. Tyto dvě předložkové skupiny tvoří jeden přízvučný úsek věty. □

Závěr:

Vymezení prvního přízvučného úseku věty není jednoduché. Ačkoli se obvykle kryje s jednou syntaktickou skupinou (odpovídající jednomu keři v závislostním stromu), lze nalézt věty, ve kterých první přízvučný úsek věty pokrývá více syntaktických skupin, a naopak nepřijatelné věty s jedinou syntaktickou skupinou stojící před příklonkou (viz Avgustinova a Oliva (1995)).

Tato nejistota ohledně určení prvního přízvučného úseku věty však nemá vliv na platnost pozorování (22), které platí ‚jedním směrem‘ (jde o implikaci, nikoli o ekvivalenci).⁶⁹ □

⁶⁹ Věta (IV.43') sice dokládá, že často platí i obrácené tvrzení, nicméně řada příkladů, které uvádí Avgustinova a Oliva (1995), je vyvrací:

Od Aše po Vladivostok se rozkládala světová socialistická soustava.
?V listopadu na Václavském náměstí se konala demonstrace.

Shrnutí

„Princip separace“ je důležitým kritériem, které se uplatňuje vzhledem k původní, neredukované větě. Umožňuje zablokovat některé nesprávné analýzy Pg, a to i takové analýzy, která by neadekvátně zaplňovaly valenční pozici jistého větného členu (a jako takové měly vysokou preferenci, viz kapitola IV.2.5.).

Také pravidlo o postavení příklonek lze využít pro stanovení hranic jednotlivých syntaktických členů. Neshodný atribut nemůže být oddělen příklonkou od substantiva, které rozvíjí.

ČÁST V.

UPLATNĚNÍ A EVALUACE NAVRŽENÝCH KRITÉRIÍ

V této části předkládáme návrh uspořádání jednotlivých typů kritérií, která byla zkoumána v části IV., a způsob jejich využití pro automatické určení větných členů, které mohou být rozvíjeny předložkovou skupinou, v mechanismu analýzy založené na principu redukce (RA).

Zopakujme zde, že stanovená kritéria pro detekci homonymie předložkových skupin pracují s plně morfologicky zpracovaným textem. Předpokládáme, že každé slovo zkoumané věty nese komplexní morfologickou, syntaktickou a sémantickou informaci extrahovanou ze slovníku – podtrhněme v tomto kontextu zejména skutečnost, že každé sloveso, substantivum a adjektivum musí mít přiřazeny (kromě morfologické charakteristiky) všechny možné valenční, případně i kvazivalenční rámce. Lexikální položky pro substantiva a předložky případně obsahují údaj o sémantických rysech. Kromě těchto údajů, vztahujících se k jednotlivým členům věty, předpokládáme, že jsou na základě slovníku identifikovány slovesně jmenné výrazy a složena ustálená spojení.

Před začátkem vlastní analýzy předložkových skupin jsou zpracovány některé jevy lokální povahy – především jsou určeny analytické tvary sloves. Pro jednoduchost také předpokládáme, že jsou složeny jednoduché jmenné skupiny obsahující číslovky a redukována zdůrazňovací slova.

Metodologický rámec pro studium závislostí předložkových skupin tvoří redukční analýza. RA je realizována jako posloupnost kroků, v nichž jsou uplatňována jednotlivá navržená kritéria. Ta budou aplikována v každé větvi, v každém kroku analýzy, ve kterém může být redukována zpracovávaná Pg. Některá tato kritéria se vztahují k původnímu znění (rozvité) věty, jiná pracují s větou už částečně redukovanou. Soustředíme se na zachycení morfematické homonymie, předpokládáme jednoznačnou morfologickou analýzu zkoumané Pg (v případě fonologické homonymie této Pg, tj. má-li tato Pg více možných morfologických analýz, se analýza rozpadne do dvou (příp. více) větví, v každé z nich je zpracována jedna kombinace sémat).

V.1. USPOŘÁDÁNÍ KRITÉRIÍ V JEDNOTLIVÝCH KROCÍCH ANALÝZY ZALOŽENÉ NA PRINCIPU REDUKCE A VYHODNOCENÍ ZÍSKANÝCH ZÁVISLOSTÍ

V každém kroku analýzy, ve kterém je redukována zkoumaná Pg (s jednoznačnou morfologickou analýzou), jsou aplikována všechna kritéria stanovená v části IV. Na jejich základě jsou určeny potenciální řídicí uzly Pg (vzhledem k aktuálnímu stavu věty). Získané analýzy jsou ohodnoceny podle typu pravidla, podle kterého byla ta která možná závislost určena. Pokud jsou v daném kroku RA zjištěny dvě (nebo více) možné závislosti, analýza se (v intencích RA) rozpadá na příslušný počet větví. Není-li určena žádná závislost, daná větev analýzy končí v daném kroku neúspěchem, RA pokračuje jinou větví (pokud není jiná větev k dispozici, končí neúspěchem analýza celé věty).

Řazení pravidel v jednom kroku analýzy

1. Jsou zkontrolovány valenční a kvazivalenční požadavky všech autosémantických sloves, substantiv a adjektiv. Zkoumaná Pg potenciálně uspokojuje všechny valenční i kvazivalenční požadavky, které (vzhledem k možnému povrchovému vyjádření) lze takovou Pg uspokojit – stanoví se všechny takové závislosti.
2. Je (nedeterministicky) vybrán slovosledný vzorec, WOP aktuálního stavu věty (tj. věty částečně redukované).
3. Aplikují se pravidla vyplývající z WOP (částečně) redukované věty, a to následujícím způsobem (podle definic v kapitole IV.1.1.):
 - závislosti, které už byly stanoveny na základě valenčních a kvazivalenčních požadavků, se znovu netvoří;
 - pro základní ‚jednoznačné‘ WOP se stanoví příslušná závislost;
 - pro základní ‚podezřelé‘ WOP se stanoví všechny potenciální závislosti Pg;
 - pokud je naplněn WOP s preferovaným čtením pro adjektiva, je Pg určena jako rozvítil adjektiva.
4. Zkontroluje se, zda stanovené závislosti respektují formální kritéria vzhledem k původní (neredukované) větě – při porušení ‚principu separace‘ nebo pravidla o postavení příklonky ve větě je analýza vyloučena.
5. Zkontroluje se, zda vytvořené závislosti splňují slovosledná omezení:
 - pokud je v klauzi sloveso, jehož valenční nebo kvazivalenční požadavek lze zaplnit zkoumanou Pg (včetně požadavků zachycujících přenesené a posunuté užití), berou se v úvahu slovosledná omezení pro (valenční, kvazivalenční i volná) rozvíjení substantiv stanovená v kapitole IV.1.1. (pozorování (1) a (2));
 - pokud se v klauzi nevyskytuje sloveso, jehož valenční nebo kvazivalenční požadavek lze zaplnit zkoumanou Pg, neberou se pro valenční a kvazivalenční doplnění substantiv v úvahu slovosledná omezení (pozorování (9) a (10) z kapitoly IV.1.2.);
 - pro naplnění valenčního rámce adjektiva platí pozorování (3) z kapitoly IV.1.1.

- adjektivum (jako shodný atribut substantiva) může být rozvíjeno Pg, která je od tohoto adjektiva oddělena substantivem (v pořadí $A_i N_i P_g$) (pozorování (12)).

Pokud některá ze závislostí nespĺňuje uvedená omezení, je vyloučena.

Pokud se reflektují slovesně jmenné výrazy, je pro slovesně jmenné výrazy navíc vyloučena analýza napĺňující případný valenční nebo kvazivalenční požadavek slovesa; je zachována pouze závislost Pg na substantivu. (Skutečnost, zda nejde o slovesně jmenný výraz, je potřeba kontrolovat vzhledem k původní, nezjednodušené větě.)

6. Zjistí se, zda v daném kroku RA vznikla alespoň jedna platná závislost Pg:

- pokud byla taková závislost určena, je Pg vypuštěna, analýza je úspěšná, RA pokračuje dalším krokem;

- pokud žádná taková závislost určena nebyla, Pg není vypuštěna, krok analýzy je neúspěšný, RA pokračuje jinou větví.

Pokud je alespoň jedna větev RA úspěšná (Pg je redukována), potom se podařilo úspěšně analyzovat Pg ve zkoumané větě. □

Poznámky:

K bodu 3. poznamenejme, že tímto způsobem zabráníme systematickému vytváření dvojí závislosti Pg jako valenčního a volného doplnění jediného větného členu. V ojedinělých případech se může stát, že takto eliminujeme platnou analýzu (věta (I.1), kde Pg *na (domácím) trhu* může být určena jako valenční i volné doplnění substantiva *zájem*, s různým významem; viz také věty (II.11a) a (II.11b), *prohrávat / válčit s nimi* a (II.11c), *sejít se s někým/něčím*).

K bodu 5. uveďme zopakujeme, že pro účely automatického zpracování Pg předpokládáme, že valenční rámce sloves popisují i jejich posunutá a přenesená užití. □

Vyhodnocení získaných závislostí

Platné závislosti vzniklé na základě valenčních nebo kvazivalenčních požadavků se ohodnotí podle jejich ‚spolehlivosti‘ stanovené v kapitole IV.2.5., ohodnotí se i závislosti určené na základě slovosledu.

1. Valenční, kvazivalenční i volná doplnění adjektiva podporovaná WOP s preferovaným čtením pro adjektiva mají nejvyšší preferenci

2. Následují závislosti uspokojující valenční a kvazivalenční požadavky, a to v pořadí daném v IV.2.5., to jest:

- ‚klasické‘ valenční rámce sloves (včetně rámců zachycujících jejich přenesená a posunutá užití, tj. kvazivalenční doplnění prvního typu) mají nejvyšší prioritu při určování možné závislosti konkrétního doplnění (v našem případě Pg);

- následují valenční požadavky adjektiv (pokud není naplněn WOP s preferovaným čtením pro adjektiva a přitom jsou splněna slovosledná omezení);

- valenční rámce substantiv má nejnižší prioritu ze standardních valenčních doplnění; jejich uplatnění pro netypickou a neprojektivní pozici Pg je podmíněno nepřítomností valenčních požadavků slovesa;

- následují doplnění daná kvazivalenčními požadavky jednotlivých slovních druhů:

- kvazivalenční doplnění substantiv,

- kvazivalenční doplnění sloves (druhý typ) a
- kvazivalenční doplnění adjektiv.

3. Závislosti vyplývající z ‚jednoznačného‘ a ‚podezřelého‘ WOP mají nižší preferenci než závislosti dané valencí, resp. kvazivalencí, přitom platí:

- pokud jsou k dispozici sémantické rysy a pravidla pro jejich využití, stanoví se preferenční čtení Pg jako volného doplnění (vzhledem k původní (neredukované) větě);
- pokud sémantické rysy a pravidla pro jejich využití k dispozici nejsou, mají všechny takto určené závislosti stejnou ‚váhu‘.

4. Valenční, kvazivalenční i volná doplnění adjektiva oddělená substantivem mají preferenci nižší než volná doplnění splňující slovosledná omezení. □

Závěr:

Stanovená posloupnost kroků zachycuje lingvistická pozorování a závěry z předchozích kapitol. Realizuje tak model automatické syntaktické analýzy, který respektuje principy obecné metody redukční analýzy.

Navržený postup umožňuje určit možné závislosti předložkových skupin na základě různých typů lingvistických kritérií a uspořádat je podle jejich ‚spolehlivost‘. □

V.2. UPLATNĚNÍ KRITÉRIÍ VE VYBRANÝCH PŘÍKLADECH

V následujících odstavcích se vrátíme k některým větám z části IV. a na konkrétních příkladech přiblížíme způsob aplikace kritérií, která jsme zkoumali v předchozí části, a jejich uspořádání. V jednotlivých větách se soustředíme vždy na některý z lingvistických jevů, kterými jsme se zabývali v předchozích kapitolách; všímáme si jeho zpracování v rámci jednotlivých kroků analýzy pracující na principu redukce. (V následující, VI. části popíšeme výsledky analýzy jednotlivých příkladů pomocí konkrétního parseru, jehož gramatiku jsme přizpůsobili tak, aby v zásadě odpovídala navrženým kritériím.)

Poznámka:

Úvahy týkající se valence jednotlivých lexikálních jednotek prezentované v oddíle IV.2. byly obecné, neopíraly se o konkrétní model pro zachycení valenčních požadavků určitých typů lexikálních jednotek (sloves, substantiv a adjektiv).

Zde předpokládáme, že pro danou lexikální jednotku budou jednotlivé položky, které tvoří valenční, příp. kvazivalenční rámec, zachyceny pomocí „slotů“ obsahujících informaci o typu doplnění a jeho možné povrchové realizaci (pro vnitřní doplnění). Je-li některý větný člen určen jako aktant zkoumané lexikální jednotky, je její příslušný valenční „slot“ označen jako zaplněný (v dané větvi analýzy nemůže být žádný další větný člen analyzován jako tento aktant, valenční požadavek je zablokován – to zabraňuje nepřipustnému dublování vnitřních doplnění). □

První příklad ilustruje práci s valenčními a kvazivalenčními rámci. Podrobný rozbor simulující automatickou analýzu umožňuje zajímavá pozorování (viz závěr příkladu).

Příklad:

(V.1) *Zákon dbá na ochranu nájemníků.* (PDT, b1101js.fs #18, zkráceno)

Na modifikaci věty (V.1) (označené jako (IV.1)) jsme ilustrovali mechanismus redukční analýzy; její slovosledné varianty jsme zkoumali vzhledem ke křížení valenčních požadavků (příklad (IV.31a)).

Sloveso *dbát* má dva valenční rámce:

*dbát*₁ ... Act | Pat (na+Acc/o+Acc)

*dbát*₂ ... Act | Pat (Gen)

Předpokládáme též následující rámce substantiva:

zákon ... Pat (o+Loc)

zákon ... ^Aim (na+Acc)

Analýza založená na principu redukce se rozpadá do několika větví, redukce Pg *na ochranu (nájemníků)* začíná v několika stavech.

1. Pro rámec *dbát*₁ postupuje analýza Pg v následujících krocích:

(a) *Zákon dbá na ochranu nájemníků.*

(Původní věta, zatím nebylo nic vypuštěno.)

1. Je zjištěn jeden valenční požadavek – *dbát na ochranu* – a jeden kvazivalenční požadavek – *zákon na ochranu*. Jsou vytvořeny obě závislosti.

2. Je vybrán WOP typu **VPg**.
3. Na základě WOP nevzniká žádná nová struktura.
4. Nejsou porušena formální kritéria.
5. Nejsou splněny podmínky pro uvolnění slovosledných omezení, je vyloučena analýza Pg uspokojující kvazivalenční požadavek substantiva.
6. Byla zachována jedna nová závislost, Pg *na ochranu* je redukována („slot“ pro patient slovesa *dbát₁* označen jako zaplněný), redukce pokračuje dalším krokem.
(Tato větev analýzy skončí později neúspěchem, neboť nebude možno spojit jmennou skupinu (*na ochranu nájemníků*.)

(b) *Zákon dbá na ochranu.*

(Částečně redukována věta, vypuštěno substantivum *nájemníci* (zaplňuje valenční požadavek substantiva *ochrana*.)

- 1.-6. Stejně jako v předešlé větvi (a).

(Tato větev vede k úspěšné analýze zkoumané věty (jmenná skupina (*na ochranu nájemníků*) složena před redukcí Pg.)

(c) *Dbá na ochranu nájemníků.*

(Částečně redukována věta, vypuštěno substantivum *zákon* (aktor slovesa *dbát₁*.)

1. Je zjištěn jeden valenční požadavek – *dbát na ochranu*.
- 2.-4. Stejně jako v předešlé větvi (b).
5. Vytvořená závislost splňuje slovosledná omezení.
6. Stejně jako v předešlé větvi.

(Tato větev analýzy skončí později neúspěchem, neboť nebude možno spojit jmennou skupinu (*na ochranu nájemníků*.)

(d) *Dbá na ochranu.*

(Částečně redukována věta, vypuštěno substantiva *zákon* (aktor slovesa *dbát₁*) a *nájemníci* (atribut substantiva *ochrana*.)

- 1.-6. Stejně jako v předešlé větvi (c).

(Tento krok vede k úspěšné analýze zkoumané věty – zkoumaná věta je redukována na predikát *Dbá*.)

Shrňme: Větve, jejichž součástí je popisovaný krok (b) a (d), vedou k úspěšné analýze, Pg *na ochranu (nájemníků)* je analyzována jako valenční doplnění slovesa *dbát₁* (jeho patient). Závislost *dbát na ochranu* má vysokou preferenci (jde o valenční doplnění).

2. Rámec *dbát₂* je vybrán ve větvi RA, která identifikuje jako patient substantivum *nájemníci*:

(a) *Zákon dbá na ochranu nájemníků.*

(Původní věta, zatím nebylo nic vypuštěno.)

1. Je zjištěn jeden kvazivalenční požadavek – *zákon na ochranu*. Je vytvořena jedna závislost.
2. Je vybrán WOP typu **VPg**.
3. Na základě WOP vzniká závislost *dbát na ochranu*.
4. Nejsou porušena formální kritéria.
5. Jsou splněny podmínky pro uvolnění slovosledných omezení, analýza Pg uspokojující kvazivalenční požadavek substantiva je zachována.
6. Byly vytvořeny dvě závislosti, Pg *na ochranu* je redukována, analýza se rozpadá do dvou větvi (v jedné má substantivum *zákon* příslušný „slot“ označen jako zaplněný).
(Tyto větve vedou k analýze zkoumané věty, kterou je vzhledem ke stanoveným kritériím nutno považovat za úspěšnou, víc viz shrnutí v závěru příkladu.)

(b) *Zákon dbá na ochranu.*

(Částečně redukována věta, vypuštěno substantivum *nájemníci* (zaplňuje valenční požadavek slovesa *dnát₂*, patient).)

- 1.-6. Stejně jako v předešlé větvi (a).

(Tyto větve vedou k analýze zkoumané věty, kterou je vzhledem ke stanoveným kritériím nutno považovat za úspěšnou, víc viz shrnutí v závěru příkladu.)

(c) *Dbá na ochranu nájemníků.*

(Částečně redukováná věta, vypuštěno substantivum *zákon* (aktor slovesa *dbát₂*)).

1. Nezjištěn valenční ani kvazivalenční požadavek na Pg-doplnění .
2. Je vybrán WOP typu **VPg**.
3. Na základě WOP vzniká závislost *dbát na ochranu*.
4. Nejsou porušena formální kritéria.
5. Vytvořená závislost splňuje slovosledná omezení.
6. Byla vytvořena jedna závislost, Pg *na ochranu* je redukována, redukce pokračuje dalším krokem.
(Tato větev vede k úspěšné analýze zkoumané věty.)

(d) *Dbá na ochranu.*

(Částečně redukováná věta, vypuštěna substantiva *zákon* a *nájemníci* (aktor a pacient slovesa *dbát₂*)).

- 1.-6. Stejně jako v předešlé větvi (c).
(Tento krok vede k úspěšné analýze zkoumané věty – zkoumaná věta je redukována na predikát *Dbá*.)

Shrnutí: V popsanych větvičích analýzy pro rámec *dbát₂* byly vytvořeny dvě závislosti – závislost *zákon na ochranu* má vysokou preferenci (kvazivalenční doplnění substantiva), závislost *dbát na ochranu* má preferenci nižší (volné doplnění slovesa).

Při výběru valenčního rámce *dbát₂* dochází k zajímavému jevu. Je vidět, že dvě slovesa s různými rámci, *dbát₁* a *dbát₂*, nejsou zcela nezávislé jednotky – zdá se, že pro člověka valenční požadavek slovesa *dbát₁* blokuje možnost neprojektivního doplnění substantiva ve větě se slovesem *dbát₂*. Při jiném slovosledném uspořádání je výběr slovesa *dbát₂* zcela relevantní, jak ukazuje následující věta:

Zákon na (jejich) ochranu dbá doporučení (Vrchního soudu).

(Jiné lexikální obsazení ovlivňuje obsahovou přijatelnost.) □

V následujícím příkladě ukážeme, jak se při analýze Pg uplatňuje povrchový slovosled a jakým způsobem jej zachycují slovosledná pravidla. Všimáme si slovosledných vzorců, která obsahují adjektiva.

Příklad:

(V.2) *Neschopnost dohodnout se na něm způsobila skluz v realizaci omezené autonomie na území Golanských výšin.* (PDT, blc02zu.fs, pozměněno)

(V.2') *Neschopnost dohodnout se na něm způsobila skluz v realizaci autonomie omezené na území Golanských výšin.* (PDT, blc02zu.fs, slovosledná varianta)

Věty (V.2) a (V.2') se liší pouze pozicí adjektiva *omezený*; tato změna je zásadní pro jejich (člověkem preferovaný) význam. Ilustrovali jsme na nich (označeny (IV.6b') a (IV.6b)) vliv slovosledných omezení a jejich oslabování na určení potenciální (morfematické) homonymie Pg. Ukázali jsme též odvození valenčního rámce deverbativního adjektiva *omezený*.

omezený ... Act (Ins) Eff (*na*+Acc)

Pg *na území (Golanských výšin)* nese víceznačnou morfologickou charakteristiku, jde o fonologickou homonymii *na*+Acc / *na*+Loc.

Podívejme se nyní na aplikaci pravidel pro RA, vždy na druhou část věty (předpokládáme, že v předchozích krocích analýzy je zjednodušena Pg *na území Golanských výšin*, pracujeme s redukovanou Pg *na území*).

Analýza vět založená na principu redukce se opět rozpadá do několika větví, redukce *na území* začíná v několika stavech.⁷⁰

Rozeberme nejprve podrobněji první krok analýzy, ve kterém je redukována Pg ve větě (V.2):

A. Pg *na území* je určena jako *na+Acc*.

... *způsobila skluz v realizaci omezené autonomie na území.*

(Část původní věty, zatím nebylo nic vypuštěno.)

1. Je zjištěn jeden valenční požadavek – *omezený na území*.
 2. Je vybrán WOP typu **VNPg**.
 3. Na základě WOP vznikají čtyři nové závislosti – *způsobit na území, skluz na území, realizace na území a autonomie na území*.
 4. Nejsou porušena formální kritéria.
 5. Je uvolněna podmínka projektivnosti pro rozvíjení adjektiva předložkovou skupinou
 6. Bylo vytvořeno 5 závislostí, Pg *na území* je redukována, analýza se rozpadá do pěti větví (v jedné má adjektivum *omezený* příslušný ‚slot‘ (Effect) označen jako zaplněný).
- (Tyto větve mohou vést k úspěšné analýze zkoumané věty.)

V dalších větvích analýzy je Pg *na území* zpracovávána později, až po vypuštění adjektiva, případně některých z předcházejících substantiv. Není však vytvořena žádná nová závislost.

Shrňme: Pg *na území* (*na+Acc*) může být ve větě (V.2) redukována jednak na základě slovosledných vzorců jako volné doplnění (vytvoří se čtyři závislosti se stejnou ‚váhou‘ volných doplnění *způsobit na území, skluz na území, realizace na území a autonomie na území*), jednak jako valenční doplnění adjektiva *omezený* (Eff), a to s nižší preferencí, neboť je v netypické pozici (tvoří neprojektivní podstrom).

B. Je-li Pg *na území* určena jako *na+Loc*, probíhá analýza založená na principu redukce obdobně, tato Pg však nesplňuje valenční požadavky adjektiva *omezený*, je tedy určena vždy jako volné doplnění.⁷¹

Redukce Pg ve větě (V.2') probíhá podobně, přiblížíme opět jen příslušné kroky některých větví RA:

A. Pg *na území* je určena jako *na+Acc*.

(a) ... *způsobila skluz v realizaci autonomie omezené na území.*

(Část původní věty, zatím nebylo nic vypuštěno.)

1. Je zjištěn jeden valenční požadavek – *omezený na území*.
 2. Je vybrán WOP typu **N_iA_iPg**.
 3. Na základě WOP nevzniká žádná nová závislost.
 4. Nejsou porušena formální kritéria.
 5. Vytvořená závislost splňuje slovosledná omezení.
 6. Byla vytvořena 1 závislost, Pg *na území* je redukována (příslušný ‚slot‘ adjektiva je označen jako zaplněný), analýza pokračuje dalším krokem.
- (Tato větev může vést k úspěšné analýze zkoumané věty.)

(b) ... *způsobila skluz v realizaci autonomie na území.*

⁷⁰ Uvádíme zde pouze některé větve, které vedou k úspěšné analýze – různé větve analýzy mohou vést ke stejnému rozboru věty (jak bylo řečeno výše, RA je nedeterministická, vzájemně nezávislé členy věty mohou být redukovány v libovolném pořadí). Určováním ostatních větví členů se nezabýváme, některé z nich mohou být redukovány v jiném pořadí.

⁷¹ Poznamenejme, že nemáme prostředky pro automatické ‚upřednostnění‘ analýzy Pg *na území* jako volného doplnění místa (Loc, kde, s morfologickou analýzou *na+Loc*) před analýzou *na+Acc*, která vyjadřuje směr (Direction, kam).

(Část původní věty, vypuštěno adjektivum *omezený*.)

1. Není zjištěn žádný valenční požadavek.
2. Je vybrán WOP typu **VNPg**.
3. Na základě WOP vznikají čtyři závislosti – *způsobit na území, skluz na území, realizace na území a autonomie na území*.
4. Nejsou porušena formální kritéria.
5. Vytvořené závislosti splňují slovosledná omezení.
6. Byly vytvořeny 4 závislosti, Pg *na území* je redukována, analýza se rozpadá do čtyř větvi.
(Tyto větve mohou vést k úspěšné analýze zkoumané věty.)

V dalších větvích analýzy je Pg *na území* zpracovávána později, až po vypuštění některých z předcházejících substantiv. Není však vytvořena žádná nová závislost.

Shrňme: Pg *na území* (*na+Acc*) může být ve větě (V.2') redukována jednak jako valenční doplnění adjektiva *omezený* (Eff), a to s nejvyšší preferencí (je naplněn WOP s preferovaným čtením pro adjektiva), jednak (s nízkou mírou spolehlivosti) jako volné doplnění – vytvoří se čtyři závislosti se stejnou ‚váhou‘ volných doplnění *způsobit na území, skluz na území, realizace na území a autonomie na území*.

B. Je-li Pg *na území* je určena jako *na+Loc*, probíhá analýza založená na principu redukce obdobně, tato Pg však nespĺňuje valenční požadavky adjektiva *omezený*, je tedy určena vždy jako volné doplnění.⁷²

Variety zkoumané věty jasně ukazují možnosti, které s sebou nese navržené vyhodnocování získaných závislostních struktur. Umožňuje stanovení preferované analýzy Pg a naopak upozadění málo spolehlivých závislostí. □

V následujícím příkladu se soustředíme na zpracování věty se slovesně jmenným výrazem.

Příklad:

(V.3) *Na nákup technického vybavení a učebnic poskytly Nadaci finanční prostředky obchodní společnosti.* (PDT, bm122zua.fs #21, pozměněn slovosled)

Věta (V.3) obsahuje slovesně jmenný výraz *poskytnout prostředky*, substantivum *prostředek* má ve svém valenčním rámci patient s povrchovým vyjádřením *na+Acc* (viz kapitola IV.2.3.; předpokládáme, že informaci o slovesně jmenných výrazech získáme ze slovníku).

poskytnout ... Act | Pat (Acc) Addr (Dat)

prostředky ... Pat (*na+Acc*)

Informaci o slovesně jmenném výrazu lze interpretovat jako pokyn redukovat celý tento výraz současně (s určením jediné závislosti Pg na jeho jmenné části, viz řešení navržené v kapitole IV.2.3.). Podívejme se, jak se tato skutečnost projeví v navrženém postupu RA.

Kvůli názornosti pracujeme s redukovanou Pg *na nákup*, dále předpokládáme, že v předchozích krocích byly též redukovány atributy *finanční a obchodní*. Analýza (už částečně redukováne) věty se rozpadá do několika větví.⁷³

¹³ Opět nelze automaticky upřednostnit analýzu Pg *na území* jako určení místa (*na+Loc*) před směrovým určením (*na+Acc*).

⁷³ V tento okamžik není zajímavá fonologická homonymie substantiv; bereme v úvahu např. větve s morfologickou charakteristikou *Nadaci*_{Dat}, *prostředky*_{Acc} a *společnosti*_{Nom}.

(a) Na nákup poskytly Nadaci prostředky společnosti.

(Výchozí věta.)

1. Je zjištěn jeden valenční požadavek – *prostředky na nákup*.
2. Je vybrán WOP typu **PgV**.
3. Na základě WOP vzniká závislost *poskytnout na nákup*.
4. Nejsou porušena formální kritéria.
5. Jsou splněny podmínky pro uvolnění slovosledných omezení, analýza Pg *prostředky na nákup*, uspokojující valenční požadavek substantiva, je zachována. Sloveso a substantivum tvoří slovesně jmenný výraz, je vyloučena závislost *poskytnout na nákup*.
6. Byla vytvořena 1 závislost, Pg *na nákup* je redukována, analýza pokračuje dalším krokem. (Tato větev může vést k úspěšné analýze zkoumané věty.)

Větné členy *na nákup, Nadaci, společnosti* jsou syntakticky nezávislé, nezáleží tedy na pořadí jejich vypouštění (na základě WOP nejsou substantiva *nadace* a *společnost* určeny jako řídicí členy Pg, není důvod uvolňovat slovosledná omezení). Přiblížíme až následující možné redukční kroky, ve kterých je zpracovávána Pg (v příslušných větvích RA).

(b) Na nákup poskytly prostředky.

(Částečně redukováná věta (např. s analýzou *společnosti* (Act), *Nadaci* (Addr).)

1.-6. Jako v kroku (a).

(Tato větev vede k úspěšné analýze zkoumané věty.)

(c) Na nákup poskytly.

(Částečně redukováná věta (např. s analýzou *společnosti* (Act), *Nadaci* (Addr), *prostředky* (Pat).)

1. Není zjištěn žádný valenční ani kvazivalenční požadavek .
2. Je vybrán WOP typu **PgV**.
3. Na základě WOP vzniká závislost *poskytnout na nákup*.
4. Nejsou porušena formální kritéria.
5. Sloveso a substantivum tvoří slovesně jmenný výraz, je vyloučena závislost *poskytnout na nákup*.
6. Nebyla zachována žádná závislost, Pg *na nákup* není vypuštěna, krok analýzy není úspěšný, (Tato větev není úspěšná – RA pokračuje jinou větví.)

Shrňme: V původní nezjednodušené větě byl identifikován slovesně jmenný výraz, *poskytnout prostředky*. Pg *na nákup* byla určena jako valenční doplnění substantiva *prostředky*, s vysokou mírou spolehlivosti. Podle navrženého řešení (viz kapitulu IV.2.3.) byla vyloučena adverbální závislost Pg, která má stejný význam (stejně pravdivostní podmínky) jako její atributivní analýza. (Pg ovšem interpretujeme jako větný člen rozvíjející celý slovesně jmenný výraz.)

Krok (c) ilustruje situaci, kdy je řídicí větný člen redukován dříve než člen závislý, v našem případě Pg. □

Následující příklad ukazuje aplikaci pravidel a určení preferencí pro jednotlivá čtení věty, ve které po slovese a substantivu následují dvě předložkové skupiny. Při vyhodnocování výsledných struktur jsou uplatněna pravidla využívající sémantických rysů substantiv.

Příklad:

(V.4) *Ministerstvo školství vypsalo výběrové řízení na historickou budovu na Brusnici.* (PDT, bm102zua.fs #18, zkráceno)

Na větě (V.4) (označená (IV.33)) jsme demonstrovali pravidlo pro substantiva s rysy ‚konkr‘ a ‚místo‘ + ‚jméno‘.

Substantivum *řízení* má jeden kvazivalenční požadavek:

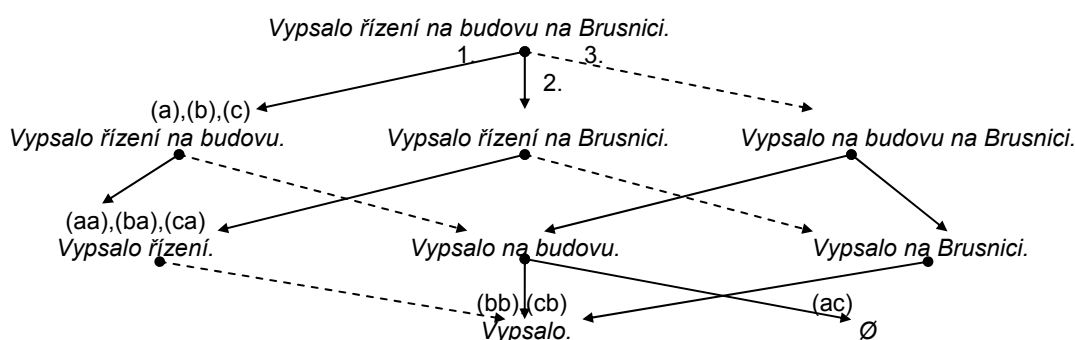
řízení ... ^Regard (na+Acc)⁷⁴

Kvůli názornosti předpokládáme, že v předchozích krocích byly redukovány atributy *školsví, výběrové a historickou*. (Nespecifikujeme všechny kombinace, ve kterých jsou vypouštěny syntakticky nezávislé větné členy.) Dále předpokládáme, že bylo vypuštěno substantivum *ministerstvo* (aktor slovesa *vypsát*.)⁷⁵

Je ovšem potřeba brát v úvahu (fonologickou) homonymii Pg *na Brusnici*, které jsou přiřazeny dvě morfologické analýzy – jde buď o *na+Acc*, nebo o *na+Loc* (výběru morfologických kategorií budou odpovídat i případná volná doplnění (Kde? pro *na+Loc* / Kam, příp. účel, Aim pro *na+Acc*).

Popíšme kroky některých větví RA, ve kterých může docházet k redukci Pg *na budovu* a Pg *na Brusnici*, a přínos sémantických preferencí.

A. Nejprve pracujeme s Pg *na Brusnici* analyzovanou jako *na+Acc*.



Obr.31: Některé kroky RA, ve kterých mohou být redukovány Pg *na budovu* a *na Brusnici*⁷⁶

První větev (RA začíná redukcí Pg *na Brusnici*):

Vypsalo řízení na budovu na Brusnici.

(Výchozí věta.)

1. Je zjištěn jeden kvazivalenční požadavek substantiva – řízení na Brusnici.
2. Je určen WOP typu **VNPg**.
3. Na základě WOP vzniknou nové závislosti *vypsát na Brusnici* a *budova na Brusnici*.
4. Nejsou porušena formální kritéria.
5. Vytvořené závislosti splňují slovosledná omezení.
6. Byly vytvořeny 3 závislosti, Pg *na Brusnici* je redukována, analýza se rozpadá do tří větví (v jedné větví má substantivum *řízení* příslušný 'slot' označen jako zaplněný).
(Tato větev může vést k úspěšné analýze, pokračuje dále jedním z kroků v bodech (a), (b) nebo (c).)

(a) Nejdříve popíšeme případ, kdy je Pg *na Brusnici* určena jako doplnění substantiva *řízení* (^Regard):

(aa) *Vypsalo řízení na budovu.*

(Částečně redukována věta – Pg *na Brusnici* kvazival. zřetel substantiva *řízení* (zaplňuje kvazivalenční požadavek).)

⁷⁴ Tento požadavek je patrný zejména v kolokaci *výběrové řízení*, viz i příklad (IV.24a).

⁷⁵ Opět se nezabýváme fonologickou homonymií substantiv, pracujeme s větvemi, kde je určeno *ministerstvo*_{Nom} a *řízení*_{Acc}.

⁷⁶ Plně šipky odpovídají krokům RA, ve kterých je redukována jedna ze zkoumaných Pg – číslice označují větvě analýzy, písmena popisované kroky. Přerušované šipky označují kroky analýzy, ve kterých se redukuje jiný větný člen než Pg.

1. Není zjištěn žádný valenční ani kvazivalenční požadavek.
2. Je určen WOP typu **VNPg**.
3. Na základě WOP vzniknou dvě nové závislosti – *vypsát na budovu* a *řízení na budovu*.
4. Nejsou porušena formální kritéria.
5. Vytvořené závislosti splňují slovosledná omezení.
6. Byly vytvořeny 2 závislosti, Pg *na budovu* je redukována, analýza se rozpadá do dvou větví.
(Struktura se závislostí *vypsát na budovu* nevede k úspěšné analýze – ‚princip separace‘ znemožní vytvoření závislosti *vypsát řízení*. Druhá větev vede k úspěšné analýze.)

(ab) *Vypsalo řízení na budovu.*

Alternativa pro bod (aa), kdy je napřed redukováno substantivum *řízení*, ‚princip separace‘ znemožní redukci Pg:

Vypsalo na budovu.

(Částečně redukováná věta – Pg *na Brusnici* kvazival. zřetel substantiva *řízení*, které je patientem slovesa.)

1. Není zjištěn žádný valenční ani kvazivalenční požadavek.
2. Je určen WOP typu **VPg**.
3. Na základě WOP vznikne závislost *vypsát na budovu*.
4. Porušen ‚princip separace‘ – analýza vyloučena.
- 5.-6. Nebyla vytvořena žádná závislost, Pg *na budovu* není redukována, analýza pokračuje jinou větví.

(b) Nyní se soustředíme na případ, kdy je Pg *na Brusnici* určena jako neshodný atribut substantiva *budova* (podporováno sémantickým pravidlem):

(ba) *Vypsalo řízení na budovu.*

(Částečně redukováná věta – Pg *na Brusnici* atribut substantiva *budova*.)

1. Je zjištěn jeden kvazivalenční požadavek substantiva – *řízení na budovu*.
2. Je určen WOP typu **VNPg**.
3. Na základě WOP vznikne nová závislost *vypsát na budovu*.
4. Nejsou porušena formální kritéria.
5. Vytvořené závislosti splňují slovosledná omezení.
6. Byly vytvořeny 2 závislosti, Pg *na budovu* je redukována, analýza se rozpadá do dvou větví (v jedné větví má substantivum *řízení* příslušný ‚slot‘ označen jako zaplněný).

(Obě větve se závislostmi *vypsát na budovu* a *řízení na budovu* mohou vést k úspěšné analýze.)

(bb) *Vypsalo řízení na budovu.*

Alternativa pro bod (ba), kdy je napřed redukováno substantivum *řízení*:

Vypsalo na budovu.

(Částečně redukováná věta – redukováno substantivum *řízení* (Pat slovesa), Pg *na Brusnici* neshodný atribut substantiva *budova*.)

1. Není zjištěn žádný valenční ani kvazivalenční požadavek.
2. Je určen WOP typu **VPg**.
3. Na základě WOP vznikne závislost *vypsát na budovu*.
4. Nejsou porušena formální kritéria.
5. Vytvořená závislost splňuje slovosledná omezení.
6. Byla vytvořena jedna závislost, Pg *na budovu* je redukována, analýza je úspěšná.

(Tento krok vede k úspěšné analýze zkoumané věty – zkoumaná věta je redukována na predikát *Vypsalo*.)

(c) Pg *na Brusnici* může být určena také jako volné doplnění slovesa.

(ca) *Vypsalo řízení na budovu.*

(Částečně redukovaná věta – Pg *na Brusnici* volné doplnění slovesa *vypsát*.)

1.-6. Jako v kroku (ba).

(Obě větve se závislostmi *vypsát na budovu* a *řízení na budovu* mohou vést k úspěšné analýze.)

(cb) *Vypsalo řízení na budovu.*

Alternativa pro bod (ca), kdy je napřed redukováno substantivum *řízení*:

Vypsalo na budovu.

(Částečně redukovaná věta – redukováno substantivum *řízení* (Pat slovesa), Pg *na Brusnici* volné doplnění slovesa *vypsát*.)

1.-6. Jako v kroku (bb)

(Tento krok vede k úspěšné analýze zkoumané věty – zkoumaná věta je redukována na predikát *Vypsalo*.)

V popsáných větvích RA bylo vytvořeno pět různých syntaktických struktur (shrňme v závěru příkladu).

Druhá větev (RA začíná redukcí Pg *na budovu*):

Analýza probíhá obdobně, vytvořena jediná nová závislostní struktura, ve kterém je Pg *na budovu* určena jako kvazivalenční doplnění substantiva *řízení* a Pg *na Brusnici* jako volné doplnění tohoto substantiva (obdoba jednoho ze stromů z větve (aa), s prohozeným valenčním a volným doplněním).

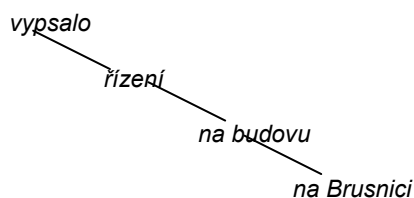
Třetí větev (RA začíná redukcí substantiva *řízení*):

Analýza probíhá opět podobně, opět nejsou vytvořeny závislostní struktury, které by nebyly vytvořeny již v popsáných krocích.

B. Pro Pg *na Brusnici* analyzovanou jako *na+Loc* probíhá analýza založená na principu redukce obdobně, tato Pg však nesplňuje kvazivalenční požadavky substantiva *řízení*, je tedy určena vždy jako volné doplnění.

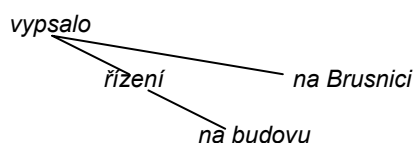
Na závěr shrňme – výsledkem všech větví RA je šest syntaktických struktur, které jsou následujícím způsobem uspořádány podle spolehlivosti uplatněných kritérií (soustředíme se zde pouze na určení ‚tvaru‘ závislostních stromů a na rozlišení valenčních a volných doplnění, nezabýváme se určováním typů volných doplnění, které souvisí s morfológickou charakteristikou zkoumaných Pg):

1. Nejvyšší preferenci má analýza získaná ve větvi obsahující krok (ba) – Pg *na Brusnici* tvoří neshodný atribut substantiva *budova* (volné doplnění, podporováno sémantickými rysy), Pg *na budovu* je určena jako kvazivalenční doplnění substantiva *řízení*.



Analýzy 2.-4. jsou ohodnoceny jako stejně spolehlivé:

2. analýza získaná ve větvi obsahující např. krok (ca) – Pg *na Brusnici* je volné doplnění slovesa, Pg *na budovu* je kvazivalenčním doplněním substantiva *řízení*;

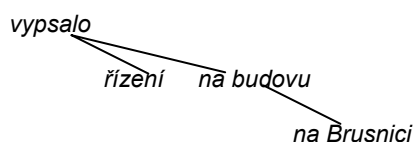


3. analýza získaná v B. – Pg *na budovu* je určena jako kvazivalenční doplnění substantiva *řízení*, Pg *na Brusnici* tvoří neshodný atribut téhož substantiva *řízení* (volné doplnění);

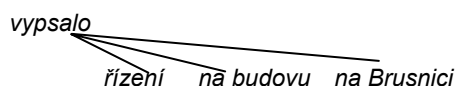
4. analýza získaná ve větvi obsahující (aa) – Pg *na budovu* tvoří neshodný atribut (volné doplnění) substantiva *řízení*, Pg *na Brusnici* je určena jako kvazivalenční doplnění téhož substantiva *řízení* (pouze pro Pg *na Brusnici* ... *na*+Acc).



5. Následuje analýza získaná např. ve větvi obsahující krok (bb) – Pg *na Brusnici* určeno jako volné doplnění substantiva *budova* (podporováno sémantickými rysy substantiv), Pg *na budovu* určeny jako volné doplnění slovesa.



6. Poslední je analýza získaná např. ve větvi obsahující krok (cb) – obě Pg *na budovu* a *na Brusnici* určeny jako volná doplnění slovesa *vypsát*.



Analýza s nejvyšší ‚váhou‘ (bod 1.) se shoduje s preferovaným čtením na základě obsahu věty (s morfologickou analýzou *na budovu* ... *na*+Acc, *na Brusnici* ... *na*+Loc).

Analýza (bod 2.) s Pg *na budovu* určenou jako kvazivalenční doplnění substantiva *řízení* (*na*+Acc) a Pg *na Brusnici* jako volné doplnění slovesa (*na*+Loc) je také přijatelná, i když méně pravděpodobná.

Analýzy 3. a 4. se liší pořadím kvazivalenčního a volného doplnění jednoho substantiva – zatím nejsou stanoveny podmínky, které by umožňovaly porovnat jejich spolehlivost.

Poslední dvě analýzy (bod 5. a 6.) nejsou přijatelné (jsou v nich upřednostněna volná doplnění před doplněními kvazivalenčními). □

Shrnutí

Jednotlivá kritéria zkoumaná v části IV. pracují s různými jazykovými informacemi, mají různý dosah a různou spolehlivost. Pro jejich využití v automatickém systému je zásadní jejich uspořádání, určení, v jakém pořadí se jednotlivé informace zpracovávají.

V této kapitole jsme předložili návrh uspořádání jednotlivých lingvistických typů kritérií a způsob jejich využití pro určení větných členů, které mohou být rozvíjeny předložkovou skupinou, v mechanismu analýzy založené na principu redukce. RA je pro

konkrétní úkol, řešení (morfematické) homonymie P_g , realizována jako posloupnost kroků, v nichž jsou uplatňována tato kritéria. Cílem je vyčerpávající analýza, která určí všechny adekvátní závislosti P_g .

Velmi významná je možnost vyhodnocovat získané syntaktické struktury podle jejich spolehlivosti, která je dána typy pravidel, které vedly k vytvoření té které závislosti. Stanovení preferencí pro jednotlivá „čtení“ zkoumané P_g umožňuje uspořádat výsledné analýzy podle jejich předpokládané adekvátnosti.

Na příkladech analýzy předložkových skupin v několika (již dříve z jazykového hlediska rozebíraných) větách byla podrobně popsána postupná aplikace jednotlivých pravidel. Ilustrovali jsme možnost uspořádat získané analýzy podle jejich spolehlivosti (podle typu použitého pravidla) a adekvátnost takového uspořádání.

Zdá se, že není vhodné stanovit předem pevnou hranici mezi analýzami ještě přijatelnými a analýzami, které jsou již velmi málo pravděpodobné nebo nepřijatelné. Tato hranice je různá pro jednotlivé věty, v závislosti na konkrétní kombinaci využitých pravidel.

Podrobný rozbor simulující automatickou analýzu vede k dalšímu zjemňování pravidel pro řešení jednotlivých dílčích problémů spojených s automatickou syntaktickou analýzou, ke shromažďování podkladů, které povedou k zpřesnění formálního popisu syntaxe.

ČÁST VI.

TESTOVÁNÍ NAVRŽENÝCH KRITÉRIÍ

VI.1. GRAMATIKA A JEJÍ INTERPRET

Šestá část je věnována možnosti aplikace zkoumaných lingvistických kritérií pro určení (potenciálních) závislostí předložkových skupin v rámci konkrétní automatické procedury. Za základ takové procedury jsme zvolili gramatiku navrženou pro již výše zmíněný Robustní parser pro češtinu, který popisuje Kuboň (1999), a její interpret (Holan (2001)). Formalismus RFODG, který tato gramatika a interpret využívají, popisují např. Plátek et al (1996), (1997). V následujících kapitolách přiblížíme úpravy, které umožňují automaticky zpracovávat předložkové skupiny podle kritérií popsaných v předchozích částech této práce.

Shrňme zde základní skutečnosti, které jsou podstatné pro využití robustního parseru ke zpracování předložkových skupin.

Robustní parser pro češtinu byl navržen pro účely automatického vyhledávání chyb (grammar checkeru), tedy pro dílčí rozpoznávání syntaktické správnosti vět, a jako takový musí zpracovávat nejen syntakticky správně utvořené věty (well-formed sentences), ale i věty se syntaktickou nekonzistencí (ill-formed sentences). Soustřeďuje se proto především na čistě syntaktická pravidla, pravidel využívajících vlastností jednotlivých lexikálních jednotek využívá v míře co nejmenší. Jeho podstatným rysem je skutečnost, že jádro jazyka popisuje relativně jednoduchou gramatikou, která umožňuje časově i prostorově nepřiliš náročné zpracování vět a odhalení případných syntaktických chyb.

Gramatika robustního parseru je vhodná pro testování navržených kritérií z jednoho prostého, nicméně podstatného důvodu – předložková skupina je podle ní vždy zpracovávána jako volné doplnění, možné závislosti Pg jsou určovány pouze na základě meta-pravidla (viz níž) pro zapojení volného doplnění zprava a meta-pravidla pro zapojení volného doplnění zleva,⁷⁷ která reflektují jen slovoslednou pozici Pg. Robustní parser nepracuje s valenčními doplněními ve formě předložkových skupin. Jeho autor zdůvodňuje toto rozhodnutí úkolem, pro který byl parser navržen – tedy možností zpracovávat i chybné věty. V takových větách může být nekonzistence způsobena právě neadekvátním výběrem valenčního členu ve formě předložkové skupiny (jeho předložky nebo pádu). Aby byl parser schopen vyřešit tento typ syntaktické chyby, jsou valenční doplnění vyjádřená Pg zpracovávána stejným způsobem jako doplnění volná. Větné členy vyjádřené předložkovou skupinou jsou tedy analyzovány jako doplnění nejbližšího slovesa nebo podstatného jména, které jim předchází v povrchové realizaci

⁷⁷ Jedinou výjimku tvoří meta-pravidlo pro selektivní konstrukce u číslovek a adjektiv v superlativu – *jeden z chlapců, nejlepší z nich*.

věty, resp. jako doplnění následujícího slovesa (viz princip ‚last attachment‘, oddíl I.2.). Přitom je ponechána možnost doplnit k této základní struktuře (která tvoří kostru, ‚skeleton‘ množiny všech možných syntaktických závislostí) i další struktury, ve kterých Pg rozvíjí jiná substantiva, příp. slovesa, která jí předcházejí. Jde svým způsobem o zachycení regularity morfematické homonymie, které je zde věnována kapitola o prototypické pozici Pg, viz IV.1.1.

Kritéria zkoumaná v předchozích kapitolách podstatným způsobem překračují intence, se kterými byla navržena gramatika robustního parseru. Základní slovosledná pravidla pro Pg doplňují především pravidla umožňující analýzu Pg jako valenčního, případně kvazivalenčního doplnění, a to i v neprototypických slovosledných pozicích. Navíc mohou být jednotlivé analýzy Pg ohodnoceny a uspořádány podle své ‚spolehlivosti‘.

Chceme-li se tedy soustředit na zpracování Pg, můžeme gramatiku robustního parseru využít pro základní analýzu celé věty. Protože všechny Pg jsou podle této gramatiky řešeny jednotně jako volná doplnění, není třeba ‚odfiltrovávat‘ takto vytvořené závislosti – stanovené závislosti (Pg potenciálně závisí na slovese i na všech substantivech ležících na cestě mezi tímto slovesem a ‚nejbližším‘ substantivem, resp. na následujícím slovese) odpovídají při vhodném nastavení parametrů (parametry ‚projekt‘ a ‚nejbližsi‘, viz níže) závislostem vytvořeným na základě navržených slovosledných vzorců (pokud nejsou překryta skutečností, že daná závislost už byla určena jako valenční doplnění). Pracujeme s verzí gramatiky z června 2000.

Interpret gramatiky byl původně implementován T. Holanem pro potřeby projektu LATESLAV (viz Plátek et al (1996)), později byl upraven pro potřeby robustního parseru. Používáme verzi z podzimu 2000, kterou popisuje Holan (2001).

Gramatika robustního parseru i její interpret umožňují volit z široké nabídky **přepínačů**, pomocí nichž lze určovat průběh analýzy. Je možné určovat složitost výsledných struktur – lze zadávat požadavky na projektivitu závislostních stromů, případně určit míru jejich neprojektivity. (Vhodné míry neprojektivity definují Holan et al (1998)). Obdobná omezení se dají vztáhnout i k jednotlivým syntaktickým meta-pravidlům. To umožňuje zvolit takové nastavení, které respektuje potřeby různých úloh.

Připomeňme zde explicitně, že používáme gramatiku robustního parseru ke zcela jiným účelům, než pro které byla vyvinuta. Odtud také pramení jiné priority a jiné požadavky, které vedou k jejím modifikacím. Nutno také říci, že využíváme pouze její pozitivní část, tj. zpracováváme pouze zaručeně správné věty, a pouze ty z možných konfigurací interpretu, které odpovídají potřebám stanoveným při zpracování Pg.

VI.1.1. Úpravy gramatiky

Formální gramatika je zachycena pomocí **meta-pravidel**, která jsou srozumitelná pro člověka a která reprezentují vždy určitou množinu pravidel RFODG (Robust Free-Order Dependency Grammar, což je formalismus navržený jako nástroj pro popis vět se syntaktickou nekonzistencí pro jazyky s vysokou mírou volnosti slovosledu; jeho přesnou definici uvádějí Plátek et al (1997)).

K jednotlivým meta-pravidlům lze stanovit omezení na pozici vstupních položek, a to pomocí parametrů ‚projekt‘ a ‚nejbližsi‘ v hlavičce meta-pravidla:

- parametr ‚projekt‘ určuje, že dané meta-pravidlo lze využít pouze k vytvoření projektivní konstrukce;

- parametr ‚nejbliži‘ znamená, že dané meta-pravidlo lze aplikovat pouze na dvě nejbližší slova, která splňují podmínky meta-pravidla (umožňuje např. vytvořit výše zmíněný ‚skeleton‘ při určování volných doplnění).

Výsledné syntaktické struktury jsou zachyceny závislostním stromem, který se ovšem poněkud liší od tradiční závislostní struktury (např. každý vstupní symbol – slovní formy, interpunkční znaménka, ... – je reprezentován jedním uzlem). Definici DR-stromu (DR-tree, popisuje historii uplatňování jednotlivých meta-pravidel) i závislostního stromu (De-tree, zachycuje výsledné struktury) lze nalézt u Plátka et al (1997). Hrany stromů vždy nesou číslo pravidla, podle kterého byly vytvořeny.

Úpravy meta-pravidel gramatiky

Pro testování kritérií navržených pro zpracování předložkových skupin jsme navrhli některé změny v gramatice robustního parseru. V následujících odstavcích tyto změny přiblížíme.

Valenční doplnění vyjádřená Pg

Valenční rámce sloves, substantiv a adjektiv jsou zachyceny v atributu *frameset*. Jde o jediný atribut, který může mít komplexní hodnoty – skládá se ze ‚slotů‘, které popisují jednotlivá valenční doplnění, jejich typ (atribut *actant*) a povrchové vyjádření (atribut *case* udává pád nominálního doplnění, příp. skutečnost, že daná pozice může být zaplněna infinitivem nebo klauzí, atribut *prep* určuje předložku v případě, že aktant je vyjádřen Pg).⁷⁸

Příklad:

Valenční rámec slovesa *způsobit*, který sestává z aktoru, pacientu a fakultativního adresátu, bude zachycen atributem *frameset* následujícím způsobem:

```
způsobit ... Act | Pat (Acc/klauze-aby, že) (Addr (Dat))
frameset: ?
( [ actant: act case: nom prep: 0 ]
  [ actant: pat case: acc prep: 0 ]
  [ actant: adr case: dat prep: 0 ] )
,
( [ actant: act case: nom prep: 0 ]
  [ actant: pat case: clause prep: aby ]
  [ actant: adr case: dat prep: 0 ] )
,
( [ actant: act case: nom prep: 0 ]
  [ actant: pat case: clause prep: že ]
  [ actant: adr case: dat prep: 0 ] )
!
```

Mezi symboly ? a ! je uveden seznam možných povrchových vyjádření valenčního rámce, oddělených čárkou (,). Kulaté závorky označují vždy jednu možnou instanci rámce, hranaté závorky odpovídají vždy jedné valenční pozici (‚slotu‘). □

⁷⁸ Pravidla výchozí gramatiky ovšem tuto možnost nevyužívají.

Požadavek umožnit analýzu Pg jako valenčního doplnění slovesa, substantiva nebo adjektiva vedl k doplnění meta-pravidla 7 (naplnění rámce zprava), meta-pravidla 8 (naplnění rámce zleva) a meta-pravidla 24 (naplnění rámce zleva vztažným zájmenem):

- Při testování podmínek na vstupní položky je připuštěna možnost doplnit za závislý člen předložkovou skupinu (atribut *syntcl* = *prephr*).

- Kontroluje se shoda předložky a pádu substantiva tvořících jádro Pg (viz oddíl II.1.), která je potenciálním valenčním doplněním.

- Původní meta-pravidlo 7 nepřipouští valenční doplnění substantiva v genitivu. Rozvíjení substantiva zprava genitivem je řešeno zvláštním meta-pravidlem na připojení genitivu zprava k substantivu (11) (je zaručena jedinečnost a projektivita takového připojení). Toto řešení je zachováno.

Takovou modifikaci meta-pravidel gramatiky lze připustit, neboť se omezujeme pouze na zpracování správně utvořených vět (nemusíme tedy brát v úvahu syntaktické nekonzistence způsobené Pg, což vedlo autora robustního parseru k rozhodnutí zpracovávat je jako volná doplnění).

Kvazivalenční doplnění

Zavedení kvazivalenčních doplnění, která mají vzhledem k danému úkolu obdobné syntaktické vlastnosti jako doplnění valenční, se (zatím) neodrazilo v meta-pravidlech gramatiky. Tato doplnění jsou zachycena ve valenčním rámci příslušného řídicího slova (slovesa, substantiva nebo adjektiva), získané závislosti jsou vyhodnocovány ručně.

Myšlenka nového atributu pro slovesa, substantiva a adjektiva (komplexního atributu *kvazi*, který by měl stejnou strukturu jako atribut *frameset*) naráží na omezení interpretu, který zatím připouští pouze jediný komplexní atribut. Pokud by bylo toto omezení odstraněno, zavedení kvazivalenčních rámců by si vyžádalo tři nová meta-pravidla:

- Meta-pravidla pro naplnění kvazivalenčního rámce sloves, substantiv a adjektiv se liší od upravených meta-pravidel pro naplnění valenčního rámce sloves, substantiv a adjektiv (upravených meta-pravidel 7, 8 a 24) pouze záměnou původního atributu *frameset* novým atributem *kvazi*, který má stejnou syntax, a testováním tohoto nového atributu.

Toto zdvojení meta-pravidel pro naplnění rámce by dovolovalo automaticky stanovit různou preferenci pro doplnění získaná na základě valenčních a kvazivalenčních požadavků.

Jinou možností, jak zachytit kvazivalenční doplnění, je uchovávat valenční i kvazivalenční doplnění v jediném rámci, přičemž by ke každému ‚slotu‘ přibyl atribut, který by určoval typ doplnění.

Zachycení obligatorních volných doplnění

V předchozích částech jsme viděli, že předložkové skupiny často vyjadřují volná doplnění. Jde především o doplnění fakultativní, ale nezájmena i o doplnění obligatorní, která jsou součástí valenčního rámce. To vede k úvahám, zda a jakým způsobem zachytit v daném systému alespoň prototypické formy takových doplnění.

Nabízí se myšlenka využít ‚primárních funkcí‘ jednotlivých předložkových skupin (viz kapitulu IV.3.) – na jejich základě lze určit typické doplnění, které daná Pg vyjadřuje. V popisovaném systému je možné takovou informaci zachytit následujícím způsobem:

- Slovníkové informace (vstupní data):

- U předložek lze uvádět ‚primární funkci‘ v komplexním atributu (např. *sem*); takový atribut by byl podobný atributu *frameset* (k jednotlivým předložkám je potřeba uvést kromě typu doplnění i pád substantiva tvořícího s předložkou jádro Pg (tedy ‚slot‘ tvořený atributy *aktant* a *case*)).

- Při uvádění valenčního rámce, *frameset*, je možné kódovat obligatorní volná doplnění např. v atributu *aktant* (atributy *case* a *prep* by v tomto případě měly hodnotu 0); hodnota atributu *case* různá od nuly v takovém případě signalizuje vnitřní doplnění, hodnota *case* = 0 znamená obligatorní volné doplnění.

- Při aplikaci meta-pravidla pro vytvoření Pg (meta-pravidlo 6, *syntcl* = *prephr*) by se ve zvláštním atributu uchovávala informace o prototypické funkci zkoumané Pg (např. *prephr_type*).

- Při zaplňování valenčních rámců (upravená meta-pravidla 7, 8, 24) by se kromě vnitřních doplnění testovala i prototypická volná doplnění; pokud se ve *frameset* vyskytuje volné doplnění (atribut *case* = 0) určitého typu (atribut *aktant*), je nutné hledat položku reprezentující Pg (*syntcl* = *prephr*) patřícího typu (*prephr_type*).

- Analogicky pro zaplňování kvazivalenčních rámců (nová meta-pravidla kopírující upravená meta-pravidla 7, 8 a 24).

Tomuto řešení zatím brání omezení interpretu (omezení na jediný komplexní atribut).

Volná doplnění adjektiv

Gramatika robustního parseru neumožňuje rozvíjení adjektiva volným doplněním. Proto bylo doplněno meta-pravidlo pro rozvíjení adjektiva zprava a meta-pravidlo pro rozvíjení adjektiva zleva.

- Meta-pravidla pro rozvíjení adjektiva zprava, resp. zleva jsou obdobou meta-pravidel pro volná doplnění zprava (9, pro slovesa a substantiva), resp. zleva (10, pro slovesa), s parametry ‚projekt‘ a ‚nejblízi‘.

Alternativou k tomuto řešení bylo povolit v původních meta-pravidlech 9 a 10 adjektivum jako člen rozvíjený volným doplněním. Dáváme však přednost vytvoření nových meta-pravidel – toto řešení umožňuje zadávat pro slovesa, substantiva a adjektiva různé podmínky na projektivitu vytvořených podstromů.

Parametry ‚projekt‘ a ‚nejblízi‘

Při testování navržených pravidel pro zachycení homonymie předložkových skupin připouštíme možnost neprojektivní konstrukce jen u Pg (jak byla popsána v kapitole IV.1.2, tedy především neprojektivitu související s valenčními doplněními); meta-pravidla, která Pg nezpracovávají, se aplikují vždy projektivně (meta-pravidla 2-5,9,10,12-18,27-30,32,33,39-43,47).

Kvůli odhalení všech (potenciálních) závislostí volných doplnění (nejen ‚skeletonu‘ množiny všech možných syntaktických závislostí, viz výš) je meta-pravidlo pro zapojení volného doplnění zprava (meta-pravidlo 9) aplikováno na všechny výskyty

vstupních položek (tedy nejen na jejich nejbližší výskyty), pouze s omezením na projektivitu výsledné struktury.

Vyloučení neadekvátních závislostí a evaluace získaných analýz

V předchozích odstavcích byly popsány úpravy gramatiky robustního parseru, které (oproti původní gramatice) umožňují analyzovat předložkové skupiny nejen jako volná doplnění, ale i jako valenční nebo kvazivalenční doplnění sloves, substantiv a adjektiv. Navržená kritéria ovšem navíc předpokládají odstranění neadekvátních analýz Pg – to se týká vyloučení duplicit valenční / volné doplnění, kontroly slovosledných omezení a splnění formálních kritérií. Závislostní stromy, ve kterých je Pg duplicitně analyzována jako volné doplnění, jsou vyloučeny automaticky (jsou zachovány pouze struktury, ve kterých je Pg určena jako valenční doplnění příslušného větného členu)⁷⁹. Ostatní kroky, ve kterých jsou odstraněny neadekvátní analýzy, jsou (zatím) na výsledných závislostních stromech prováděny ručně.

Na závěr jsou výsledné závislostní stromy předběžně automaticky utříděny, výsledné ohodnocení podle zásad popsaných v oddíle V.1. je stanoveno manuálně.

⁷⁹ Za praktické provedení děkujeme autorovi interpretu T. Holanovi, který se postaral i o předběžné automatické třídění.

VI.2. PŘÍKLADY AUTOMATICKÉ ANALÝZY

Vraťme se k příkladům podrobně rozebíraným v části V. a ukažme, jakým způsobem budou tyto věty zpracovány pomocí přizpůsobené gramatiky robustního parseru a jaké konfigurace jejího interpretu je vhodné zvolit. Soustředíme se zde opět pouze na určení ‚tvaru‘ závislostních stromů a na rozlišení valenčních a volných doplnění, nezabýváme se určováním typů volných doplnění.

Příklad:

(VI.1) *Zákon dbá na ochranu nájemníků.* (PDT, bl101js.fs #18, zkráceno, viz (V.1))

Příklad (VI.1) ilustruje práci s valenčními a kvazivalenčními rámci slovesa a substantiva. (Připomeňme, že sloveso *dbát* má dva valenční rámce, *dbát*₁ ... Act | Pat (*na+Acc/o+Acc*) a *dbát*₂ ... Act | Pat (Gen). Substantivum *zákon* má valenční i kvazivalenční doplnění, *zákon* ... Pat (*o+Loc*) a *zákon* ... ^Aim (*na+Acc*).

Ve vstupních datech (po sloučení morfologických a syntaktických údajů pro jednotlivé slovní formy) jsou všechny valenční i kvazivalenční informace (týkající se všech rámců) zachyceny v jediném atributu *frameset* – jednotlivé ‚sloty‘ popisují všechny možné kombinace doplnění, se všemi povrchovými formami.

V kapitole IV.1.2. byla navržena relaxace podmínky projektivnosti pro valenční a kvazivalenční doplnění. Proto je zajímavá analýza, která neprojektivní konstrukce umožňuje. Při takovém zadání je vytvořeno pět výsledných závislostních stromů.

Jeden z pěti získaných stromů (strom 4) je automaticky vyfiltrován (duplicita valenční – volné doplnění).

The screenshot shows two windows: 'F0brazek' and 'FSeznam'.

F0brazek displays a dependency tree for the sentence 'zakon dba na ochranu najemniku tecka'. The root node is 'dba' with frameset 'B:0(A:8,D:7(C:6,E:11),F:48) 1000'. It branches into three children: 'zakon' (frameset 8), 'ochranu' (frameset 7), and 'tecka' (frameset 48). The 'ochranu' node further branches into 'na' (frameset 6) and 'najemniku' (frameset 11).

FSeznam displays a list of framesets for the sentence 'zakon dba na ochranu najemniku tecka'. It lists 13 DR-Tvary and 5 De-Tvary. The 'Vyhazuji:' section shows two framesets being filtered out: '4: 0 B:0(A:8,D:9(C:6,E:11),F:48)' and '2: 0 B:0(A:8,D:7(C:6,E:11),F:48)'. The 'Tridim:' section shows the remaining framesets: '0: 0 B:0(A:8(D:7(C:6),E:7,F:48) 1000', '1: 0 B:0(A:8(D:7(C:6,E:11),F:48) 1000', '2: 0 B:0(A:8,D:7(C:6,E:11),F:48) 1000', and '3: 0 B:0(A:8,D:9(C:6),E:7,F:48) 100'.

Manuální vyhodnocení získaných analýz (zpětně rozlišený dva valenční rámce):

1. Pro sloveso *dbát₁* je relevantní pouze strom 2, Pg *na ochranu (nájemníků)* je určena jako valenční doplnění tohoto slovesa (Patient).

2. Pro sloveso *dbát₂* jsou vytvořeny stromy 0 a 3, ve kterých je substantivum *nájemníci* analyzováno jako Patient slovesa, Pg *na ochranu* je určena jednou jako kvazivalenční doplnění substantiva *zákon* (strom 0), podruhé jako volné doplnění slovesa (strom 3).

Ve stromu 1 lze identifikovat *dbát₁* i *dbát₂* (zaplněn pouze ‚slot‘ pro aktor, Pg *na ochranu (nájemníků)* určena jako kvazivalenční doplnění substantiva *zákon*). Protože oba valenční rámce slovesa *dbát* jsou považovány za nezávislé, bude tento strom pro *dbát₁* odstraněn (valenční požadavek slovesa blokuje relaxaci slovosledných omezení), pro *dbát₂* ovšem zůstane zachován.

Ve vstupních datech ovšem nejsou striktně rozlišený možné realizace různých rámců. Pokud se ve více rámcích vyskytuje jedno doplnění se stejnou formou, je možnost, že jde o jediné (povrchově realizované) valenční doplnění, popsána jediným ‚slotem‘. Potom je ovšem potřeba pro relaxaci slovosledných omezení brát v úvahu valenční požadavky všech rámců daného slovesa. Zdá se však, že toto řešení neodporuje intuitivnímu postupu. □

Příklad:

(VI.2) *Neschopnost dohodnout se na něm způsobila skluz v realizaci omezené autonomie na území Golanských výšin.* (PDT, blc02zu.fs, pozměněno, viz (V.2))

(VI.2') *Neschopnost dohodnout se na něm způsobila skluz v realizaci autonomie omezené na území Golanských výšin.* (PDT, blc02zu.fs, slovosledná varianta, viz (V.2'))

Slovosledné varianty této věty se liší pouze pozicí adjektiva *omezený*. Ukazují vliv slovosledných omezení a jejich oslabování na určení možných závislostí Pg.

Podívejme se na automatické zpracování druhé části varianty (VI.2), pracujeme s redukovanou Pg *na území*.

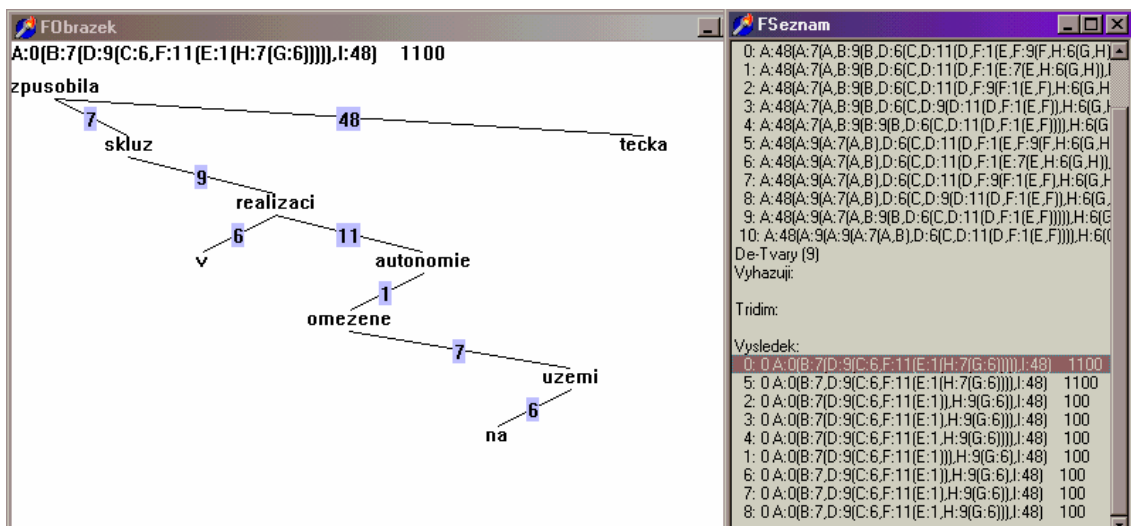
... *způsobila skluz v realizaci omezené autonomie na území.*

The screenshot shows a window titled 'F0brazek' displaying a dependency tree for the phrase 'způsobila skluz v realizaci omezené autonomie na území'. The root node is 'způsobila' (7), which branches into 'skluz' (9) and 'tečka' (48). 'skluz' (9) branches into 'realizaci' (6) and 'uzemi' (11). 'realizaci' (6) branches into 'v' (1) and 'autonomie' (9). 'autonomie' (9) branches into 'omezene' (1) and 'uzemi' (9). 'uzemi' (9) branches into 'na' (6). The window also shows a window titled 'FSeznam' with a list of variants (DR-Tvary and De-Tvary) and their corresponding dependency structures and scores.

DR-Tvary (9)	De-Tvary (7)
0: A:48[A:7(A,B:9[B:D:6(C,D:11(D,F:1(E,F:9(F,H:6(G,I:1(A:48[A:7(A,B:9[B:D:6(C,D:11(D,F:9(F:1(E,F:6(G:2: A:48[A:7(A,B:9[B:D:6(C,D:9(D:11(D,F:1(E,F:6(C:3: A:48[A:7(A,B:9[B:9[B:D:6(C,D:11(D,F:1(E,F:6(4: A:48[A:9(A:7(A,B),D:6(C,D:11(D,F:1(E,F:9(F,H:6(5: A:48[A:9(A:7(A,B),D:6(C,D:11(D,F:9(F:1(E,F:6(6: A:48[A:9(A:7(A,B),D:6(C,D:9(D:11(D,F:1(E,F:6(7: A:48[A:9(A:7(A,B:9[B:D:6(C,D:11(D,F:1(E,F:6(8: A:48[A:9(A:7(A,B),D:6(C,D:11(D,F:1(E,F:6(DR-Tvary (7) Vyhazují: Tridím: Vysledek: 0: 0 A:0[B:7(D:9(C:6,F:11(E:1)),H:9(G:6),I:48) 100 1: 0 A:0[B:7(D:9(C:6,F:11(E:1)),H:9(G:6),I:48) 100 2: 0 A:0[B:7(D:9(C:6,F:11(E:1)),H:9(G:6),I:48) 100 3: 0 A:0[B:7(D:9(C:6,F:11(E:1),H:9(G:6),I:48) 100 4: 0 A:0[B:7(D:9(C:6,F:11(E:1)),H:9(G:6),I:48) 100 5: 0 A:0[B:7(D:9(C:6,F:11(E:1),H:9(G:6),I:48) 100 6: 0 A:0[B:7(D:9(C:6,F:11(E:1),H:9(G:6),I:48) 100

Pokud je zadáno omezení na projektivní analýzu, je vytvořeno sedm závislostních stromů, se čtyřmi různými závislostmi Pg *na území* – *způsobit na území*, *skluz na území*, *realizace na území* a *autonomie na území*. Pg je ve všech případech určena jako volné doplnění.

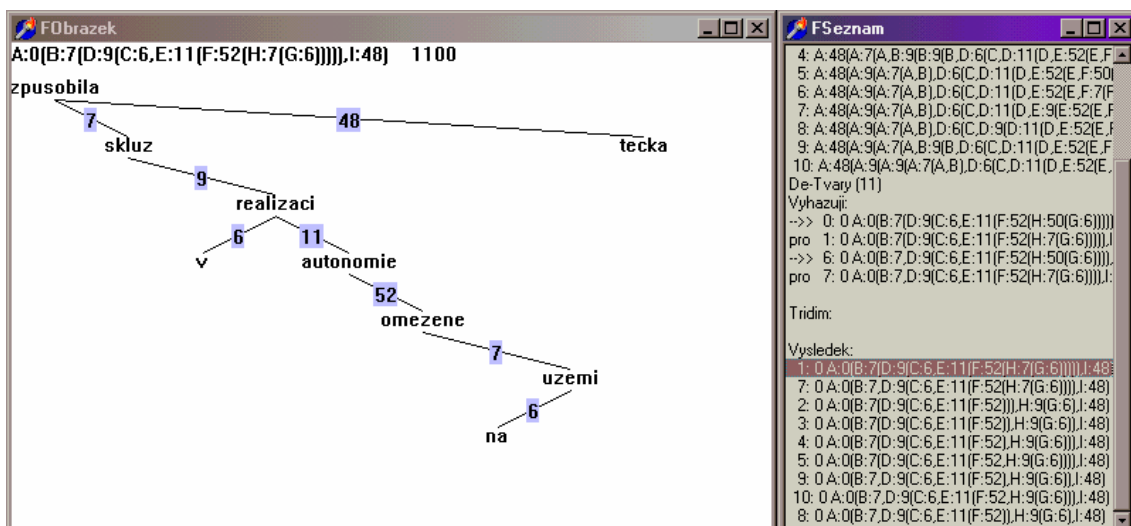
Jsou-li umožněny neprojektivní konstrukce, je navíc k těmto čtyřem závislostem vytvořena závislost *omezený na území*, a to jako valenční doplnění příslušného adjektiva. Výsledkem analýzy je 9 stromů.



Vyhodnocení: Všechny čtyři analýzy Pg jako volného doplnění – *způsobit na území*, *skluz na území*, *realizace na území* a *autonomie na území* – mají stejný stupeň spolehlivosti, závislost *omezený na území* má nižší preferenci, neboť tvoří neprojektivní podstrom.

Nyní se soustředíme na automatické zpracování druhé části varianty (VI.2'), opět pracujeme s redukovanou Pg *na území*.

... *zpusobila skluz v realizaci autonomie omezené na území*.



Zpracování této věty pomocí (upravené) gramatiky robustního parseru s sebou nese problém – původní gramatika neumožňuje rozvíjení substantiva adjektivem zprava. Pro rozbor (V.2') postačuje prosté připsání meta-pravidlo pro zpracování shodného adjektivního přívlastku zprava, které ovšem obecně vede k vytváření velkého množství špatných závislostí (pokazí i analýzu varianty (VI.2)).

Výsledkem projektivní analýzy je jedenáct závislostních stromů s šesti rozborů Pg *na území*: Pg je určena jednak jako valenční doplnění adjektiva *omezený*, jednak jako jeho volné doplnění (automaticky vyloučeno jako duplicitní). Dále jako volné doplnění rozvíjející substantiva *autonomie*, *realizace*, *skluz*, nebo sloveso *způsobit*.

Analýza umožňující neprojektivní konstrukce přináší stejné výsledky

Vyhodnocení:

Nejvyšší preference je přidělena závislosti *omezený na území*, která naplňuje valenční požadavek adjektiva. Ostatní čtyři závislosti – *způsobit na území*, *skluz na území*, *realizace na území* a *autonomie na území* – mají stejnou spolehlivost volných doplnění. □

V následujícím příkladu se soustředíme na zpracování věty se slovesně jmenným výrazem.

Příklad:

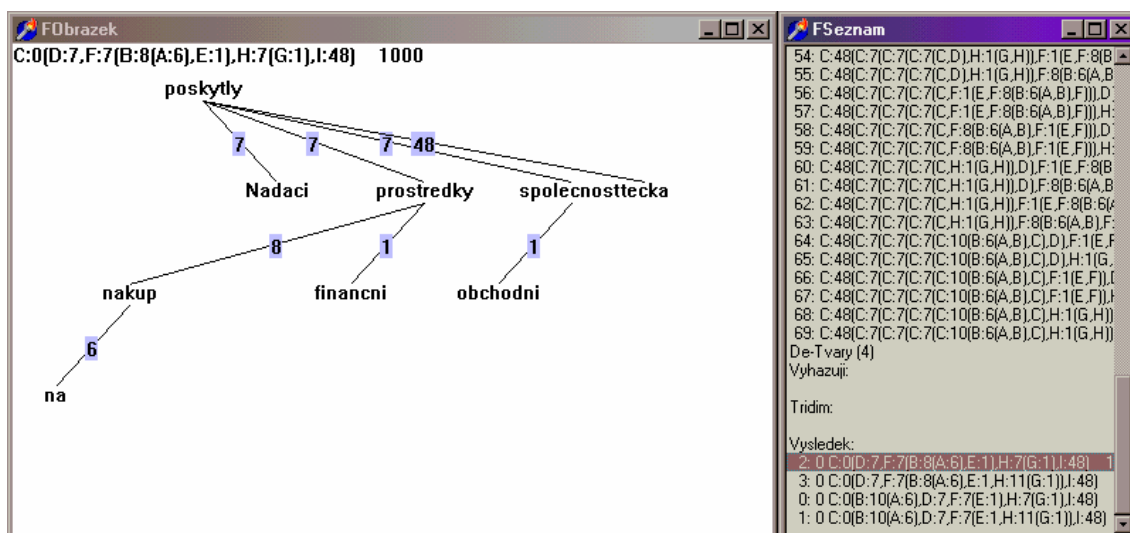
(VI.3) *Na nákup technického vybavení a učebnic poskytly Nadaci finanční prostředky obchodní společnosti.* (PDT, bm122zua.fs #21, pozměněn slovosled, viz (V.3))

Předpokládáme, že je identifikován slovesně jmenný výraz *poskytnout prostředky* (substantivum *prostředek* má valenční doplnění vyjádřené Pg (Pat, na+Acc)).

Je zpracovávána věta s redukovanou Pg *na nákup*.⁸⁰

Na nákup poskytly Nadaci finanční prostředky obchodní společnosti.

Protože slovesně jmenné výrazy patří k jevům, pro které jsou oslabena slovosledná omezení (viz kapitola IV.1.2.), je zajímavá analýza, která umožňuje neprojektivní konstrukce. Při takové analýze jsou vytvořeny čtyři stromy, Pg *na nákup* je zpracována dvěma způsoby – jako volné doplnění slovesa *poskytnout* nebo jako (kvazi)valenční doplnění⁸¹ substantiva *prostředek*.



⁸⁰ Technická omezení interpretu (omezení maximálního počtu derivačních stromů (DR-stromů)) nás vedlo k ručnímu snížení homonymie adjektiva *finanční*, ve valenčním rámci slovesa *poskytnout* nejsou uvedeny všechny možné kombinace doplnění.

⁸¹ Připomeňme, že z technických důvodů zde nejsou rozlišena valenční a kvazivalenční doplnění, pracuje se s nimi shodně (viz kapitola VI.1.1.)

Manuální vyhodnocení:

Protože byl v původní větě identifikován slovesně jmenný výraz, *poskytnout prostředky*, je vyloučena adverbální analýza. Je zachována pouze závislost Pg na substantivu, *prostředky na nákup*, se spolehlivostí valenčního doplnění. □

Příklad:

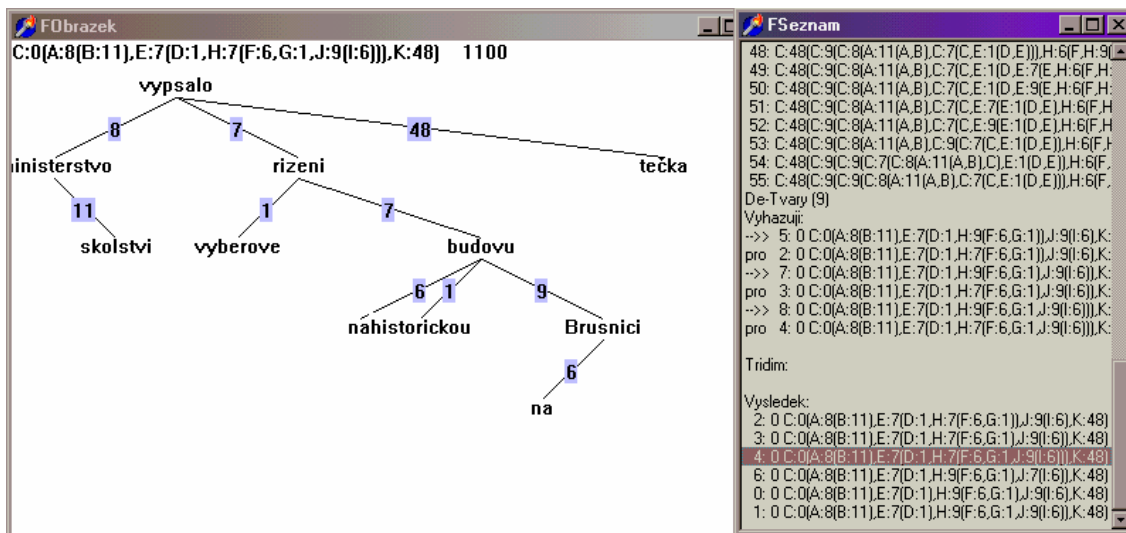
(VI.4) *Ministerstvo školství vypsalo výběrové řízení na historickou budovu na Brusnici.* (PDT, bm102zua.fs #18, zkráceno, viz i (V.4))

Tato věta obsahuje dvě předložkové skupiny, které mohou naplňovat kvazivalenční požadavek substantiva. Ukazovali jsme na ní použití pravidel využívajících sémantické rysy substantiv. Tato pravidla nejsou implementována, lze je však použít pro ruční ohodnocení získaných volných doplnění.

Při projektivní analýze je vytvořeno devět stromů.

- Pg na *(historickou) budovu* je určena jako (kvazi)valenční doplnění substantiva *řízení*, Pg na *Brusnici* je postupně určena jako volné doplnění rozvíjející sloveso *vypsát*, substantivum *řízení*, nebo substantivum *budovu* (stromy 2, 3 a 4);
- Pg na *Brusnici* je určena jako (kvazi)valenční doplnění substantiva *řízení*, Pg na *(historickou) budovu* je určena jako volné doplnění rozvíjející substantivum *řízení* (strom 6);
- obě Pg jsou určeny jako volná doplnění – Pg na *(historickou) budovu* jako neshodný atribut substantiva *řízení*, Pg na *Brusnici* postupně rozvíjí sloveso *vypsát*, substantivum *řízení*, nebo substantivum *budova* (stromy 5, 7 a 8);
- obě Pg jsou určeny jako volná doplnění – Pg na *(historickou) budovu* rozvíjí sloveso *vypsát*, Pg na *Brusnici* rozvíjí sloveso *vypsát* nebo substantivum *budova* (stromy 0 a 1).

Tři analýzy (stromy 5, 7 a 8) jsou automaticky vyloučeny (duplicita valenční / volné doplnění substantiva *řízení* pro Pg na *budovu*).



Vyhodnocení:

Závislostní strom (strom 4), ve kterém je Pg na *budovu* určena jako kvazivalenční doplnění substantiva *řízení*) a Pg na *Brusnici* tvoří neshodný atribut substantiva *budova* (volné doplnění),

má nejvyšší preferenci (výběr této závislosti je dán sémantickými rysy substantiv). To je ve shodě s preferovaným čtením na základě obsahu věty.

Následující analýzy (stromy 2 a 3) se stejně určenou Pg *na budovu* a s Pg *na Brusnici* rozvíjející postupně sloveso *vypsát* nebo substantivum *řízení*, resp. analýza (strom 6) s oběma Pg určenými jako atribut substantiva *řízení*, a to s kvazivalenční Pg *na Brusnici* a volnou Pg *na budovu*.

Analýzy reprezentované stromy 0 a 1 budou označeny jako málo spolehlivé.

Jsou-li umožněny neprojektivní struktury, je navíc vytvořen jediný závislostní strom – Pg *na Brusnici* (*na+Acc*) je určena jako kvazivalenční doplnění substantiva *řízení* a Pg *na budovu* jako volné doplnění slovesa *vypsát* (strom 9). Tato analýza však nesplňuje formální omezení („princip separace“, viz IV.4.1.), a jako taková bude vyloučena. □

Shrnutí

V šesté části jsme ukázali možnost využít kritéria zkoumaná v předchozích kapitolách v rámci konkrétní automatické procedury, za jejíž základ byla použita gramatika Robustního parseru pro češtinu. Byly popsány úpravy této gramatiky umožňující analyzovat předložkové skupiny jako valenční, resp. kvazivalenční doplnění sloves, substantiv a adjektiv. Dále byly navrženy vhodné konfigurace interpretu této gramatiky.

Vyhodnocení získaných závislostí předložkových skupin podle typu pravidel, která vedla k jejich stanovení, případně odstranění neadekvátních analýz (vyloučení duplicit valenční / volné doplnění, kontroly slovosledných omezení a splnění formálních kritérií) je prováděno manuálně pro všechny úplné položky (tj. položky reprezentující syntaktickou strukturu celé vstupní věty).

ZÁVĚR

Tato práce se zabývá možností zachycení homonymie předložkových skupin v rámci automatické syntaktické analýzy, tedy automatickým určením jejich možných řídicích slov ve větě.

Vymezuje **dva zdroje** homonymie předložkových skupin – několikeré ‚čtení‘ Pg může být dáno jednak víceznačnou morfologickou analýzou (**fonologická homonymie** Pg, která je ‚avizována‘ již během morfologického zpracování věty, na jehož výstupu je několik kombinací sémat), jednak obecnými syntaktickými pravidly, která umožňují vytváření několika závislostních struktur. Ukázali jsme, že oba typy homonymie se mohou kombinovat v jediné předložkové skupině. Z hlediska automatické syntaktické analýzy je zajímavý především druhý typ homonymie, reprezentovaný **morfemickou homonymií** (případná fonologická homonymie Pg způsobí rozpad analýzy do větví, v nichž je morfologická charakteristika Pg jednoznačná). Proto jsme se v této práci soustředili na stanovení lingvistických pravidel, která mohou přispět k (automatickému) určování adekvátních větných členů, které jsou v konkrétní větě řídicími členy zkoumané předložkové skupiny, a na možnost vyhodnocení získaných analýz.

Významným zdrojem dat je pro nás **Pražský závislostní korpus** – anotace na analytické rovině (popisující syntaktickou strukturu povrchových realizací vět) umožnila získat množství výskytů předložkových skupin nejen v prototypických pozicích, ale i v postaveních netypických, které nelze jednoduše detekovat pouze na základě slovosledu. (Sekundárním výsledkem naší práce je ilustrace možností, které poskytuje PDT jako zdroj dat při vyhledávání poměrně složitých syntaktických struktur.)

Formulovali jsme obecné principy pro postup při zpracování přirozeného jazyka na základě porozumění větě, **metodu redukční analýzy**, spočívající v postupném zjednodušování analyzované věty vypouštěním (redukcí) jednotlivých větných členů. Tuto metodou jsme využili při zkoumání různých typů lingvistických kritérií, která se dají uplatnit při určování morfemické homonymie předložkových skupin. Zaměřili jsme se přitom na čtyři oblasti – na slovosledná kritéria, na přínos valence pro disambiguaci Pg, na možnou slovoslednou pozici zkoumané Pg vzhledem k ostatním doplněním konkrétní lexikální jednotky a na pravidla využívající sémantické rysy předložek a substantiv tvořících Pg. Shrňme zde v krátkosti výsledky těchto zkoumání.

Základní **strategie**, se kterou jsme přistupovali ke zkoumání možných závislostí předložkových skupin, spočívá v postupném zjemňování pravidel pro jejich analýzu. Nejprve byla stanovena základní omezení popisující prototypické postavení Pg. Tato omezení byla ve druhé fázi výzkumu pro některé jazykové jevy oslabena. V poslední fázi byly vyloučeny nepřijatelné tvary závislostních stromů.

Ukazuje se, že pro homonymii předložkových skupin jsou typické určité pozice Pg v povrchové realizaci věty – tyto pozice byly popsány pomocí **slovosledných vzorců** (word order patterns, WOP). Taková slovosledná omezení, charakteristická zejména pro volná doplnění, jsou pro některé jazykové konstrukce rozvolněna (v zájmu získání analýzy, která by lépe odpovídala jazykové intuici). Navrhujeme uvolnit slovosledná omezení pro slovesné jmenné výrazy, pro valenční a kvazivalenční doplnění substantiv (za dodatečného předpokladu,

že zkoumaná Pg nenaplnuje valenční požadavky slovesa) a pro neprojektivní rozvíjení adjektiva.

Pro určení adekvátního řídicího členu předložkové skupiny hrají klíčovou roli **valenční požadavky** jednotlivých lexikálních jednotek – uvažujeme valenční rámce sloves, substantiv a adjektiv. Koncept valenčních rámců interpretujeme pro potřeby analýzy povrchového zápisu vět (tato interpretace se týká především vypustitelnosti jednotlivých členů rámce) a rozšiřujeme ho o tzv. kvazivalenční doplnění, tedy taková doplnění, která nespĺňují podmínky kladené na valenční doplnění, nicméně popisují jejich přenesené nebo idiomatičké užití, příp. zachycují ‚obvyklé užití‘ konkrétních lexikálních jednotek. Při zkoumání slovesně jmenných výrazů je naznačena možnost přechodu od valenčních rámců jednotlivých slov ke zkoumání požadavků víceslovných spojení. Zvláštní pozornost je věnována křížení valenčních požadavků, je stanovena prioritá doplnění jednotlivých slovních druhů.

Testujeme možnost pracovat s pravidly využívajícími ‚**primárních funkcí**‘ jednotlivých předložek a **sémantických rysů** substantiv. Jejich využití je ovšem problematické (ačkoliv v jednotlivých případech můžou být jejich výsledky povzbuzující), získané ohodnocení jednotlivých závislostí Pg není zcela spolehlivé. Zdá se, že sémantická kritéria lze využít především pro určení jednotlivých typů volných doplnění (valenčních, tedy obligatorních, i fakultativních).

Navrhovaná uvolnění slovosledných omezení pro některé jazykové jevy mohou vést k vytváření neprojektivních konstrukcí. Ve snaze zamezit nepatřičnému uvolňování podmínky projektivnosti zkoumáme další principy, které mohou přispět k vyloučení neadekvátních analýz valenčních i volných doplnění (‚**princip separace**‘). Také pravidlo o **postavení příklonek** lze využít pro stanovení hranic jednotlivých syntaktických členů.

Vyšetřovaná pravidla pro zpracování předložkových skupin mají různou ‚váhu‘ – na jejich základě je možné stanovit preference pro jednotlivé potenciální závislosti předložkových skupin, umožňují **evaluovat** získané závislostní vztahy. Tento rys přístupu k analýze, založený na zkoumání lingvistických kritérií, je pro automatické zpracování přirozeného jazyka velmi podstatný. Umožňuje dostatečně spolehlivě zpracovávat věty obsahující jevy, které nejsou plně popsány, ať už ve slovníku nebo v gramatice – to je důležité zejména vzhledem ke skutečnosti, že žádný jazykový model použitelný pro automatickou analýzu nemůže obsahovat všechny informace (včetně sémantických a pragmatických znalostí), na jejichž základě se rozhoduje člověk. Navíc odráží skutečnost, že mnohé věty jsou v obvyklých kontextech jednoznačné, přitom však mohou existovat kontexty, ve kterých je nutno preferovat jejich jiné ‚čtení‘.

Ukázali jsme, že – pokud překročíme hranice relativně jednoduché analýzy založené převážně na syntaktických pravidlech – každé zpřesňování gramatiky vyžaduje velmi podrobná lingvistická pozorování, která daleko přesahují rámec čisté syntaxe.

Vytvořili jsme **model automatické syntaktické analýzy předložkových skupin** založený na principech obecné metody redukční analýzy. Redukční analýza je realizována jako posloupnost kroků, ve kterých jsou aplikována jednotlivá kritéria pro stanovení adekvátních řídicích členů předložkových skupin a pro případné určení jejich homonymie.

Možnost využití zkoumaných lingvistických kritérií byla **testována v prostředí konkrétní automatické procedury**, jejímž základem se stala gramatika a interpret Robustního parseru pro češtinu. Ukázali jsme, jakým způsobem lze upravit tuto gramatiku tak, aby zachycovala navržená kritéria. Výsledné syntaktické struktury testovacích vět (reprezentované závislostními stromy) byly částečně automaticky vyhodnoceny na základě přijatého uspořádání.

Přestože je tato práce tematičce omezená na zkoumání homonymie předložkových skupin a na možnosti automatického určování jejich potenciálních řídicích členů, lze ji

považovat za metodologický základ pro konkrétní parser využívající jemné lingvistické zákonitosti, který by pokrýval (alespoň) jádro češtiny a adekvátním způsobem zachycoval homonymii různých větných členů.

Široká škála popisovaných jevů a pozorování založená na shromážděném lingvistickém materiálu dává možnost dalšího využití – jmenujme například budování bohatého slovníku češtiny, zkoumání valenčních rámců jednotlivých lexikálních jednotek i víceslovných spojení, podněty k pravidly řízené morfologické disambiguaci atd.

V širším pohledu lze tuto práci chápat jako ‚odrazový můstek‘ pro další výzkum lingvistických kritérií využitelných v rámci různých aplikovaných úkolů, jejichž základem je automatická syntaktická analýza.

LITERATURA

- J. Allen (1987) *Natural Language Understanding*. The Benjamin/Cummings Publishing Company.
- T. Avgustinova, K. Oliva (1995) *Wackernagel Position and related phenomena in Czech*. In: *Wiener Slavistisches Jahrbuch* (eds. G. Holzer, G. Huttl-Folter, R. Katicic, F. V. Mares, J. Vintr, P. Wincher), volume 41, Verlag des Oesterreichischen Akademie der Wissenschaften, Wien.
- A. Bémová (1979) *Syntaktické vlastnosti prefigovaných sloves*. In: *Explicite Beschreibungder Sprache und automatische Textbearbeitung V.*, pp. 69-194.
- A. Bémová et al. (1997) *Anotace na analytické rovině: Návod pro anotátory*. ÚFAL Technical Report TR-1997-03.
- E. Brill, P. Resnik (1994) *A Rule-Based Approach to Prepositional Phrase Attachment Disambiguation*. In: *Proceedings of COLING-94, Kyoto, Japan*.
- E. Buráňová (1980) *Ob odnoji možnosti semantičeskoi klassifikacii suščestvitel'nyx*. *PBML* 34, Prague, pp. 33-43.
- Fr. Daneš (1964) *Opyt teoretičeskoi interpretacii grammatičeskoi omonimii*. *Voprosy jazykoznanija* 6, pp. 3-16.
- Fr. Daneš, Z. Hlavsa (1981) *Větné vzorce v češtině*. Academia, Praha.
- A. Dikovsky (1999) *On dependency syntax: abstract vs. concrete*. (nepublikovaný rukopis)
- A. Dikovsky, L. Modina (2000) *Dependencies on the other side of the Curtain*. *Traitement Automatique des Langues (TAL, ed. S. Kahane)*, vol. 41, No. 1, Hermes, pp. 79-111.
- Ch. J. Fillmore (1968) *The Case for Case*. In: *Universals in Linguistic Theory* (eds. E. Bach, R. Harms), New York, 1-90.
- Ch. J. Fillmore (1977) *The Case for Case Reopened*. In: *Syntax and Semantics 8* (eds. P. Cole, J.M. Sadock), New York-San Francisco-London, pp. 59-81.
- J. Hajič (1998) *Building a Syntactically Annotated Corpus: The Prague Dependency Treebank*. In: *Issues of Valency and Meaning. Studies in Honour of Jarmila Panevová* (ed. E. Hajičová), Karolinum, Charles University Press, Prague, pp. 106-132.
- J. Hajič, E. Hajičová, J. Panevová, P. Sgall (1999) *Specifikace stromových struktur na tektogramatické úrovni značkování a základní kroky překladu z analytické na tektogramatickou úroveň*. Technical report of the project GAČR 405-96-0198.
- J. Hajič, K. Ribarov (1997) *Rule-Based Dependencies*. In: *Workshop Notes of the ECML/MLnet Workshop on Empirical Learning of Natural Language Processing Tasks* (eds. W. Daelemans, A. van den Bosch, A. Weijters), Prague, pp. 125-136.
- E. Hajičová (1979) *Agentive or Actor/Bearer*. *Theoretical Linguistics* 6, pp. 173-190.
- E. Hajičová, J. Panevová (1984) *Valency (Case) Frames of Verbs*. In: *Contributions to Functional Syntax, Semantics, and Language Comprehension* (ed. P. Sgall), Prague, Academia and Amsterdam, John Benjamins, pp. 147-188.
- E. Hajičová, J. Panevová, P. Sgall (ve spolupráci s A. Böhmovou, m. Ceplovou a V. Řezníčkovou) (2000) *A Manual for Tectogrammatical Tagging of the Prague Dependency Treebank*. ÚFAL/CKL Technical Report TR-2000-09.
- K. Hausenblas (1958) *Syntaktická závislost, způsoby a prostředky jejího vyjádření*. *Bulletin Vysoké školy ruského jazyka*, 2, pp. 23-51.

- D. Hindle, M. Rooth (1993) Structural ambiguity and Lexical Relations. *Computational Linguistics*, Vol. 19, No.1, pp. 103-120.
- G.E. Heidorn (1975) Augmented Phrase Structure Grammars. In: *Theoretical Issues in Natural Language Processing* (eds. Nash-Weber and Schank), Association for Computational Linguistics.
- T. Holan (2001) Nástroje pro vývoj závislostních analyzátorů přirozených jazyků s volným slovosledem. *Disertační práce*, MFF UK, Praha. (in prep.)
- T. Holan, V. Kuboň, K. Oliva, M. Plátek (1998) Two Useful Measures of Word Order Complexity. *Proceedings of Dependency-Based Grammars. Proceedings of the Workshop* (eds. A. Polguere, S. Kahane), Conference COLING – ACL '98, Montreal, Canada, pp. 21-28.
- T. Holan, V. Kuboň, K. Oliva, M. Plátek (2000) On Complexity of Word Order. In: *Les grammaires de dépendance - Traitement automatique des langues (TAL)*, Vol. 41, No. 1 (special issue on dependency grammars of the journal *Traitement automatique des langues*, guest editor Sylvain Kahane), pp. 273-300.
- P. Jančar, F. Mráz, M. Plátek, J. Vogel (1999) On Monotonic Automata with a Restart Operation. *Journal of Automata, Languages and Combinatorics*, Vol. 4, No. 4, pp. 287-311.
- K. Jensen, J.-L. Binot (1987) Disambiguating Prepositional Phrase Attachment by Using On-Line Dictionary Definitions. *Computational Linguistics*, Vol. 13, No. 3-4, pp. 251-260.
- Z. Kirschner (1994) CZECKER – a Maquette Grammar-Checker for Czech. *The Prague Bulletin of Mathematical Linguistics* 62, pp.5-30.
- V. Kuboň (1999) A Robust Parser for Czech. *ÚFAL Technical Report TR-1999-06*.
- V. Kuboň, T. Holan and M. Plátek (1997) A Grammar-Checker for Czech. *ÚFAL Technical Report TR-1997-02*.
- J. Kuryłowicz (1936) Dérivation lexicale et dérivation syntaxique. *Bulletin de la Société linguistique de Paris* 37, pp. 79-92.
- S. Marcus (1965) Sur la notion de projectivité. In: *Zeitschr. f. math. Logik und Grundlagen d. Math.* XI, pp. 181-192.
- V. Mathesius (1942) Řeč a sloh. In: *Čtení o jazyce a poesii*. (Eds. B. Havránek, J. Mukařovský), *Družstevní práce*, Praha. (2. vydání)
- I. Mel'čuk (1988) *Dependency Syntax: Theory and Practice*. NY:SUNY Press, Albany.
- L. Nebeský (1972) A Projectivity Theorem. *Prague Studies in Mathematical Linguistics* 3, (Academia, Prague), pp.165-169.
- J. Novotný (1980) Valence dějových substantiv v češtině. *Sborník Pedagogické fakulty v Ústí n. Labem*, Praha.
- M. Novotný (1988) S algebrou od jazyka ke gramatice a zpět. *Academia*, Praha.
- K. Oliva (1989) A Parser for Czech Implemented in Systems Q. In: *Explizite Beschreibung der Sprache und automatische Textbearbeitung*, MFF UK, Prague.
- K. Oliva (1996) A Grammar Checker for Czech. JRP PECO 2824 *Language Technologies for Slavic Languages*, part of Final Research Report, Prague.
- J. Panevová (1966) Nesoglasovannoe opredelenie s točki zrenija analiza dlja mašinnogo perevoda. *Prague Studies in Mathematical Linguistics* 1, *Academia*, Prague. pp. 219-239.
- J. Panevová (1974, 1975) On Verbal Frames in Functional Generative Description. Part I, *The Prague Bulletin of Mathematical Linguistics* 22, pp.3-40, Part II, *The Prague Bulletin of Mathematical Linguistics* 23, pp. 17-52.

- J. Panevová (1980) *Formy a funkce ve stavbě české věty*. Academia, Praha.
- J. Panevová (1981) *K otázkám homonymie a neutralizace ve stavbě věty*. *Jazykovedné štúdie* 16, pp. 85-89.
- J. Panevová (1992) *O nekotoryx tipax obobschennyx aktantov*. In: *Festschrift fur V. J. Rozencvejg, Wiener Slawistischer Almanach* (ed. T. Reuther), Sonderband 33, Wien, pp. 73-78.
- J. Panevová (1994) *Valency Frames and the Meaning of the Sentence*. In: *The Prague School of Structural and Functional Linguistics* (ed. Ph. L. Luelsdorff), Amsterdam-Philadelphia, John Benjamins, pp. 223-243.
- J. Panevová (1998a) *Funkční styly a automatické zpracování jazyka*. In: *Česká slavistika. České přednášky pro XII. mezinárodní sjezd slavistů, Krakov 1998* (ed. H. Bláhová, S. Wollmann a kol.), Slavia, Slovanský ústav AV ČR, pp. 161-167.
- J. Panevová (1998b) *Ještě k teorii valence*. *Slovo a slovesnost* 59, pp. 1-14.
- J. Panevová (1999) *Česká reciproční zájmena a slovesná valence*. *Slovo a slovesnost* 60, pp. 269-275.
- J. Panevová (2000) *Poznámky k valenci podstatných jmen*. *Čeština - univerzália a specifika 2*, Sborník z konference ve Šlapanicích u Brna, Masarykova Univerzita, Brno, pp. 173-180.
- J. Panevová (2001) *Valency Frames: Extension and Re-examination*. (in press)
- J. Panevová, P. Sgall (1979) *Towards an automatic parser for Czech*. In: *International Review of Slavic Linguistics*, pp. 433-445.
- J. Panevová, P. Sgall (1980) *On Some Issues of Syntactic Analysis of Czech*. *The Prague Bulletin of Mathematical Linguistics* 34, pp. 21-32.
- J. Panevová, M. Straňáková (1999) *Some Types of Syntactic Ambiguity; How to Treat them in an Automatic Procedure*. In: *TSD'99, Proceedings* (eds. V. Matoušek, P. Mautner, J. Ocelíková, P. Sojka), *Lecture Notes in Artificial Intelligence* vol.1692, Springer, pp. 50-55.
- J. Panevová, M. Straňáková (2000) *Selected Types of Ambiguity of Prepositional Groups; Classification, Criteria and Method for Automatic Processing (Abstract)*. In: *Proceedings of the SLE Conference, Pognan, 31.8.-2.9. 2000*, pp. 61-63.
- J. Panevová, V. Řezníčková (2001) *K možnému pojetí všeobecnosti aktantu*. In: *Čeština - univerzália a specifika 3*. Sborník konference ve Šlapanicích u Brna, 22.-24.11.2000 (ed. Zdeňka Hladká, Petr Karlík), pp. 139-146.
- P. Pantel, Dekang Lee (2000) *An Unsupervised Approach to Prepositional Phrase Attachment using Contextually Similar Words*. In: *ACL 2000 – Proceedings*, pp. 101-108.
- P. Piřha (1981) *On the Case Frames of Nouns*. *Prague Studies in Mathematical Linguistics* 7, Academia, Prague, pp. 215-224.
- P. Piřha (1982) *K otázce valence u adjektiv*. *Slovo a slovesnost* 43, pp.113-118.
- M. Plátek, T. Holan, V. Kuboň, J. Hric (1996) *Grammar Development & Pivot Implementation*. JRP PECO 2824 *Language Technologies for Slavic Languages*, part of Final Research Report, Prague.
- M. Plátek, V. Kuboň, T. Holan (1997) *Formal Tools for Separating Syntactically Correct and Incorrect Structures*. In: *Proceedings of the IWPT'97, MIT, Boston*, pp. 247-248.
- H. Prouzová (1983) *K valenčním vlastnostem primárních adjektiv v češtině*. *Slovo a slovesnost*, 44, pp. 265-274.
- A. Ratnaparkhi (1998) *Statistical Models for Unsupervised Prepositional Phrase Attachment*. In: *COLING/ ACL'1998. Proceedings*, pp.1079-1085.

- K. Ribarov (1996) Automatická tvorba gramatiky přirozeného jazyka. Diplomová práce, ÚFAL MFF UK, Praha.
- J. J. Robinson (1969) Case, Category and Configuration. *Journal of Linguistics* 6, pp. 57-80.
- P. Sgall (1960) Soustava pádových koncovek v češtině. *AUC-Slavica Pragensia*, pp.65-84.
- P. Sgall (1967) Generativní popis jazyka a česká deklinace. *Academia*, Praha.
- P. Sgall (1992) Underlying Structure of Sentences and its Relations to Semantics. In: *Festschrift für Viktor Jul'evič Rozencvejk zum 80. Geburtstag*, Wiener Slawistischer Almanach 33 (ed. T. Reuther), Vienna, pp. 273-282.
- P. Sgall (1998) Revisiting the classification of the dependents; Interesting results and tempting topics for further research. In: *Issues of Valency and Meaning. Studies in Honour of Jarmila Panevová* (ed. E. Hajičová), Karolinum, Charles University Press, Prague, pp. 15-26.
- P. Sgall et al (1986a) Úvod do syntaxe a sémantiky. Některé nové směry v teoretické lingvistice. *Academia*, Praha.
- P. Sgall, E. Hajičová, J. Panevová (1986b) The Meaning of the Sentence in Its Semantic and Pragmatic Aspects (ed. by J. Mey), Dordrecht:Reidel and Prague:Academia.
- S. M. Shieber (1986) An Introduction to Unification-Based Approaches to Grammar. *CSLI Lecture Notes* 4, Center for the Study of Languages and Information, Stanford University, Stanford.
- K. Sikkel, A. Nijholt (1997) Parsing of Context-Free Languages. In: *Handbook of Formal Languages* (eds. G. Rosenberg, A. Salomaa), Springer, Berlin and Heidelberg, pp. 61-100.
- H. Skoumalová (1994) Czech Dictionary for the Grammar Checker. *JRP PECO 2824 Language Technologies for Slavic Languages*, Research Report. Saarbruecken, pp. 130-139.
- H. Skoumalová (1997) Verb Frames in the Czech Hierarchical Lexicon. *TELRI Newsletter* 6, pp. 18-32.
- H. Skoumalová (1998) Derived frames and the lexicon. In: *Issues of Valency and Meaning. Studies in Honour of Jarmila Panevová* (ed. E. Hajičová), Karolinum, Charles University Press, Prague, pp.154-168.
- H. Skoumalová (2001) Czech syntactic lexicon. PhD thesis, Charles University, Faculty of Arts, Prague. (submitted)
- J. Stetina, M. Nagao (1997) Corpus Based PP Attachment Ambiguity Resolution with a Semantic Dictionary. In: *Proceedings of the Fifth Workshop on Very Large Corpora*, Beijing and Hong Kong, pp. 66-80.
- M. Straňáková (1998) Ambiguity in Czech Sentences, its Classification and Searching for it. In: *Proceedings of WDS'98* (ed. J. Šafránková), Matfyz-press, pp. 165-170.
- M. Straňáková (1999) Selected Types of Pg-Ambiguity. *The Prague Bulletin of Mathematical Linguistics* 72, pp. 29-58.
- M. Straňáková (2000) Selected Types of Pg-Ambiguity; Processing Based on Analysis by Reduction. In: *TSD2000, Proceedings* (eds. P. Sojka, I. Kopeček, K. Pala), *Lecture Notes in Artificial Intelligence* vol.1902, Springer, pp. 139-144.
- M. Straňáková-Lopatková (2001) Ambiguity of Prepositional Groups; Classification, Criteria and Method for Automatic Processing. In: *sborník university v Oldenburgu*. (in press)
- M. Straňáková-Lopatková (2001) Některé typy syntaktické homonymie (z hlediska možnosti automatického zpracování). In: *Čeština - univerzália a specifika 3. Sborník konference ve Šlapanicích u Brna, 22.-24.11.2000* (ed. Z. Hladká, P. Karlík), pp. 183-196.
- Vl. Šmilauer (1966) *Novočeská skladba*. SPN, Praha (2. vydání).

- Fr. Štícha (1987) Komunikativní a jazykové funkce lexikálního nevyjadřování objektu děje ve větě. *Naše řeč*, 70, pp. 184-193.
- V. Řezníčková (2000) Valency Frames of Nouns and Adjectives from PDT Point of View. In: *Proceedings of WDS 2000* (ed. J. Šafránková), Matfyz-press, pp. 466-471.
- L. Tesnière (1959) *Éléments de syntaxe structurale*. Paříž.
- M. Těšitelová (1966) *O morfologické homonymii v češtině*. Praha.
- M. Těšitelová a kol. (1974) *Otázky lexikální statistiky*. Academia, Praha.
- B. Trnka (1931) O homonymii, její terapii a profylaxi. *Časopis pro moderní filologii*, 17, pp. 141-147.
- B. Trnka (1990) *Kapitoly z funkční jazykovědy / Studies in functional linguistics* (ed. Jiří Nosek), Univerzita Karlova, Praha.
- L. Uhlířová (1972) On the Non-Projective Constructions in Czech. *Prague Studies in Mathematical Linguistics* 3, (Academia, Prague), pp.171-181.
- L. Uhlířová (1987) *Knížka o slovosledu*. Academia, Praha.
- D. Zeman (1997) *Pravděpodobnostní model významových zápisů vět*. Diplomová práce, ÚFAL MFF UK, Praha.
- D. Zeman (1998) A Statistical Approach to parsing of Czech. *PBML* 69, pp.29-38.

PŘÍLOHA
POUŽITÉ VALENČNÍ RÁMCE

Valenční a kvazivalenční rámce sloves

sloveso	valenční / kvazivalenční rámec	číslo	příklady, poznámky
<i>bydlet</i>	Act ^Local		
<i>čekat</i>	Act Pat (Acc/na+Acc/klauze) -že, zda, Pron, Adv	IV.14b	č. <i>se zahájením.Reg, od někoho.Dir</i>
<i>číst</i>	Act (Pat (o+Loc) (Addr(Dat)) Eff (Acc/klauze) -že,Pron,Adv	IV.16	? Addr – fakultativní X oblig., všeob.
<i>dát₁</i>	Act Pat (Acc) Addr (Dat)	IV.17	
<i>dbát₁</i>	Act Pat (na+Acc/o+Acc)	IV.31a,V.1	<i>dbát na ochranu</i>
<i>dbát₂</i>	Act Pat (Gen)	IV.31a,V.1	<i>dbát zákonů</i>
<i>hřešit₁</i>	Act	IV.20a	<i>h. skutkem.Means, h. proti příkázání.Ben</i>
<i>hřešit₂</i>	Act ^Pat (na+Acc)	IV.20a	<i>h. na jeho dobrotu</i>
<i>hřešit₃</i>	Act ^Pat (na+Loc)	IV.20a	<i>h. na dětech</i>
<i>informovat</i>	Act Pat (o+Loc/klauze) Addr (Acc) -že, zda, Pron, Adv	IV.18	
<i>jednat₁</i>	Act ^R (Pat (o+Loc)) Addr ^R (s+Ins)	IV.22c	
<i>jednat₂</i>	Act	IV.22c	
<i>jednat₃</i>	Act Pat (s+Ins) Mann	IV.22c	(=zacházet)
<i>jít</i>	Act ^Direction (kam, odkud)		
<i>klást</i>	Act Pat(Acc) Direction	IV.27,29b	<i>klást nároky na ...</i>
<i>léčit₁</i>	Act Pat (Acc)	IV.23c	prim. <i>lék lékař léčí Pavlínu bylinkami.Means na chřipku.Regard</i>
<i>léčit₂</i>	Act Pat (Acc)	IV.23c	prim. <i>lék lékař léčí angínu bylinkami.Means na chřipku.Regard</i>
<i>mít</i>	Act Pat (Acc) (Orig (od+Gen/z+Gen))	IV.26a	
<i>mít</i>	Act Pat (Acc/klauze) ^Locative (na myslí) -že, aby, Pron, Adv	IV.21c	(fraz.)
<i>mít na myslí</i>	Act ^Pat (Acc/klauze) -že, aby, Pron, Adv	IV.21c	(fraz.)

sloveso	valenční / kvazivalenční rámec	číslo	příklady, poznámky
<i>nakládat</i> ₁	Act Pat (Acc) Direction (kam)	IV.23b	<i>nakládat</i> ₁ písek na auto / okurky do octa
<i>nakládat</i> ₂	Act Pat (Acc)	IV.23b	<i>nakládat</i> ₂ auto pískem/lopatou.Means
<i>nakládat</i> ₃	Act Pat (s+Ins) Mann	IV.23b	(=zacházet)
<i>navrhnout</i> ₁	Act Pat (Acc/inf/klauze) (Addr (Dat)) -že, aby, at', Pron, Adv	IV.14c, 23a	
<i>navrhnout</i> ₂	Act Pat (Acc) (Addr (Dat)) Aim (na+Acc)	IV.14c, 23a	<i>kritik navrhuje (porotě) film na ocenění</i>
<i>nominovat</i>	Act Pat (na+Acc) Addr (Acc)	IV.22b	
<i>odejít</i>	Act Direction (odkud)	IV.40	
<i>odpovídat</i>	Act (Pat (na+Acc)) (Addr (Dat)) Eff (klauze) -že, Pron, Adv / General	IV.22a	? Addr – fakultativní X oblig., všeob. Eff – často všeob. <i>odpovídat odpověď</i>
<i>ohodnotit</i>	Act (Nom) Pat (Acc)	I.3	
<i>omezit</i>	Act Pat (Acc) (Eff (na+Acc))	IV.6b	<i>omezit žvýkačkami.Means, v kouření.Regard</i>
<i>pobývat</i>	Act Loc	IV.24a	
<i>pomoci</i> ₁	Act (Pat (s+Ins/inf/klauze) Addr (Dat))	IV.20c	<i>p. radou.Means, ke zdraví/od nemoci.Aim</i>
<i>pomoci</i> ₂	Act ^Pat (na+Acc)	IV.20c,32	<i>p. někomu.Ben na problémy</i>
<i>poskytnout</i>	Act Pat (Acc) Addr (Dat)	IV.26b,29a,V.3	
<i>postihnout</i> ₁	Act Pat (Acc) ^Regard (na+Acc)	IV.30a	<i>postihnout na zdraví</i>
<i>postihnout</i> ₂	Act Pat (Acc)	IV.30a	<i>postihnout (pouhým) okem.Means</i>
<i>potřebovat</i>	Act Pat (Acc) ^Aim (na+Acc)	IV.21b	
<i>prohrávat</i>	Act Pat (Acc) (Addr (s+Ins/proti+Dat))	II.11b	patient X doprovod
<i>předělat</i>	Act Pat (Acc) (Addr (Dat)) Orig (z+Gen) Eff (na+Acc)	IV.15	
<i>překvapovat</i>	Act Pat (Ins) Addr (Acc)	II.17i	
<i>přeměnit</i>	Act Pat (Acc) Orig (z+Gen) Eff (na+Acc/v+Acc)	IV.14d	
<i>přeměnit se</i>	Act Pat (na+Acc/v+Acc) Orig (z+Gen)	IV.14d	refl. tantum
<i>převést</i>	Act Pat Direction (kam)	IV.31b	
<i>přijet</i>	Act Direction (kam)	IV.41	
<i>přimět</i>	Act Pat (k+Dat/inf/klauze) Addr (Acc) - aby	IV.40	
<i>reagovat</i>	Act Pat (na+Acc)	IV.14a	<i>r. smíchem.Means, r. živě.Mann</i>
<i>rozšířit</i>	Act Pat (Acc) Eff (na+Acc)	IV.31c	

sloveso	valenční / kvazivalenční rámec	číslo	příklady, poznámky
<i>sejít se</i>	Act ^R Pat ^R (s+Ins)	II.11c	patient X doprovod
<i>směřovat</i>	Act Direction (kam)	IV.14e	
<i>tlačit₁</i>	Act Pat (Acc)	IV.20b	<i>t. někam.Dir</i>
<i>tlačit₂</i>	Act ^Pat (na+Acc)	IV.20b	<i>na tužku i činit nátlak na někoho, čím.Means</i> (idiom <i>dítě sedí na nočníku a tlačí</i>)
<i>upozornit</i>	Act Pat (na+Acc) (Addr (Acc))	II.16d	
<i>uvalit</i>	Act Pat (Acc) Direction (kam)		<i>uvalit embargo na dovoz</i>
<i>válčit</i>	Act Pat (s+Ins / proti+Dat)	II.11a	patient X doprovod
<i>vynakládat</i>	Act Pat (Acc) Eff (na+Acc)	IV.23b	
<i>vyslýchat</i>	Act Pat (Acc)	II.10	
<i>zajímat se</i>	Act Pat (o+Loc)		tantum
<i>získat</i>	Act Pat (Acc) (Orig (od+Gen))	IV.19	<i>z. pro postižené.Ben,</i> <i>z. na brigádu/za přítele.Aim</i> <i>z. za něco.Subst, z. darem.Means</i>
<i>způsobit</i>	Act Pat (Acc/klauze) (Addr (Dat)) - aby, že		
<i>zvyhodnit</i>	Act Pat (Acc) ^Regard (v+Loc/na+Loc)	IV.21a,30b	<i>z. na daních</i>
<i>žít₁</i>	Act (Pat (Acc))	IV.28a	<i>žít život</i>
<i>žít₂</i>	Act Locative	IV.28a	

Valenční a kvazivalenční doplnění substantiv

subst.	valenční doplnění	číslo	odkazy	poznámky, příklady
<i>akcie</i>	^Aim	IV.25a, 31b		<i>vydat/ akcie na majetek</i>
<i>bratr</i>	App			obligatorní (relační jméno)
<i>certifikát</i>	Pat (na+Acc) ¹			<i>c. na potraviny</i>
<i>článek</i>	Pat (o+Loc) Orig (od+Gen) ¹	IV.10c,13	<i>psát</i>	artefakt <i>článek na téma/k tématu</i> Pat (fraz.)
<i>daň</i>	^Pat (na+Acc/za+Acc/z+Gen)	IV.31c		
<i>dědeček</i>	App (Gen _{subj} / Adj _{pos}) oblig.	II.16d		
<i>dopis</i>	Pat (o+Loc) Orig (od+Gen) ¹	IV.10b,11	<i>psát</i>	artefakt
<i>důraz</i>	Pat (na+Acc)			<i>klást důraz na ...</i>
<i>embargo</i>	Pat (na+Acc)			<i>uvalit embargo na ...</i>
<i>hektar</i>	Mat (Gen)			obligatorní
<i>jednání</i>	Act Pat Addr	IV.22c	<i>jednat₁</i>	<i>jednání něčí_{pos-subj}/někoho_{Gen-subj} (s někým_{Ins}) (o něčem_{Loc})</i> <i>jednání někoho_{Gen-subj}^R (o něčem_{Loc})</i> <i>jednání (někoho_{Gen-subj} a někoho_{Gen-subj}) (o něčem_{Loc})</i> <i>jednání (mezi někým_{Ins} a někým_{Ins}) (o něčem_{Loc})</i>
<i>klíč</i>	Orig (od+Gen)	IV.8a		<i>klíč od zahrady</i> .Orig
<i>láska</i>	Pat (k+Dat) ¹		<i>milovat</i>	supletivum
<i>léčení₁</i>	Act Pat (Gen)	IV.23c	<i>léčit₁</i>	<i>léčení₁ Pavlína / Pavlínino léčení₁</i>
<i>léčení₂</i>	Act Pat (Gen)	IV.23c	<i>léčit₂</i>	<i>léčení₂ anginy</i>
<i>lék</i>	Pat (na+Acc/proti+Dat)	IV.23c,32	primární	<i>lékařů_{subj}/Pavlínin_{obj}.App lék</i>
<i>licence</i>	Pat (na+Acc)	IV.10d		
<i>náklad_i</i>	Act Pat Direction (kam)	IV.23b	<i>nakládat₁</i>	<i>náklad_i písku (na auto)</i>
<i>náklad_j</i>	Act Pat	IV.23b	<i>nakládat₂</i>	<i>náklad_i auta</i>
<i>náklad_k</i>	Act Eff (na+Acc)	IV.23b	<i>?vynakládat</i>	zabudován Pat <i>náklady_k na živobytí</i>
<i>náklad_l</i>	Ø	IV.23b		?abstrakce <i>náklad_l knihy</i> (frazolog. <i>nákladem vlastním / x výtisků</i>)
<i>nárok</i>	Pat (na+Acc)	IV.2,11,27		<i>mít nárok na něco / klást nároky na ...</i>
<i>návrh₁</i>	Act Pat Addr	IV.23a	<i>navrhovat₁</i>	<i>něčí_{pos-subj} návrh₁ něčeho_{Gen-obj} / na něco_{Acc} / inf / že ...</i> <i>něčí_{pos-subj} návrh₁ někomu_{Dat}</i> <i>něčí_{pos-subj}/někoho_{Gen-subj} návrh₁ na něco_{Acc} / inf / že ...</i> <i>něčí_{pos-subj}/někoho_{Gen-subj} návrh₁ někomu_{Dat}</i>

subst.	valenční doplnění	číslo	odkazy	poznámky, příklady
<i>návrh₂</i>	Act Pat Addr Aim	IV.23a	<i>navrhovat₂</i>	<i>něčí</i> _{pos-subj} <i>návrh₂</i> <i>něčeho</i> _{Gen-obj} <i>na něco</i> _{Acc} <i>něčí</i> _{pos-subj/obj} <i>návrh₂</i> <i>na něco</i> _{Acc} <i>návrh₂</i> <i>někoho</i> _{Gen-subj/obj} <i>na něco</i> _{Acc} <i>něčí</i> _{pos-subj} / <i>někoho</i> _{Gen-subj} <i>návrh₂</i> <i>někomu</i> _{Dat} <i>mít návrh na něco X navrhovat něco</i> <i>mít názor na něco</i>
<i>návrh</i>	Pat (<i>na</i> +Acc)			
<i>názor</i>	Pat (<i>na</i> +Acc)			
<i>nominace</i>	Act Pat Addr	IV.22b	<i>nominovat</i>	<i>někoho</i> _{Gen-subj} / <i>něčí</i> _{pos-subj} <i>n. někoho</i> _{Gen-obj} <i>na něco</i> _{Acc} <i>někoho</i> _{Gen-subj/obj} / <i>něčí</i> _{pos-sub/obj} <i>nominace na něco</i> _{Acc} <i>někoho</i> _{Gen-obj} / <i>něčí</i> _{pos-obj} <i>nominace na něco</i> _{Acc} (<i>někým</i> _{Ins})
<i>odchod</i>	Act (Gen _{subj} /Adj _{pos}) Direction (odkud)	IV.40	<i>odejít</i>	
<i>odpověď</i>	Act Pat Addr Eff	IV.22a	<i>odpovídat</i>	<i>někoho</i> _{Gen-subj} / <i>něčí</i> _{pos-subj} <i>o. (někomu)</i> _{Dat} <i>na něco</i> _{Acc} , <i>že ...</i>
<i>odpověď</i>	Pat (<i>na</i> +Acc)	IV.10a		<i>najít odpověď X odpovědět</i>
<i>odvolání</i>	Pat (<i>proti</i> +Dat)	IV.12		<i>podat odvolání proti ... X odvolat se</i>
<i>peníze</i>	Pat (<i>na</i> +Acc)	IV.26c		<i>mít / poskytovat / potřebovat / shánět / získat peníze na něco</i>
<i>pobyt</i>	Act ^Loc	IV.24b	<i>pobývat</i>	
<i>právo</i>	Pat (<i>na</i> +Acc)	IV.9,12		<i>mít právo na něco</i>
<i>prostředky</i>	Pat (<i>na</i> +Acc)	V.3		<i>mít / poskytovat / potřebovat / shánět / získat prostředky na ...</i>
<i>přítel</i>	App			obligatorní (relační jméno)
<i>rozšíření</i>	Act (Adj/Gen) Pat (Gen) Eff (<i>na</i> +Acc)	IV.31c	<i>rozšířit</i>	<i>vypsát (výběrové) řízení na ...</i>
<i>řízení</i>	^Regard (<i>na</i> +Acc)	IV.24a,V.3		obligatorní
<i>řada</i>	Mat (Gen)	IV.10b		<i>stanovit sazbu na ...</i>
<i>sazba</i>	Pat (<i>na</i> +Acc)			supletivum
<i>sen</i>	Pat (<i>o</i> +Loc) ¹		<i>zdat se, ale i snít</i>	obligatorní
<i>skupina</i>	Mat (Gen)			<i>poskytnout slevy na ...</i>
<i>sleva</i>	Pat (<i>na</i> +Acc)	IV.26b		<i>podat stížnost na .../ proti ... X stěžovat si na ...</i>
<i>stížnost</i>	Pat (<i>na</i> +Acc, <i>proti</i> +Dat)			<i>mít šanci na něco</i>
<i>šance</i>	Pat (<i>na</i> +Acc)	IV.26a		
<i>tlak</i>	Act ^Pat (<i>na</i> +Acc)	IV.24c	<i>tlačit₂</i>	<i>dostat trojku z mravů.^Regard</i>
<i>trojka</i>	^Regard (<i>z</i> +Gen)	IV.8b		<i>dostat / poskytnout úvěr na stavbu</i>
<i>úvěr</i>	^Aim	IV.25b		

subst.	valenční doplnění	číslo	odkazy	poznámky, příklady
<i>vliv</i>	Pat (<i>na</i> +Acc)			
<i>zájem</i>	Act (Gen _{subj} /Adj _{pos}) Pat (<i>na</i> +Loc/ <i>o</i> +Acc)	I.1		<i>mít zájem</i> (bez Act) X <i>zajímat se</i> (s Act)
<i>zákon</i>	Pat (<i>o</i> +Loc) ¹ ^Aim (<i>na</i> +Acc)	IV.31a,V.1 IV.31a,V.1		artefakt <i>z. o zdraví lidu</i> .Pat <i>z. na ochranu spotřebitele</i> .Aim

Poznámky:

Tvoří-li substantivum slovesně jmenný výraz, je v jeho valenčním rámci eliminován aktor – ten se ‚přesouvá‘ ke slovesu.

V rámci substantiva tvořícího slovesně jmenný výraz může ‚zmizet‘ i jiný aktant – např. *mít návrh* nemá Adresát.

¹ Alternativně, při dějovém chápání těchto substantiv, mají tato substantiva aktor vyjádřený jako Gen_{subj} nebo Adj_{pos-subj}.

Valenční a kvazivalenční rámce adjektiv

adjektivum	valenční rámec	obsazen	číslo	odkaz	příklady
<i>kladené</i>	Act Direction (kam)	Pat	IV.29b	<i>klást</i>	<i>klást nároky na ...</i>
<i>omezený</i>	Act (Ins) Eff (<i>na</i> +Acc)	Pat	IV.6b, V.2	<i>omezit</i>	<i>o. v něčem.Regard něčím.Means</i>
<i>poskytnuté</i>	Act Addr	Pat	IV.29a	<i>poskytnout</i>	<i>poskytnout slevy</i>
<i>postižený₁</i>	Act ^Regard (<i>na</i> +Acc)	Pat	IV.30a	<i>postihnout₁</i>	
<i>postižený₂</i>	Act (Ins)	Pat	IV.30a	<i>postihnout₂</i>	
<i>shromážděné</i>	Act Loc	Pat	IV.28b	<i>shromáždit</i>	<i>shromáždit prostředky na ...</i>
<i>zvýhodněný</i>	Act ^Regard (<i>na</i> +Acc)	Pat	IV.30b	<i>zvýhodnit</i>	
<i>žijící₁</i>	Pat	Pat	IV.28a	<i>žít₁</i>	
<i>žijící₂</i>	Locative	Act	IV.28a	<i>žít₂</i>	
<i>žitý</i>	Act (Ins)	Pat	IV.28a	<i>žít₁</i>	