

## K nové úrovni bohemistické práce: Využití anotovaného korpusu

### Část 2\*.

Eva Hajičová, Jarmila Panevová a Petr Sgall

#### 2.2. Přímá a nepřímá řeč

Hranice mezi přímou a nepřímou, popř. polopřímou řečí, není vždy ostrá; v PDT zachycujeme všechny druhy jako PAT, popř. EFF slovesa uvozovací věty. Za objekt byla považována i dřív, viz zejm. Šmilauer (1947, 246nn); ten také upozornil, že predikátem uvozovací věty nejsou vždy jen slovesa dicendi. Zejména přímou řeč uvozují i slovesa, u nichž označení řeči jako PAT/EFF není vhodné.

Slovesa pravení (v užším i širším smyslu), na kterých závisí nepřímá řeč, mají ve své obsahové struktuře zpravidla původce výpovědi, adresáta výpovědi, téma výpovědi a její diktum (srov. i Konečná, 1966), popř. ještě zdroj informace, ztvárněné na rovině významové stavby po řadě jako ACT, ADDR, PAT, EFF, popř. ORIG. Všechna slovesa, která zahrnujeme do sloves pravení, se však nechovají stejně co do strukturace svého valenčního potenciálu: liší se v počtu realizovaných míst valenčního pole, v jejich obligatornosti a fakultativnosti i v jejich morfématickém ztvárnění.

Sloveso *vypravovat* (a několik dalších jemu podobných) se štěpí do dvou lexíí (dvou významů polysémnní jednotky), lišících se valenčním rámcem. V jednom významu (*vypravovat*<sub>1</sub>) má rámeček: ACT (nom), ADDR (dat), [PAT (o+lok)], EFF (akuz/Kl)<sup>7</sup>; ve významu 2 spojuje téma a diktum do jediného aktantu a má tento rámeček: *vypravovat*<sub>2</sub> ACT (nom), PAT (o+lok), ADDR (dat).

(21) Vyprávěla<sub>1</sub> nám.ADDR o svých záměrech.PAT, že po obhájení disertace odjede.EFF na rok do USA.

(22) Vyprávěla<sub>2</sub> nám.ADDR o svých plánech.PAT

(23) Vyprávěla<sub>1</sub> nám.ADDR o svém vnukovi.PAT, že ho letos vzala.EFF poprvé na hory

(24) Vyprávěla<sub>2</sub> nám.ADDR, že letos vzala.PAT vnuka poprvé na hory.

Podle tohoto modelu (štěpení na dvě lexie) se chovají také např. slovesa *říkat*, *povídat*, *informovat*, ale také slovesa, která se jen volně řadí ke slovesům pravení (a verbálního sdělování), jako např. *číst*, *psát*, *(u)slyšet*, *dozvídat se*.

To umožňuje odlišit i konstrukce: *četl<sub>1</sub> Dostojevského.EFF*, *četl<sub>2</sub> o Dostojevském.PAT* a *četl<sub>1</sub> o Dostojevském.PAT knihu.EFF*, *řekl<sub>1</sub> nám definici.EFF aktantu*, *řekl<sub>2</sub> nám o definici.PAT aktantu*, *řekl<sub>1</sub> nám o valenci.PAT*, *že aktantem rozumí.EFF*.... Kde je diktum vyjádřeno metajazykovým pojmenováním v akuzativu (*řekl pravdu*, *vykládal nové zprávy z fronty*, *objasňoval podstatu věci*), tam jde o EFF u lexie s číslem 1. Některá taková slovesa mají ORIG, ale ne ADDR (*dozvědět se co od koho*, *slyšet o čem co od koho*).

U těch sloves z této skupiny, která se štěpí na dvě lexie, lze ve významu 1 vyjádřit odděleně téma a diktum, ve významu 2 se téma a diktum realizují v jediné valenční pozici, srov. (25) a (26):

(25) Vyprávěla<sub>1</sub> nám o svém vnukovi.PAT, jak jel.EFF poprvé k moři.

(26) Vyprávěla<sub>2</sub> nám o první cestě.PAT svého vnuka k moři.

Často je možné téma vyjmout z dikta do samostatné syntaktické pozice v souladu se zvýrazněním aktuálního členění, srov. (27):

(27) O svém nadřízeném.PAT předpokládal<sub>1</sub>, že přijde.EFF na party taky.

Nebudeme zde komentovat rozdíly mezi valenčními rámci jednotlivých sloves, protože rozdílů není málo a je potřeba slovesa zpracovat jednotlivě; složitou otázku, jak ekonomicky zachytit všechny možné realizace valenčních rámců konsistentním způsobem, zde jen naznačujeme. Sloveso *mluvit* např. má fakultativní PAT a obligatorní ADDR (ve formě *s+Instr*), srov. Panevová (2001).

Některá jiná slovesa mají jak rámec s fakultativním PAT a obligatorním EFF (lexie 1), tak i rámec bez EFF a s obligatorním PAT (lexie 2), ve kterém vyjadřují spojené téma a diktum formou akuzativu, např. *předpokládat*, *připomenout*, viz (28)(a)-(c), jiná formou *o+Lok*, jako *dozvědět se*, *říkat*, viz (29)(a)-(c):<sup>8</sup>

(28)(a) O té party.PAT předpokládal<sub>1</sub>, že na ni jeho nadřízený přijde.EFF taky.

- (b) Předpokládal<sub>1</sub> 0<sub>PAT</sub>, že jeho nadřízený na party přijde.EFF taky.
- (c) Předpokládal<sub>2</sub> příchod.PAT svého nadřízeného na party.
- (29)(a) Dozvěděl<sub>1</sub> se o své manželce.PAT, že je.EFF mu nevěrná
- (b) Dozvěděl<sub>1</sub>se 0<sub>PAT</sub>, že je mu manželka nevěrná.EFF
- (c) Dozvěděl<sub>2</sub> se o manželčině nevěře.PAT

Podobný soubor aktantů, včetně PAT a EFF (obligatorních nebo fakultativních) mají slovesa *sdělovat, dokazovat, přesvědčit, oznamovat, dohodnout se s, vyprávět, připomínat, snad i ptát se, dotazovat se (na koho/co, zda...), dodávat, zdůrazňovat (o kom/čem, že), upozorňovat (na koho/co, že...)*; bez ADDR jsou *zveřejnit, konstatovat, poznamenat, uvést, myslet, vědět, domnívat se, předpokládat* aj.; slovesa *mluvit, poukazovat nač* patrně nemají EFF. V nepřímé řeči a v dalších obsahových větách rozbíráme např. *Bál se toho, že zaprší tak, že zaprší* má funktor PAT, *toho* se jako odkazovací slovo v TGTS vynechá. Podmínky výskytu odkazovacích slov (obligatorních tam, kde rekční apod. pád není nominativ ani bezpředložkový akuzativ) bude třeba podrobně vymezit, s odlišením od zájmen řídicích vztažné klauze, viz Panevová (1971).

P ř í m á řeč není zdaleka tak pevně zakotvena v jazykovém systému jako řeč nepřímá. Jde o "včlenění neorganické části" (Šmilauer, loc. cit.) a její chápání jako PAT, popř. EFF (které zachováváme v typických případech) není vhodné tam, kde přímá řeč má podobu celého řetězu vět, takže není jednotným syntaktickým celkem, i když v jistém smyslu obsazuje valenční pozici uvozovacího slovesa. Existují i případy s valenční pozicí PAT (EFF) obsazenou jiným členem, než je přímá řeč, viz např. (30):

- (30) "Nechci být jen nástrojem tvých vášní, ... chci, aby sis vážil mé duše, mé inteligence, mého mozku!" recitovala v divadle malé formy mladá herečka moderně pojatého klasika. (H. Pawlowská: Zoufalé ženy dělají zoufalé věci. Praha: MOTTO, 1993, s. 19.)

Jako PAT u *recitovala* tu lze vidět slovo *klasika* a přímá řeč se dá chápat jako PAT elidovaného sémanticky prázdného slovesa, které

by mělo postavení doplňku. Zatím můžeme jen konstatovat, že obě zvláštnosti přímé řeči patří k aspektům, ve kterých jazyk výjimečně překračuje hranice jinak zachovávaných syntaktických pravidelností. Potvrzuje to vhodnost přístupu, který se nesnaží předem obecně vymezit celý prostor možných souborů jevů, ale postupuje i při explicitním popisu od prototypického, bezpříznakového jádra k vrstvám kontextově omezeným a specifickým, až po jednotlivé výjimky, které narušují i hodně obecné principy (srov. Sgall 2001).

### 2.3. Valence substantiv

O otázce hranic mezi restriktivním přívlastkem (RSTR), např. poslanec Svoboda a doplněním identity (ID), např. pojem subjekt(u) jsme se zmínili v odd. 2.1.2. U RSTR i u deskriptivního (nerestriktivního) přívlastku (DES), u kterého jsou tak časté chyby v interpunkci, jde o fakultativní volná doplnění. Opozice mezi takovou valencí v širším slova smyslu a valencí ve smyslu užším (tj. souboru aktantů a těch volných doplnění, která jsou u daného řídicího slova obligatorní) je dosud otevřenou otázkou v několika bodech:

(i) dvojznačnost mezi volným doplněním ID (např. *otázka kvadratury kruhu*) a aktantem PAT (*otázka jeho ubytování u nás*), popř. ACT (*otázky našich čtenářů*), podobně u *problém, teorie, seminář gen/o+lok*;

(ii) funkce členu vyjádřeného od+gen (např. *zařízení od firmy X*), který rozvíjí nejrůznější substantiva, takže tu nejde o aktant ORIG; vidět tu volné doplnění odkud (DIR.1) patrně nelze vzhledem k distribučním kritériím (nemožnost koordinace); aktant ORIG rozvíjí slovesa *obdržet, vypůjčit* apod., i odpovídající dějová substantiva (např. u spojení *po obdržení potřebných údajů od pojišťovny* jde o homonymii mezi adverbálním ORIG a volným doplněním substantiva *pojišťovna*); ACT můžeme vidět u dějových substantiv a u názvů výsledků děje jako *dar, příspěvní, půjčka*;

(iii) vyjádření ACT u dějových substantiv instrumentálem a PAT genitivem (*vytvoření díla umělcem*); pokud základové sloveso (a tedy ani dějové substantivum) nemá PAT, vyjadřuje genitiv jeho ACT (příchod vašeho otce); genitiv chápeme jako vyjádření PAT i u jmen artefaktů (*pomník T. G. Masaryka*); přesnější vymezení těchto skupin, i otázka fakultativnosti PAT (zejm. v opozici mezi jeho absencí a všeobecným PAT) předpokládá podrobnější zpracování jednotlivých valenčních rámců ve slovníku;

(iv) hranice mezi APP a MAT ve spojeních jako *část člověka.APP* a *část lidí.MAT*, i *tisíc(e) lidí*.

## 2.4. Některé specifické syntaktické otázky

### 2.4.1. Vyjádření „stavu“ a tzv. „neshodný doplněk“

U slovesa *být* (u kterého v TGTS nerozlišujeme mezi sponou a slovesem existence) nacházíme kromě valenčního PAT (*Je učitel/učitelem, Obloha je modrá*) a volných adverbialních určení (*Je na zahradě*) celou řadu předložkových určení, pro něž by se obecně hodil funktor *s t a v* (STATE). Ani zavedení tohoto funktoru samo situaci neřeší, protože sémantických diferencí je tu velmi mnoho: např. *pacient byl při smyslech, je na řadě, je po operaci, je bez obtíží, košile je bez límečku, byl pod tlakem, je v podmínce*. Je třeba tu zkoumat zejm. ustálenost předložkových výrazů, i uplatnění opozice + vs. – (*být v provozu, být mimo provoz*). Zřejmě by bylo třeba pracovat s celou řadou syntaktických gramatémů; zatím volíme užití funktoru významově nejbližšího (včetně přenesení významů, např. LOC pro *je v krizi*, TWHEN *byl po operaci*) se zvláštním symbolem \*, signalizujícím, že jde o stav. Obdobné doplnění se vyskytuje i v pozicích *d o p l ň k o v ý c h* (s funktolem COMPL), srov. (alespoň v jednom z významů) např.:

(31) Zastihli ho s aparátem v ruce.

(32) Přepadl prodavačku s tržbou v kufříku.

(33) Přepadl prodavačku s pistolí v ruce.

Takové předložkové výrazy nebývají k doplňku řazeny, protože nevykazují shodu. Bylo by možné doplnit tu elidované prázdné sloveso (srov. "*mající, maje*") s koreferencí typu kontroly, jako u doplňku; výběr interpretace je podmíněn lexikálním a pragmatickým kontextem, srov. rozdílnou interpretaci vět (32) a (33).

#### 2.4.2. Otázky dalších slovních druhů

Z okruhu *z á j m e n* vyžaduje podrobnější řešení na základě rozsáhlejšího materiálu především problematika tzv. 'pronouns of laziness' (víceznačného, popř. vágního užití zájmen); např. v (34) není jasné, zda tvar *ji* zastupuje spojení *svou ženu* nebo *jeho ženu*:

(34) Můj bratr svou ženu políbil a já jsem ji políbil taky.

Je-li referenční vztah z kontextu zřejmý, řeší se ve vzorovém souboru v atributu COREF, i tam, kde je zájmeno (i se slovesem) elidováno.

*P ř í s l o v c e* mají často primární funktor způsob (MANN), ale jsou tu ovšem možnosti sekundární (je třeba formulovat kritéria pro jejich rozlišení) a příslovce s bezpříznakovým funktorem jiným, jako *včera*.TWHEN, *tady*.LOC, *bohužel*.ATT. Seznam v Manuálu by měl zahrnout všechna příslovce s jiným funktorem než MANN.

*U č á s t i* c není jasná hranice mezi rematizátory (RHEM, 'focalizers') a funktorem postoje ('Attitude', ATT). K tomu se stručně vrátíme v příštím oddílu, kde probíráme otázky aktuálního členění .

### 2.5. Aktuální členění věty

#### 2.5.1 Úvodní poznámky

Teoretická koncepce aktuálního členění věty (AČ) je ve FGP soustavně propracovaná, viz pro češtinu zejm. Sgall ad. (1980) a s diskusí o formální sémantice AČ i o některých dalších důležitých aspektech Hajičová ad. (1998). Důraz se vždy kladl na formulaci konsistentních operativních kritérií a na ověřování na rozsáhlém

materiálu v češtině i ve srovnání s jinými jazyky. Při anotování PDT jsme proto mohli formulovat instrukce pro anotátory poměrně jednoznačně. Ty ovšem jen specifikují hypotézu, kterou je třeba na základě dalšího materiálu ověřovat, upřesňovat a třeba i měnit.

Uzlům závislostního stromu se přiřazuje jedna z následujících hodnot atributu TFA ('topic-focus articulation'); podle nich se určuje základ a ohnisko věty (viz Hajičová – Sgall, 2001):

T: pro uzly kontextově zapojené (KZ) nekontrastivní

C: pro KZ uzly stojící v kontrastu

F: pro uzly kontextově nezapojené (KN)

Základním vodítkem je postavení výrazů v povrchovém slovosledu, respektive v ATS. Členům závislým na slovesu zleva je přiřazena hodnota T nebo C (k té viz níže), s výjimkou jasně (v mluvené podobě intonací typickou pro vlastní ohnisko) vyznačeného členu vyjadřujícího „novou informaci“ (nebo nový vztah), jenž dostane hodnotu F. Tu dostane i člen závislý na slovesu nejvíce zprava, a F zpravidla dostanou i členy závislé na slovesu a stojící mezi slovesem a tímto posledním uzlem, pokud není kontextem dáno, že jsou KZ (např. se tento člen opakuje nebo je zahrnut v antecedentu apod.); takový KZ uzel dostane hodnotu T. Hluběji závislé členy (jako např. přívlastek) považujeme (s výjimkami obdobnými jako v předchozím bodě) za KN, mají hodnotu F.

Tato pravidla lze ilustrovat následujícím příkladem, kde (35b) je velmi zjednodušenou linearizovanou podobou věty (35a); uvádíme u každého uzlu jen hodnotu lexikální, syntaktickou a aktuálně členskou; závorky nahrazují hrany: levá závorka značí hranu dolů, pravá závorka návrat k řídícímu uzlu.

(35)(a) České radiokomunikace musí v tomto roce rychle splatit dluh televizním divákům.

(b) (radiokomunikace.ACT.T (České.RSTR.F)) ((v tomto.RSTR.T) roce.TEMP.T) musí-splatit.F (rychle.MANN.F) (dluh.PAT.F (divákům.RSTR.F (televizním.RSTR.F))).

### 2.5.2. Kontrastivní základ (C)

Podrobná analýza příkladů s rematizátory (RHEM),<sup>9</sup> např. *jenom, pouze, také, i, a(le)spoň* (v. Hajičová ad., 1998, 151) vedla k vyčlenění zvláštní hodnoty pro kontrastivní část základu (k prvním motivacím takového postupu srov. Koktová, 1999). Rozhodující argumenty byly dva; jeden je založen na hledisku sémantickém, druhý na využití výrazových prostředků v češtině:

(i) Z hlediska sémantického ukazují na nutnost rozlišování kontrastivního základu, C, především příklady s rematizátory v základu, protože je takové rozlišování relevantní pro jeho dosah.

(36) (Kdo kritizoval i papeže jako nástroj kapitalismu?) I papeže kritizoval JIRKA.

(36') (i.RHEM.T) (papeže.PAT.C) kritizoval (Jirka.ACT.F)

Slovo *papež*, v předcházející větě rematické, je tu kontrastivním tématem. Jak ukazuje příklad (37) z PDT, nemusí být nositel C výrazem totožným s antecedentem v předchozí větě:

(37) (Před několika týdnů však Francii opustil korunní svědek.) Právě Jacques Glassmann.C ... v květnu 1993 veřejně potvrdil snahu funkcionářů Olympique uplatit utkání s Valenciennes.

(ii) Má-li hodnotu C zájmeno, užívá se výhradně silný tvar, jak ukazuje obměna příkladu (36):

(38) (Kdo kritizoval i papeže jako nástroj kapitalismu?) I jeho.C kritizoval JIRKA.

*Toho lze využít i jako operativního kritéria pro hodnotu C. Kritériem je i (snad fakultativní) prozodické zvýraznění prvku s C stoupavým přízvukem; v psaném korpusu je jeho vyznačení (např. typem písma) vzácné, ale vždy je možné si představit rozebíraný text v jeho mluvené podobě. O to větší důležitost jistě má vytváření a rozbor korpusu mluveného, i akustická analýza řeči.*

Srovnání příkladů (39) a (40) naznačuje, že se C nevyskytuje jen v konstrukcích s rematizátory.

(39) (Kde se mluví česky?) Česky se mluví V ČESKU.

(39') Česky.T (se) mluví.T (v) Česku.F



(40) (Mluví se česky v Česku nebo na Slovensku?) Česky se mluví v ČESKU, na Slovensku se mluví SLOVENSKY.

(40') Česky.C (se) mluví.T v-Česku.F, na-Slovensku.C se-mluví.T slovensky.F

Podobně lze srovnat větu (41) s větou (42) a větu (43) s větou (44):

(41) (Petr ji nazval KONZERVATIVCEM.) Potom jeho urazila ONA.

(41') Potom.T jeho.C urazila.T ona.F

(42)(Petr ji nazval KONZERVATIVCEM.) Potom ho (ona) OPUSTILA.

(42') Potom.T ho.T [ona.T] opustila.F

(43) (Kdy jsi ho viděl?) Viděl jsem ho VČERA.

(43') [já-T] ho-T viděl.T včera.F

(44) Jeho jsem viděl VČERA (a ji už PŘED TÝDNEM).

(44') Jeho.C [já.T] viděl.T včera.F (a ji.C už.F před-týdnem.F).

Hodnotu C nemusí mít jen jeden člen; v češtině jsou členy s C

zpravidla umístěny na začátku věty:

(45) (Rodiče odjeli na dovolenou a děti svěřili příbuzným.) Syna.C v sobotu.C

odvezli k babičce a dceru.C v neděli.C k tetě.F

(46) Syna.C v sobotu.C otec.C odvezl k babičce.F a dceru.C v neděli.C matka.C

k tetě.F

Kontrastivní může být i hlouběji zapuštěný KZ člen ohniska, srov.

silný tvar *tebe* v příkladu (47); celá vedlejší věta v daném kontextu

patří do ohniska, ovšem vlastním ohniskem je jen sloveso *mrzelo*.

(47) (Všem se vývoj událostí líbil, ale ty jsi měl odlišný názor.) Přiznal ses, že tebe

osobně to MRZELO.

(47') [ty.T] přiznal-ses.T tebe.C osobně to.T mrzelo.F

Kontrastivní základ je vybrán z množiny alternativ, podle faktorů

textových i situačních. Tato množina bývá vyjádřena s různým

stupněm explicitnosti: od silného typu kontrastu, kdy je vymezena

výslovně (jako v příkladu (48)) nebo aspoň nějakým zahrnujícím

výrazem (příklad (49)), přes vymezení slabší (příklad (50)) až

k případům sémantického asociativního vztahu části k celku, v. (51).

(48) (Dalšími uchazeči ... jsou Korejská republika a Mexiko,) přičemž Korea.C již

svou žádost PŘEDLOŽILA.

(49) (O ženský titul bojovaly na severu Čech dva celky z Moravy.) Zkušenější

Olomouc.C momentálně zápasí s KRIZÍ ..., (a tak jen s obtížemi potvrdila roli

favorita.)

(50) (V německé, anglické i italské fotbalové lize pokračují dosud vedoucí celky neohroženě ve své cestě za titulem.) Vedoucí celek nejvyšší anglické divize Blackburn Rovers.C sice jen remizoval se čtrnáctým týmem tabulky Norwich CITY, ale svůj náskok před druhým Manchester United přesto zvýšil na tři body.  
(51) (Loni jsme renovovali celý dům.) Omítku.C jsme natřeli na žluto, střechu.C jsme taky vyspravili, všechna okna.C jsme vyměnili za dvojitá.

Kontrast je vnímán citlivěji, pokud jsou prvky s C při koordinaci či v těsném kontextu v analogické větné pozici:

(52) (Včera se hrál zápas mezi Brnem a Ostravou.) Domácím.C se dařilo ZE ZAČÁTKU. Hostům.C se povedl až druhý POLOČAS.

*Všeobecně se soudí, že v ohnisku věty jde vždy o výběr z množiny alternativ. Pokud bychom tedy hledali specifický kontrast v ohnisku, muselo by jít o člen zdůrazněný navíc, nějakým příznakovým způsobem. Zatím jsme nenašli argumenty pro takovou distinkci v češtině, a proto tuto otázku necháváme otevřenou. Pro některé jazyky však taková opozice může být nutná; např. pro němčinu se upozorňuje (v. i Hajičová, 2000), že prvek ohniska se může přemístit doleva jen tehdy, jde-li o ohnisko kontrastivní.*

### 2.5.3 Určení hranice mezi základem a ohniskem a pořadí uzlů

Přiřazení hodnot atributu TFA uzlům v PDT slouží pro určení základu a ohniska věty. Součástí základu jsou uzly závislé přímo na slovesu, pokud mají hodnotu T nebo C, a všechny uzly na nich přímo či nepřímo závislé; ostatní závislé uzly jsou v ohnisku. Sloveso patří do ohniska, pokud se neopakuje nebo není obsaženo v antecedentu apod. Pokud sloveso i všechny uzly na něm přímo závislé mají T, pak hranice mezi základem a ohniskem je před prvním uzlem s F ležícím na poslední (tj. nejpravější) cestě dolů od slovesa, srov.:

(53) (Kterého učitele ti řekl, že včera potkal?) Řekl, že potkal učitele matematiky.

(53') (on.T) řekl.T, že (on.T) (potkal.T) ((učitele.T (matematiky).F)

Hlavní sloveso této věty i jeho konatel a patiens jsou KZ, stejně tak oba členy (konatel i patiens) závislé na slovese vedlejší klauze; jediným KN členem je restriktivní přívlastek *matematiky*; ten je také

jediným prvkem ohniska. Hajičová ad. (1998, 135) zavedli termín 'proxy focus' pro takový KZ uzel odlišný od slovesa hlavní klauze, jemuž je podřízeno (tj. závisí na něm) ohnisko celé věty. Proxy focus sám je ovšem součástí základu věty.

V dosavadních úvahách jsme předpokládali, že KZ není jen bezprostředně řídicí člen proxy focusu, ale i všechny uzly ležící na cestě vedoucí od něho k vrcholu stromu. Tento předpoklad se však nepotvrdil, jak dokládá příklad (54), kde pozice intonačního centra ukazuje, že vlastním ohniskem je uzel podřízený KZ uzlu, i když je v horním keři jiný uzel s hodnotou F.

(54) Prozradíte.F DALŠÍ.F elitní.T jména.T?

Na nutnost dalšího zjemnění definice proxy focusu ukazují i obměny věty (53) uvedené zde jako (55)(b) a (c) s kontextem v (a):

(55)(a) (Kterého učitele ti řekl, že včera potkal?)

(b) Řekl, že před školou potkal učitele matematiky.

(c) Řekl, že potkal před školou učitele matematiky.

Ve větě (b) je člen před školou zřejmě KZ a ohniskem je jen přívlastek *matematiky* (závislý na proxy-focusu *učitele*). Ve větě (c) postavení členu *před školou* za slovesem ukazuje spíš na to, že toto místní určení je KN; je otázkou, jak určit charakter patientu; kontext ukazuje, že vlastním ohniskem je jeho přívlastek *matematiky*.

Při rematické platnosti p o m o c n é h o slovesa (včetně vlastních sloves modálních) řadíme celé sloveso do ohniska; ačkoli je jeho lexikální hodnota KZ, má sloveso v TGTS hodnotu F; analýza podrobnější (s Petkevičovým formálním rámcem) označí příslušnou gramatickou hodnotu, např. gramatém RES v (56), jako KN.

(56) Uděláno to ještě NEMÁ.

(56') on.T to.T ještě.F ne.F uděláno-má.F

Má-li sloveso jednoduchý tvar (jako v *Ještě to neudělal, ale určitě to udělá*), povrchový slovosled ovšem takový rozdíl (zde mezi KZ *udělat* a KN futurem) nevyjadřuje, ale TGTS ho i pak reflektuje.

Specifické je i vyjádření AČ u přímé řeči: sloveso uvozovací klauze má zpravidla T, ačkoli tato klauze bývá na konci souvětí.

Pro pořadí uzlů v TGTS platí primárně, že KZ uzly na řídícím uzlu závisejí zleva a uzly KN zprava (opačné pořadí může být u rematizátorů, u sester v koordinaci a u proxy focusu). Uzly s C jsou vlevo od svých sesterských uzlů. Předpokládané pořadí uzlů v tématu v TGTS zpravidla odpovídá povrchovému slovosledu: částice odkazující k předcházejícímu kontextu, jako *tedy*, *proto*, *přesto* (s funktořem PREC) předcházejí před slovy s INTF a s ATT (např. *On našťestí Jirka nepřišel*), za těmi následuje vlastní téma, pak kulisy a dál zbývající ('nevlastní') část tématu; v obou těchto skupinách jsou vlevo doplněná slova s ELID nebo ELEX (viz odd. 3.2) a pak příklony. Sloveso (s T, C i F) zpravidla stojí na hranici mezi svými KZ a KN dcerami.

Tektogramatické značkování přináší zvlášť cenný materiál pro ověření hypotézy o systémovém uspořádání prvků (o jeho zjišťování srov. Sgall ad., 1980, 1995). Bude zajímavé porovnat (zčásti snad i automaticky) pořadí v předpokládaném ohnisku se systémovým uspořádáním postulovaným pro češtinu a vyhledat odchylky od dosavadní hypotézy (třeba pro jednotlivá slovesa, popř. i z hlediska generačního zařazení mluvčích apod.), z čehož asi vyjdou i podněty pro překlasifikování funktořů. Podle dřív získaných výsledků řadíme v PDT rematizátory mimo stupnici systémového uspořádání, jako prototypicky míň dynamické než sloveso; k tomuto závěru nás vede zjištění, že v dosahu rematizátorů je vždy celé ohnisko věty. Bylo by množné považovat rematizátory za gramatémy a nikoli za funktořy (srov. už Sgall, 1967).

Postavení rematizátorů zasluhuje ještě další zpracování. Zatím chápeme RHEM primárně jako sestřu uzlu, který uvozuje: např. ve větě *Viděli jen ČERVENÉ auto závisí jen* i adjektivum na slovu *auto*. V koordinaci (apozici) RHEM v TGTS závisí na spojce,

pokud uvozuje celou konstrukci (např. ve *vysílání zejména televizní a rozhlasové* visí RHEM *zejména* na spojce *a*), ale jsou i jiné možnosti: ve spojení *zejména televizní.F a rozhlasové.F vysílání.T* je RHEM sestrou spojky a konstrukce *vysílání televizní a zejména rozhlasové* je koordinace dvou substantiv, s elipsou druhého výskytu, na kterém závisí i RHEM. Závisí-li RHEM na KZ slovesu, je umístěn vlevo od něho (s T) a uvozuje svou matku; pokud KZ sloveso není v dosahu RHEM, závisí na něm RHEM zprava (s F); závisí-li RHEM na slovesu KN, má F a uvozuje svou matku i jí podřízené členy s F, ale v TGTS závisí zleva (v. Hajičová, 1975; 1984).

Nejasné je i rozlišení RHEM od spojek: máme považovat např. spojení *a také* celé za spojku? A jak má být určeno např. *ani* ve větě *Nepřišel tam ani PAVEL*? Funkce RHEM a spojky se někdy kombinují v jednom slovu, srov. rozdíl mezi spojkami *a* a *i*.

## 2.6. Koordinace a apozice

### 2.6.1 Hranice mezi druhy koordinace

Rozdíly mezi spojovacími výrazy *o d p o r o v a c í m i*, jako *ale*, jenomže, ovšem, jsou zčásti jistě sémanticky relevantní. V TGTS, kde spojky mají svůj uzel, to lze zachycovat různými lematy. Pro teoretické chápání funkcí spojek bude nutné zvažovat uplatnění syntaktických gramatémů, popř. i jemnější třídění funktorů.

Za souřadné spojení považujeme *i d ů s l e d e k*, CSQ ('consequence'); relevantním měřítkem je, že se jeho spojky nevyskytují na začátku souvětí, že tedy mají postavení mezi spojovanými klauzemi, typické pro koordinaci:

(59) Potom ... přesune na noční provoz prostřednictvím počítačového systému, takže na něj budou mít přístup obchodníci ze všech časových pásem.

(60) Móda je stále víc přivádí z přírody do měst, a tak i pro RAV 4 platí označení městský terénní automobil.

Bude třeba se zabývat např. *i s l u č o v a c í m* užitím

disjunktivních spojek *nebo*, *či*: Jde tu snad o postup od kontextů, ve

kterých spojka sama vyjadřuje disjunkci, ale obsahově z celého spojení vyplývá platnost slučovací, jako (57), (58), k případům, kdy už se spojka (jen v určitých typech promluv v publicistice ap., nebo i obecněji?) chápe jako slučovací. Co z toho se týká jen knižního *či*, v. (59), a co i základní spojky nebo, jako v (60), popř. dalších spojek?

(57) Uložili tam ty, kteří jsou ranění nebo nemocní. (= ...ty, kteří jsou ranění a ty, kteří jsou nemocní.)

(58) Jako příklad můžeme uvést raněné nebo nemocné.

(59) V Praze sice ubylo Američanů, ale bohatě je nahradili Italové či Rusové.

(60) Uložili tam raněné nebo nemocné.

Všimněme si, že v (57) může být i *...bud' ranění nebo nemocní*,

kdežto v (60) by užití *bud'* vylučovalo slučovací význam věty.

Otázka možné dvojznačnosti vět jako (58) nebo (59) je, pokud víme, otevřená.

### 2.6.2. Apozice

Zatím u apozice rozvití zprava chápeme jako rozvití jejího pravého členu, např. ve spojení *český král, Karel, který...* vztažná klauze závisí na *Karel*; obecně tedy rozvití závisí na bližším členu apozice. Bude ale nutné rozlišení jemnější, např. u specifikující apozice se asi pravostranné rozvití může týkat apoziční konstrukce jako celku.

### 2. 6. 3. Reciprocita

Uzel s lexikální hodnotou Rcp zaplňuje místo „nepřítomného“ valenčního členu, tj. členu přítomného obsahově, ale ve větné stavbě přesunutého do koordinace s jiným aktantem, jako v (61), nebo zahrnutého v plurálovém aktantu, jako v (62); v. Panevová (1999).

(61) Jan.ACT a Marie.ACT se setkali Rcp.PAT před školou.

(62) Chlapci.ACT se (spolu) prali Rcp.PAT před školou.

Reciprocita se týká i volných doplnění (srov. (63), kde *po sobě* bude v TGTS zastoupeno uzlem s hodnotou Rcp.LOC, i (64)) a vztahů adnominálních, (65):

(63) Honza.ACT a Kuba.ACT se po sobě válejí

(64) Družstva.ACT se k sobě (= Rcp.DIR3) přiblížila.

(65) Došlo k dohodě mezi členy.ACT (Rcp.PAT) ODS a ČSSD.<sup>10</sup>

Reciprocita úzce souvisí s otázkou hypotaktické koordinace; jsou tu případy jasnější (jako (66) a (67)) a méně jasné ((68) a (69)):

(66) Jan.ACT se líbal s Marií.PAT

(67) Jan.ACT a Marie.ACT se líbali Rcp.PAT

(68) Jan.ACT s Marií.PAT se líbali.

(69) ?Jan s Marií se líbal.

Věta (69) se nejeví jako plně akceptovatelná, mezi (67) a (68) vidíme sémantický rozdíl spočívající v míře aktivity Jana a Marie na ději, proto je analyzujeme různě (*líbat se s někým*.PAT je jiná lexikální jednotka než *líbat někoho*.PAT). Pro další výzkum ponecháváme otevřené, jak zapsat různé druhy reciprocit (inherentní i vyjádřenou explicitně, pomocí *spolu, navzájem*) ve slovníkovém hesle slovesa, jehož doplnění se na reciprocitě podílejí.

## 2. 7. Koreference

Tu koreferenci, kterou nazýváme *g r a m a t i c k o u* (srov. Hajičová ad., 1985-7, Panevová, 1991), je nezbytné v PDT zachytit, neboť jde o jevy systémové. U koreference se vztažným zájmenem a u reflexív (osobních a přivlastňovacích, u tzv. vázání) do atributu COREF vyplňuje lema jejich antecedentu. U infinitivu a některých nominalizací (v. Panevová, 1996, Panevová ad., 2002) se doplňuje nevyjádřený „kontrolovaný“ člen (subjekt infinitivu, aktant dějového jména) s lematem *Cor* (v atributu COREF má lema antecedentu, tj. kontrolujícího členu). I u tzv. slovanského akuzativu s infinitivem vidíme vztah kontroly, srov. větu (70) a náznak její anotace:

(70) Honza slyšel Karla otvírat dveře.

(70') Honza.ACT slyšel Karla.PAT Cor.ACT.ELID.COREF:Karel otvírat.EFF dveře.PAT.

Podrobnější zpracování vyžadují jak adnominální infinitivy (*snaha, schopnost, idea, nápad odejít*), tak nominalizace typu (71):

(71) Matěje podezírají z krádeže slepic.

(71') Matěje.ADDR podezírají z krádeže.PAT slepic.PAT

Cor.ACT.ELID.COREF:Matěj

Seznamy slov spojených s kontrolou podřízené predikace (včetně údajů o jejich gramatických vlastnostech) se postupně sestavují. S kontrolou bude třeba pracovat i u přechodníků a u všech doplňků (i těch, které nevykazují tvaroslovnou shodu, např. *přišel naboso* jako *přišel bos*, snad i *našli ho s cigaretou*).

Abychom vytvořili dobrou materiálovou základnu i pro studium *p r o m l u v y*, zachytíme aspoň zčásti i některé jevy koreference (a anaforické návaznosti) *t e x t o v é*. Tu ostatně můžeme vidět nejen u zájmen a substantiv, ale i u sloves, resp. klauzí a výpovědí, srov. Hlavsovo (1975, s. 23-26) pojetí denotace.

## 2.8. Pasívum

Pasívum se od aktiva často (zejména v angličtině, francouzštině ap.) liší slovosledným pořadím aktantů; proto jsme původně počítali s tím, že významový rozdíl mezi oběma rody je dán rozdílem AČ. Později jsme se však přesvědčili, že nelze ani při shodném AČ počítat s plnou synonymií (obecnou zaměnitelností), nýbrž jen s kvazisynonymií aktiva a pasíva, protože se významově liší aspoň v kontextech s doplněním typu ATT; doplnění jako dobrovolně, rád, s chutí mají totiž určitý sémantický vztah k entitě, ke které referuje ("povrchový") podmět: *Spatřili ho dobrovolně. - Byl jimi spatřen dobrovolně*. Vzhledem ke vzácnosti takových pasivních vazeb v češtině zapisujeme prozatím v TGTS pasívum stejně jako aktivum, takže rozdíl se pozná pouze ze vztahu mezi ATS a TGTS dané věty. Úplnějším řešením by bylo zaznamenávat, který aktant je podmětem (není-li jím ACT); může to být PAT, ale i ADDR (*Byl o tom*



*informován*). Bylo by možné zavést gramatém SB (*On.ADDR.SB* Gen.ACT *informovat to.PAT*).

## 2.9. Slovní zásoba a tvoření slov

Zatím se v PDT nezpracovává složení lexikálního významu, většinou ani ne tam, kde je zřejmé z morfématické derivační struktury. Bude třeba zpracovat produktivní *o d v o z o v á n í*, tzn. k odvozenému slovu se ve slovníku automaticky najde jeho derivační základ. Dnes tak zpracováváme jen několik typů, zejména s Kuryłowiczovou "syntaktickou derivací" (beze změny lexikálního významu), jako:

### (i) substantivum - adjektivum:

např. *otcův* lze rozbírat jako *otec.APP (otcův stůl)*, popř. ACT (*otcův příchod*) aj., *matčin* jako *matka.APP (ACT aj.)*;

### (ii) adjektivum - adverbium:

např. *ostře* se rozbírá jako *ostrý.MANN*; podobně *statečně* v TGTS derivujeme od *statečný* atd., ale tam, kde lze vidět vztah opačný, máme zatím i v TGTS samostatná slova, např. *dnes* a *dnešní*, *tam* a *tamní*, *jinak* a *jiný*; u adverbii zřetelových řadíme zatím např. *názorově* k lexikální jednotce *názorový*, ale ne k *názor*;

### (iii) sloveso - substantivum, adjektivum (nominalizace):

např. *psaní*, *napsání*, *psaný*, *napsaný*, *píšící* se rozbírají jako N, A s valenčním rámcem odvozeným z rámce slovesa; produktivně tvořená slova jako výše uvedená (i *vyrábění*, *vyrobení* atd.) snad bude možné v TGTS derivovat od slovesa (s neutralizací času); odvozeniny ne úplně produktivní, např. *výroba*, je nutné vidět jako zvláštní lexikální jednotky, i proto, že nerozlišují vid; výhledově lze i substantiva jako *příchod* vs. *přicházení*, kde je vid rozlišen, derivovat od slovesa;

### (iv) sloveso - sloveso:

např. *šířit* a *šířit-se* jsou v TGTS a ve slovníku jako dvě slova; hranice mezi touto derivační částicí *se* a tvarem zájmena (např. v *pozorovat se/sebe*) nejsou zcela jasné, jde o široké přechodné pásmo.

**(v) zájmeno - zájmeno:**

podobně jako u přivlastňovacích adjektiv, v. (i), je možné tvar s APP (i s ACT u dějových jmen ap.) chápat jako derivovaný od základního slova i u zájmen, přičemž by se hodnota gramatému rodu určovala podle základu: její by se zapisovalo jako *on.FEM.APP*, *jejich* jako *on.PL.APP* s rodem podle kontextu; i zájmena *jehož*, *jejíž* atd. by se mohla derivovat od *který* nejen v příkladech jako *kolega*, *jehož.PAT jsme se nedočkali*, ale i *kolega*, *jehož knihu čtu* (kde tvaru *jehož* v TGTS odpovídá *který.APP*, závisející na *knihu*).

V TGTS bude třeba zachycovat *r o d* u substantiv (vzhledem k možným personifikacím, k mytologii aj.), zatímco tam, kde má rod jen platnost shody, s ním pracujeme jen v ATS; pozornost vyžadují kontexty, ve kterých je antecedent shody elidován (např. *Přišli/y/a včas*); tady může být rod určen po doplnění antecedentu elidovaného zájmena (což při textové koreferenci nebude snadné). U tzv. bezrodých zájmen je třeba rod v TGTS na základě shody doplnit (*Vy.ANIM jste přišli* - *Vy.FEM jste přišly*).

Vnitřní syntax složených vlastních *j m e n* chápeme jako závislostní podstromy, ale nerozlišujeme v nich specifické funktoři, tj. ve jménech jako *Žďár nad Sázavou*, *Jindřichův Hradec* mají závislá slova funktoři RSTR. Relevantní tu je rozdíl mezi vlastním jménem v užším smyslu a názvem, např. v názvu *Rada pro Českou televizi* je třeba funktoři rozlišit (*televize.AIM český.RSTR*).

Přechodná pásma mezi lexikem a gramatikou je třeba dále zkoumat v oblastech, jako jsou méně produktivní odvozovací typy, jednotlivé typy frazémů (v. odd. 3.3) a nezcela gramatikalizovaných kategorií (vid, stupňování aj.).

Z dalších otázek lexika připomeňme jen specifickou otázku výrazu intenzifikace (s funktoři INTF): např. *on* rozbíráme jako INTF ve větách jako *Ono prší*, *On Jirka přijde*, ale jako ACT např. v *On přijde*, kde nedochází k tzv. zdvojení podmětu. Příznakové užití

zájmena *ten* (snad jen tvaru *to*) jako v *On si to (Jirka) šine; To on si to šine* je snad frazeologicky omezené na nevelkou skupinu sloves.

## 2. 10. Gramatémy

### 2.10.1. Vid a druhy děje

Kategorii vidu (v užším smyslu) pokládáme za gramatickou a na TGTS rozlišujeme tři její hodnoty: (i) PROC (procesuálnost, primárně vyjadřovanou tvary nedokonavými), (ii) CPL (komplexnost, s dokonavými tvary) a (iii) rezultativnost (RES), která odpovídá Mathesiovým perfektním časům a Hausenblasovu výslednému stavu (v. Panevová ad., 1971, Sgall ad., 1986, odd. 2.43). Řazení RES k vidu je poměrně ekonomické řešení, umožněné tím, že se RES s opozicí vidu nekombinuje (jde o dokonavé tvary). Připomínáme, že i prostý dokonavý tvar (uvařil jsem) má někdy stejný význam RES jako mám uvařeno; toto užití bude možné rozlišit až později, stejně jako příklady typu mám vařeno, které snad lze rozbírat jako RES, i když obsahují morfém nedokonavosti.

Do gramatického vidu v širším slova smyslu počítáme díky její produktivitě opakovanost (IT, 'Iterativeness'); tento gramatém, zachycovaný jen ve vzorovém souboru, je kompatibilní se všemi třemi hodnotami vlastního vidu (pro CPL.IT v. *Pokaždé, když přijde, zapne rádio*). Jiné způsoby slovesného děje počítáme do slovo tvorby a v PDT je zatím nezachycujeme. Nově tuto problematiku z jiného hlediska zpracovala Filip (1999); porovnání obou přístupů by bylo závažným přínosem pro bohemistiku.

### 2.10.2. Modalita

Hranice mezi analytickou modalitou a syntaktickým spojením není neprostupná, o čemž svědčí různé jevy týkající se zejm. slovesa *chtít*. I ve spojení s infinitivem, kde je protějškem *chtít* gramatém modality, jsou okrajové možnosti výjimek jako *včera ráno to chtěl*

*udělat večer, ale stihl to už v poledne*, nebo zápor u infinitivu (*To musí nevidět; může nepřijít*). Snad tu objektová platnost infinitivu (PAT) proniká pod tlakem morfologické formy (nejde-li jen o jednotlivé odchylky). Zatím není zpracováno ani spojení dvou modálních sloves (v. Benešová v Panevová ad., 1971, s. 127), ani jejich epistemické užití (*tady někde musí/může být*) a otázka jeho synonymie s částicemi jako *asi, určitě*.

### 2.10.3. Stupňování

Ačkoli gramatikalizace stupňování není úplná (nestupňují se všechna adjektiva a adverbia, valence jednotlivých stupňů se liší, vyjadřovacími prostředky jsou afixy), řadíme ho do tvarosloví. Elativ (*nejpřednější autoři*) ani absolutní užití komparativu (*starší pán*) zatím nespecifikujeme; jedním z možných řešení je vidět tu doplnění s hodnotou Gen a s nulovým vyjádřením; srov. i odd. 2.11.

### 2.10.4. Syntaktické gramatémy (u funktorů)

Je známo, že některé funktoři vyžadují jemnější sémantické členění. Ve FGP se s ním už dříve počítalo u funktorů lokálních a směrových (viz níže), i časových; TWHEN – *kdy* se člení na TWHEN.AFT – *kdy-před* (následnost závislého členu po řídícím), TWHEN.NIL – *kdy* (s bezpříznakovou současností), TWHEN.BEF – *kdy-po* (předčasnost), srov. Panevová ad. (1971, kap. 3). Dalším gramatémem u funktoři TWHEN je JBEF ('just before'), tj. bezprostřední předčasnost (*jakmile*). Nově sem řadíme i hodnoty APPX (*kolem poledne*) a INTV (*mezi pondělkem a pátkem*).

S b i n á r n í opozicí hodnot syntaktického gramatému jsme už dříve pracovali u určení doprovodu (ACMP, s předložkami *s* a *bez*) a benefaktivu (prospěchu, BEN, *pro* vs. *proti*). I účel (AIM) se štěpí na dvě hodnoty (*pomoci k čemu* vs. *od čeho, krém na opalování* vs. *prášky proti bolesti*), a také u zřetelového určení (REG) je dvojice

*se zřetelem/ohledem k/na* vs. *bez zřetele*; jemnější členění vyžadují další sémantické diference sekundárních zřetelových předložek.

Sémantické opozice u jiných funktorů mají nepříznakovou hodnotu NIL a v í c e hodnot příznakových, jako u srovnání (CPR, v. odd. 2. 11) a míry (EXT); ta se uvádí „přesně“ (EXT.NIL: *zaplatil hodně, zaplatil na halíř*), nebo „přibližně“ (EXT.APPX: *váží kolem 4 kg*), „nad uvedenou míru“ (EXT.MORE) a „pod ni“ (EXT.LESS). Také u funktorů ACT a PAT (popř. EFF, viz odd. 2.2) máme gramatémy pro sémantické distinkce nesené genitivem *p a r t i t i v n í m a n e g a t i v n í m* (ACT.GPART: *Ubylo vody*, ACT.GNEG: *Není peněz*, PAT.GNEG: *Na vesnici není vody*, PAT.GPART: *Dodej soli!* i ACT.GMULT: *Tam bylo lidí!*, ACT.DISTR: *Na každé větvi viselo po jablíčku*, ACT.APPX: *Na stovky odpůrců zákona přišlo*). Významový rozdíl přísudkového nominativu a instrumentálu u spony (označovaného jako druhý aktant, tj. PAT) je vyjádřen gramatémem NIL pro nominativ (*Dědeček byl švec*) a PNREL pro *r e l a č n í* instrumentál (*Pokladníkem je ted' Horák*).

U funktorů *l o k á l n í c h* a směrových se základní vztahy členění na řadu významů vyjadřovaných primárními i sekundárními předložkami. Rozdíl *položit věci na stůl* vs. *do stolu* se jeví jako rozdíl v syntaktickém gramatému funktoru DIR3, avšak rozdíl *šel do lesa* vs. *na pole* se chápe jako frazeologicky vázané vyjádření téhož funktoru i téhož gramatému DIR3-*do*; přesněji říkáme DIR3-*v*, protože tabulka v Manuálu, která převádí lokální a směrové předložky na gramatémy, zachycuje invariant pro daný význam u různých funktorů. Převedení na invariant je průhledné pro základní významy typických předložek (*na* – povrch, *v* – vnitřek, *nad*, *pod*, *před*, *za* a *mezi*, není už však jasné pro další odstíny lokace, nesené předložkami jako *uprostřed čeho*, *nedaleko od čeho*, *poblíž čeho*).

## 2. 11. Srovnání

I funktoru CPR ('Comparison') rozlišujeme aspoň tři syntaktické gramatémy: od hodnoty bezpříznakové (CPR.NIL, s výrazy *jako*, *v porovnání s čím*) rozlišujeme srovnání rozdílem (CPR.DFR), srov. i Šmilauer (1947, s. 292) – *Petr je větší než Jirka*.CPR.DFR (toto určení jistě má valenční povahu u komparativu), další gramatém je CPR.AGST ('Against'), *v protikladu s čím*. Téměř všechny srovnávací obraty jsou v podstatě „zkrácenou“ predikací, což vede k doplňování elidovaných uzlů. Za přímá srovnání (bez „zkrácení“) pokládáme pouze typ (72), (73), s gramatémem DFR pak (74):

(72) Bydleli v chaloupce jako dlaň.CPR.NIL

(73) Upekla moučnick jako báseň.CPR.NIL

(74) Pařez byl těžší než 10 kg.CPR.DFR

Ve větách jako (75), (76) už vzniká potřeba doplnit elipsu:

(75) Chlapec pozdravil profesora zdvořileji než kamarád (= než profesora.PAT.ELID pozdravil.CPR.DFR.ELID kamarád.ACT)

(76) Chlapec pozdravil profesora zdvořileji než kamaráda (= chlapec.ACT.ELID pozdravil.CPR.DFR.ELID kamaráda.PAT)

TGTS věty (77) vypadá shodně, ať je ve větě adverbium *stejně* na povrchu přítomno, nebo je vypuštěno:

(77) Udělal to (stejně) jako Karel (= stejně.MANN (to.PAT.ELID udělal.CPR.NIL.ELID Karel.ACT)).

Srovnávací obraty jsou stále ještě spojeny s mnoha nedořešenými otázkami, např.

(a) jaké sloveso rekonstruovat u některých elips, srov. (78)

(78) Pověsil kalendář níž než obrázek (než pověsil/visí/je obrázek)

(b) kolik uzlů zrekonstruovat podle predikace řídící, srov. (79):<sup>11</sup>

(79) V neděli to doma udělám rychleji než u vás (než to udělám u vás (v neděli?))

(c) jak zachytit specifické srovnávací obraty, kde jde o škálu od jednotlivých čistých frazémů (*přiběhla jako na koni* ≠ *jako by přiběhla na koni*, podobně *jde to jako po másle*, *chová dítě jako v bavlnce*) až po ustálenost, která není zcela lexikalizovaná, srov. (80), (81), (82), kde je kvantitativní údaj uveden výrazem *víc(e) než*.

(80) Byl tam více než tisícíhlavý zástup

(81) Drak byl více než sedmihlavý

(82) Mzda se víc než ztrojnásobila

V konstrukcích typu *větší město než Praha* je možné vidět elipsu velké části závislé klauze, tj. na TR by měly zápis stejný jako *město větší, než jak velké město je Praha*. Ale v konstrukci odpovídající spojení *město větší, než jak je velká Praha* nelze komparativ přehodit do povrchového postavení neodpovídajícího projektivitě (důležité je, že tady by místo slova *Praha* mohlo být i slovo, které neoznačuje žádné město, např. *než celé Česko*).

Ani konstrukce dalších typů se neanalyzují snadno:

(83) Na rodičovské sdružení chodí (spíš) maminky (spíš) než tatínkové.

(84) Na křižovatce se dejte spíš rovně než doprava.

(85) Položte to spíše (raději) pod postel (spíše) než na postel.

Pozice slova *spíš* (ať je od *než* odděleno, nebo ne) nemá vliv na sémantiku celého výrazu, avšak věta (85), v níž mají porovnávaná směrová doplnění odlišné syntaktické gramatémy, ukazuje, že je patrně třeba větu rekonstruovat jako: *Položte to spíše pod postel, než abyste to položili na postel*. Taková rekonstrukce se zdá těžkopádná, ale je nutná i u jiných typů srovnání; rozliší např. víceznačnost často diskutovaných vět jako (86), v. i Panevová (1980, s. 153):

(86) Zním lepšího právníka než dr. Novák.

(a) = než je.CPR.DFR dr. Novák.ACT (dobrý právník)

(b) = než je právník, kterého zná dr. Novák

## 2. 12. Některé specifické problémy při anotování

Při analýze souvislých textů na základě explicitně formulovaného pojmového rámce se setkáváme s problémy, které při popisu neformálním zpravidla neřeší. Zavedli jsme např. nový typ elipsy pro názvy artefaktů užitých bez svého rodového pojmu:

(87) Na stole leží Proti všem.

(88) Adéla ještě nevečeřela bylo shlédnuto jen několika tisíci diváků.

Do stromu se doplní uzel s lematem Idph pro *i d e n t i f i k a č n í* výraz a s funktorem ACT v (87) a PAT v (88); zastupuje rodové

jméno typu *román/kniha, film*. Sám název se rozloží do příslušného počtu uzlů, z nichž základní má funktor ID.

Nový funktor zavádíme pro řešení otázky spíše technické, totiž pro konstrukce s významem matematických operací (násobení, poměr, interval aj.); je to funktor OPER se syntaktickými gramatémy MULT (*pozemek o rozměrech 10 x 15 m*), RATIO (poměr, *zvítězili 5 : 3*), INTV (*léta 1989 – 1999*). OPER odpovídá především grafickému vyjádření z jiné znakové soustavy, než je slovní zásoba češtiny, ale také některým synonymním verbálním vyjádřením (např. *v období 1989 až 2000*).

K otevřeným otázkám patří vztah mezi pozicí INTV a složenými konstrukcemi *od.DIR1... (přes.DIR2...) do/po.DIR3...*, i obdobnými konstrukcemi časovými.

### **3. Vztahy mezi TR a morfématickým zápisem**

Jak už řečeno, zabývali jsme se i v odd. 2 vztahy mezi rovinou tektogramatickou (TR) a morfématickou, protože vymezení jednotek TR pro účely analýzy a anotování musí vycházet z morfématických vyjadřovacích prostředků (ke kterým počítáme i povrchový slovosled). Zatímco v odd. 2 šlo především o otázky spojené s jednotlivými hodnotami funktorů a gramatémů, v této části článku probereme hlavně některé obecnější aspekty, jako vztah mezi slovosledem hloubkovým a povrchovým (3.1), problémy spojené s elipsou (3.2) a s frazémou (3.3), i některé specifické otázky (3.4).

#### **3.1. Slovosled hloubkový a povrchový (morfématický)**

K základním principům slovosledu v češtině (podobně jako v řadě jiných jazyků) patří Ertlova a Mathesiova členská sounáležitost, jejímž přesněji zpracovaným protějškem ve formální gramatice je podmínka projektivity. Ta vyžaduje, aby závislostní syntagma nebylo přerušeno členy, které nejsou jeho řídicímu slovu podřízeny.



Bezpříznakový slovosled, vyhovující této podmínce (stejně jako její obdobě pro koordinaci a apozici), umožňuje vidět zápis věty v TGTS jako projektivní strom (na TR, pokud jde o kombinace závislostních vztahů s koordinačními a apozičními, jako vícerozměrnou síť), s možností jednoznačné linearizace. Povrchový slovosled patří k morfématickému zápisu, který nemá podobu stromu, ale jen řetězu těsněji nebo méně těsně spojených morfématických jednotek; tady už syntaktická závislost není přítomná, takže tu k přímému porušení projektivity nedochází. Rozdíl mezi úrovněmi slovosledu se netýká jen těch kontextů, ve kterých jde o projektivitu, v. např. postavení adjektiva před substantivem i tam, kde je adjektivum dynamičtější, a taky ovšem o postavení pomocných slov aj. Pořadí uzlů v TGTS odpovídá stupňům výpovědní dynamičnosti (v. odd. 2.5), tj., kde je pomocný strom ATS neprojektivní, je třeba postavení aspoň jednoho uzlu změnit. Často jde o první (nejlevější) člen ATS, který je v TGTS označen jako kontrastivní část základu a umístěn jako projektivně závislý zleva na svém řídicím uzlu, který může být částí ohniska (srov. už Sgall ad. 1986, zejm. s. 241); tak je tomu např. ve větě (89), které v TGTS odpovídá pořadí vyznačené v (89'):

(89) Karla plánujeme poslat do Francie.

(89') my.T plánujeme Cor.T Karla.C poslat do Francie.F

Příznakového slovosledu se týká i dlouho diskutovaná, nedořešená problematika odděleného přívlastku a hranic mezi ním a doplňkem.

### 3.2. Elipsa

Problematika elipsy (ke které je možné v širším smyslu počítat i nulové vyjádření všeobecného aktantu) je velmi široká. Pro slova v morfématické podobě věty nepřítomná doplňujeme v TGTS uzly především tam, kde jde o o b l i g a t o r n í aktant systémově nevypustitelný, elidovatelný jen ve specifických kontextech (jde o elipsu textovou, "aktuální"; není gramaticky dáno, jaký lexém je

elidován). Někdy se tu dá doplnit jen zájmeno (nebo ani to ne), někdy konkrétní slovo, např. v (90) by přicházelo v úvahu i anotování s členem Jirka.ELID:

(90) (Potkals Jirku?) - Potkal (v TGTS je *já*.ANIM.SG.ELID.ACT  
*on*.ANIM.SING.ELID.PAT.COREF:Jirka).

(91) (Potkals našeho Jirku?) - Potkal (*já* jako v (90), *on* má ELEX namísto ELID) Symbolem ELID vyznačujeme, že výraz je elidován (a v TGTS rekonstruován), pokud kontext ukazuje, že nejde o výraz rozvitý; symbol ELEX označuje rekonstrukci rozvitého výrazu, ale bez specifikace jeho rozvití. Není totiž možné v kontextech jako (92) určit, že *potkal* je tu za *potkal před domem*. Pokračování by mohlo být třeba i: *Jenže ne před domem, ale u řeky*.

(92) (Potkals Jirku před domem?) Potkal.

Textově elidované fakultativní členy se doplňují, jen pokud k tomu přímo vede koordinační ap. kontext. Při elipse slovesa zatím v TGTS doplňujeme Emp.PRED, jde-li o řídicí sloveso věty (nejen závislé klauze), např. *Nač* Emp.PRED *ten spěch* (na rozdíl od závislé klauze ve větě *Ptali se*, Emp.PAT *nač ten spěch* nebo od nevětného pojmenování jako *Samoobsluha*.DENOM; ani hranice mezi eliptickou větou a pojmenováním není předem jasná). Větu po spojce jako *a to*, *a sice*, chápeme jako aponovanou; např. *Milan přijel do Prahy včera, a to* (Emp.ELEX.COREF:přijet) *vlakem*. Symbol COREF (a další) dává východisko pro studium vztahů v d i s k u r z u, které bude mít velmi závažnou úlohu.

Při elipse aktantů u dějových substantiv je třeba pečlivě vybírat mezi odkazem k antecedentu a nulou vyjádřeným Gen:

(93) Po vstupu do místnosti Honza rozsvítil.

Podle kontextu se ke *vstup* doplní buď *on*.ACT (s rodem a číslem podle kontextu, např., předchází-li věta (94), s ANIM.PL a s *Honza* v COREF), nebo Gen.ACT, popř. Unsp.ACT.

(94) Přišli Milan a Honza.

I spojení s posesivním *z v r a t n ý m* zájmenem ukazují, že bude potřeba důkladně zvážit, kde jsou hranice elipsy:

(95) Brzo jsem viděla člověka s cedulí se svým jménem.

(96) Po atentátu na svého předchůdce se stal vrchním velitelem armády.

Pokud *svým* v (95) znamená *mým*, jde o obecně předpokládanou koreferenci s podmětem; při míň pravděpodobném čtení, kde je *svým* koreferenční s *člověka*, je třeba předpokládat elipsu slov ze závislé klauze  *který měl* (nebo kondenzaci této klauze do podoby *s+instr*). Elipsa v kondenzované závislé klauzi ve větě (96) jedním ze svých čtení (pro daný kontext nevhodným) budí dojem, že pachatelem atentátu byl ten, ke komu referuje podmět celé věty (v dané zprávě egyptský prezident Mubarak); slovo *atentát* tu funguje jako dějové jméno (významově odvozené od výrazu *spáchat atentát*); další příklady viz u Panevové (1986) a u Čmejrkové (1998).

U komplexů *z á v i s l ý c h* členů obecnými pravidly těžko

vymežitelných umožní symbol ELEX pozdější studium, v. např.:

(97) Potkal tvého mladšího bratra, ale mého [mladšího bratra] ne[potkal.ELEX].

Elidován je tu nejen chybějící obligatorní člen, ale i slova, která jsou v první části věty (popř. v otázce nebo v jiném odpovídajícím specifickém typu kontextu) přítomná jako syntakticky podřízená tomuto uzlu (např. *mladšího bratra* v (97)); k elidovaným členům ovšem nepatří *ty*, jejichž syntaktická pozice je (jinými lexémy) v druhé (elipsou "zkrácené") části věty obsazená (např. *tvého* v (97)). Často se doplňuje zájmenný uzel s lematem *on* (pro všechny rody a čísla, i pro posesivní *jeho, její* atd.), zejména tam, kde ze stavby dané věty (nebo z těsného kontextu, jako je zodpovídaná otázka v (90)) není jasné, co je antecedentem elipsy, viz (93).

Kdy je vhodné doplnit *ten*, ne *on*, např.: [*ty*] *kulaté dej doprava*, [*ty*] *hranaté doleva*, to musí být upřesněno na základě dalšího výzkumu.

Řídící slovo se nedoplňuje u slov jako *nemocný, raněný, ten(to), některý, všichni*, která mají i substantivní funkci. Seznamy těchto

slov uvedené v Manuálu bude třeba postupně doplňovat; podstatná je ovšem otázka kritérií výběru.

S řadou problémů se setkáváme u elipsy v *k o o r d i n a c i*. Při homonymii koordinace členské a větné dáváme přednost strukturám členským, pokud nejsou kontextem přímo vyloučeny. U elipsy závislých členů rozlišujeme mezi koordinací členskou a větnou i podle takových měřítek, jako je rozdíl mezi (98) a (99):

(98) Jirka potkal a pozdravil Marii.

(99) Jirka potkal Marii a pozdravil.

Zatímco v (98) vidíme koordinaci přísudků společně rozvitých jediným ACT a jediným PAT, v (99) se aspoň někdy dá podle kontextu vybrat, kdo byl pozdraven (v COREF bude *Marie*, Unsp nebo Gen). Ani v (100), kde nejde o elipsu, není totožnost s Marií jedinou možností, takže rozbíráme *ji* jako *on.FEM.SG.ELID.PAT*.

(100) Jirka potkal Marii a pozdravil ji.

### 3.3. Frazémy

Oblast frazémů patří k těm, kterým se dosud zdaleka nedostalo potřebné míry pozornosti, viz zejm. Čermák (1994) a Hnátková (2002). Aby PDT mohl co nejlíp sloužit k jejich dalšímu studiu, označujeme ty z nich, které byly při anotování zjištěny, jako PHR<sub>i</sub>, kde *i* je 1 u všech složek prvního frazému ve větě, popř 2 (u druhého frazému) atd. Konkrétní hodnoty funktorů zatím zaznamenáváme tam, kde je to trochu možné (aspoň "v přeneseném smyslu"). To jistě není teoreticky podepřené řešení, ale může to prospět při dalším zkoumání složení frazémů a jejich přesnějšího odlišení od jiných syntaktických spojení. Uvnitř PHR<sub>i</sub> však je možné mít slovo bez specifické hodnoty funktoru (což zapisujeme jako DPHR, 'dependent part of phraseme'), závislé na hlavní složce frazému, která funktor má, např. *křížem.MANN krážem.DPHR*.

Zdá se, že dnes v češtině rychle roste váha slovesných sousloví a podobných vazeb, jako mít zájem o, mít názor na, cítit nutnost čeho, být doma v čem. Aktuální je problematika sloves fázových (začít, přestat) a ovšem i hranice mezi slovesy čistě modálními (chápanými jako gramatémy) a jinými (jako plánovat, zamýšlet, dovolit si aj.), viz i odd. 2.10.2.

Otevřené jsou technické otázky zachycení frazémů ve slovníku, kde jistě musí být uvedena jejich povrchová podoba (včetně předložek a dalších pomocných slov). O frazémech ve srovnávacích konstrukcích jsme mluvili v odd. 2.11.

#### 3.4. Další otázky morfémů a jejich funkcí

Závažným a často diskutovaným problémem je, jak chápat funkci genitivu při *n u m e r a c i*. Spojení jako pět lidí, lidí přišlo pět (několik, mnoho ap.) jsou v TGTS strukturována stejně jako tři lidé, lidé přišli tři; protějškem genitivu je tu funktor ACT (popř. PAT, EFF aj.). Důležité měřítko se opírá o to, že dativ a instrumentál si tu své funkce zachovávají: dal to pěti (mnoha) lidem.ADDR, soutěží s deseti lidmi; tvary číslovky (nebo jiného kvantitativního výrazu) tu vystupují jako shodný přívlstek (ale např. lidí je pět rozbíráme jako sponu s ACT a PAT, stejně jako lidé jsou tři nebo lidé jsou zdraví).

Naproti tomu slova sto, tisíc, milión, řada, spousta, trocha/u apod. chápeme jako substantiva, na nichž závisí jiné substantivum s funktorem MAT. Jak už řečeno, považujeme za relevantní tvar v pozici dativu a instrumentálu (chybí to spoustě lidí, přišel s celým stem lidí), popř. i plurálové tvary: byly tam tisíce lidí. Pokud se vyskytnou spojení jako se sto lidmi, rozbírají se jako s pěti lidmi.

Další specifické otázky se týkají např.:

(a) příslovcí, u kterých není jasné, jestli rozvíjejí sloveso, nebo substantivum, jako vstupte dveřmi vlevo, mají vstup zdarma;

(b) vztahu lokace a přeneseného významu (existuje gramatická hranice mezi spojeními jako ustoupit od zdi a ustoupit od názoru?);  
(c) doplňování repertoáru funktorů: např. doplnit nový funktor "odplata" pro spojení jako za vaši ochotu vám půjčím auto; počítáme se dvěma významy slova jak: např. Slyšel jsem, jak přicházel chápeme jako synonymní s ...ho přicházet (jiné je ...že přichází) a vedle toho je jak jako způsobové příslovce: Podle toho, jak přicházel, poznali, že mu není dobře.

#### 4. Závěrečné úvahy

Viděli jsme, že z práce na syntaktických anotacích v PDT vyplývá řada nadějných, lákavých námětů pro bohemistický výzkum. Protože anotování postupuje od formy (výrazu, "označujícího") k funkci (významové nebo syntaktické, k "označovanému"), je třeba podrobně specifikovat primární a sekundární funkce výrazových prostředků, tj. zejména (i předložkových) pádů, spojek, infinitivu. Jejich specifikace vyžaduje formulaci a ověření operativních kritérií pro popis hranic mezi různými jevy ap. Cílem je získat východisko pro komplexní mluvnické zpracování češtiny, opřené jak o rozsáhlou empirickou základnu ze souvislých textů, tak o soustavný teoretický podklad.

Nemohli jsme zde tuto problematiku ani zdaleka vyčerpat. Naléhavě volá po principiálním zpracování, jak jsme se zmínili, výstavba p r o m l u v y, stejně jako rozšiřování s l o v n í k u a jeho obohacování v údajích valenčních rámců, repertoáru funkcí příslovcí apod., i všestranné studium f r a z é m ů. V syntaxi věty jde o přesné vymezení z á v i s l o s t i (v základě určené endocentricitou na úrovni slov nebo slovních tvarů, v. odd. 2.1.3), včetně jejího rozlišení od souřadnosti. Sem patří vymezení hranic mezi koordinací a subordinací včetně tzv. podřadného vyjádření koordinace (v. odd. 2.6.2), i např. tzv. obrácený poměr vět (klauzí). Chybí i specifikace funkcí i n t e r p u n k c e (včetně středníku), pomlček, uvozovek aj.

Další p e r s p e k t i v y výzkumu struktury češtiny nad materiálem z Českého národního korpusu v PDT vyžadují důsledněji navázat na morfématickou analýzu a (zdokonalené) značkování i na anotace na analytické rovině. Základní snahou bude i rozšíření automatických složek anotování, včetně soustavnějšího využívání statistických postupů. Pro ty je zatím v mnoha oblastech syntaxe příliš málo dat; bude třeba zpracovat mnohem rozsáhlejší korpusový (popř. i jiný) materiál, aby bylo umožněno jemnější třídění jevů.

Především je ale nutné osvojit si nový pohled na jazykový materiál a novou metodu jeho zpracování do podoby mluvnické soustavy: místo snahy po plně obecném základním východisku, do jehož rozměrů by bylo možné uvádět všechny jevy jazyka i mluvy, je třeba od začátku, už v základním pohledu (a tedy i v teoretickém pojmovém aparátu, ve formálním rámci popisu) počítat s tím, že jazyk má své j á d r o, své vrstvy bezpříznakové, základní, tvořící pravidelné vzorce, ale vedle toho oblasti příznakové, omezené kontextovými podmínkami a tedy více nebo méně specifické, až po jednotlivé výjimky; někdy se narušují i hodně obecné principy (v. zejm. o přímé řeči, odd. 2.2). Pojmový rámec by měl počítat s existencí rozsáhlých přechodných vrstev bez ostrých hranic jak mezi jednotlivými rovinami a oblastmi jazykového systému, tak mezi systémem a jevy patřícími do oblasti užívání, fungování jazyka nebo do oblastí kognitivních (ontologických, popř. pragmatických).

Jak zachytit zároveň pravidelnost jádra jazykového systému (které je pro dítě tak snadno osvojitelné) a všechny vrstvy a typy nepravidelnosti jeho rozsáhlých příznakových oblastí, to je právě hlavní otevřený úkol. K jeho řešení se můžeme přiblížit podrobným rozbohem nejrůznějších jevů, se kterými se teoreticky založená lingvistická koncepce setkává při snaze o její podrobné uplatnění na nejrůznější jevy přítomné v otevřených textech. K takovému podrobnému studiu poskytuje vhodné východisko a řadu aktuálních

námětů právě syntakticky anotovaný korpus.

### *Literatura*<sup>12</sup>

- ČERMÁK, Fr.: Idiomatics. In: *The Prague School of Structural and Functional Linguistics* (red. P.L.Lueldorff). Amsterdam-Philadelphia, J.Benjamins 1994, s. 185-195.
- ČMEJRKOVÁ, S.: Syntactic and discourse aspects of reflexivization in Czech: The case of the reflexive pronoun *svůj*. In: *Issues of valency and meaning. Studies in honour of Jarmila Panevová* (red. E. Hajičová), Praha, Karolinum 1998, s. 65-87.
- FILIP, H.: *Aspect, Eventuality Types and Nominal Reference*. New York – London, Garland Publishing 1999.
- Hajičová, E.: *Negace a presupozice ve významové stavbě věty*. Praha:Academia 1975.
- HAIČOVÁ, E.: Postavení rematizátorů v aktuálním členění věty, *SaS* 56, 1995, s. 241-251.
- HAIČOVÁ, E. Teorie optimality a aktuální členění věty. *SaS* 63, 2000, 161-169.
- HAIČOVÁ, E. – PANEVOVÁ, J. – SGALL, P.: Coreference in the grammar and in the Text. Part 1. *PBML* 44, 1985, 3-22. Part 2, *PBML* 46, 1986, 1-11. Part 3, *PBML* 48, 1987, 3-12.
- HAIČOVÁ, E. – SGALL, P.: Topic-focus and salience. In: *Proceedings of the 39th Annual Meeting of the association for Computational Linguistics and the 10th Conference of the European Chapter*, Toulouse:CNRS, 2001, s. 268-273.
- HLAVSA, Z.: *Denotace objektu a její prostředky v současné češtině*. Praha, Academia 1975.
- HNÁTKOVÁ, M.: Značkování frazémů a idiomů v Českém národním korpusu s pomocí Slovníku české frazeologie a idiomatiky. *SaS* 63, 2002, 117-126.
- KONEČNÁ, D.: K otázce druhů objektu podle významu. In: *AUC- Slavica Pragensia* 8, 1966, s. 311-316.
- PANEVOVÁ, J.: Vedlejší věty obsahové, *SaS* 32, 1971, 289-300.
- PANEVOVÁ, J.: K voprosu o reflexivnoj pronominalizacii v češskom jazyke. *Linguistische Arbeitsberichte* 54/56. Leipzig 1986, 44-56.



PANEVOVÁ, J.: Koreference gramatická nebo textová? In: *Etudes de linguistique romane et slave (A M. St. Karolák à l'occasion de son 60e anniversaire)* (red. W. Banyś, L. Bednarczuk, K. Bogacki), Cracovie 1991, s. 495-506.

PANEVOVÁ, J.: Česká reciproční zájmena a slovesná valence. *SaS* 60, 1999, s. 269-275.

PANEVOVÁ, J. – ŘEZNÍČKOVÁ, V. – UREŠOVÁ, Z.: The theory of control applied to the Prague Dependency Treebank (PDT). In: *Proceedings of the 6th International Workshop on Tree Adjoining Grammars and Related Frameworks. Università di Venezia*, 2002, s. 175-180.

ROOTH, M.: *Association with focus*. Amherst: University of Massachusetts 1985.

SGALL, P.: Functional sentence perspective in a generative description. *Prague Studies in Mathematical Linguistics*, 2, 1967, s. 203-225.

SGALL, P. – PFEIFFER, O. – DRESSLER, W. U. – PŮČEK, M.: Experimental research on Systemic Ordering. *Theoretical Linguistics* 21, 1995, s. 197-239.

STRAŇÁKOVÁ-LOPATKOVÁ, M. - ŽABOKRTSKÝ, Z.: *Valency Dictionary of Czech Verbs: Complex Tectogrammatical Annotation*. In: *Proceedings of the Third International Conference on Language Resources and Evaluation. Las Palmas, 2002*.

### **Poznámky:**

\* Základem tohoto článku, jehož první část vyšla v minulém čísle *SaS*, je výzkum v rámci projektu MŠMT ČR LN00A063.

7 Index u lematu identifikuje lexii, tj. příslušný význam polysémní jednotky; do kulatých závorek zapisujeme morfematickou formu příslušného aktantu, symbol *Kl* označuje vedlejší větu (klauzi, bez další specifikace, zda jde o klauzi indikativní, tázací nebo rozkazovací); aktant zapsaný v hranatých závorkách je fakultativní.  
8 Symbol *O<sub>PAT</sub>* vyznačuje, že fakultativní pozice *PAT* nebyla využita.

9 S tímto pojmem pracoval především J. Firbas; o rematizátorech v FGP viz Hajičová (1995); do středu pozornosti se ve světové literatuře dostává tento jev především v pracích formálních sémantiků počínaje Roothem (1985) pod názvy focus sensitive (focusing) particles, focalizers.

10 Interpretaci, že může jít o dohodu uvnitř členů ODS a uvnitř členů ČSSD, zde necháváme stranou; není ani jasné, zda tu jde spíš o vágnost než o víceznačnost, srov. i násobená čtení u Honza a Pavel se viděli (v představách) ve svých domech.

11 Relevantní tu je i postavení slov v AČ; v (79) z tematické pozice slova neděle plyne, že se srovnání jako celek primárně týká neděle.

12 Neopakujeme tu údaje uvedené u první části článku.

### **Resumé**

#### ***Towards a new level of work in the study of Czech: Working with an Annotated Corpus. Part 2.***

*Individual values of syntactic and morphological attributes used in the syntactic annotation of the Prague Dependency Treebank are discussed, as well as certain more general issues concerning the relationships between the underlying and the morphemic levels (word order, deletion and others).*

### **Příloha 2: Poznámky k manuálu tektogramatického značkování**

Veronika Řezníčková

V článku mnohokrát zaznělo, že hranice mezi jazykovými jevy často není jasná. V lingvistických úvahách nejasnost můžeme připustit, při konkrétním anotování však máme úkol sebesložitější větu rozebrat a musíme uplatnit jedno z možných řešení. Manuál (Hajičová ad. 2000) má anotátorům dát jednoznačné instrukce, jak postupovat. Pro

širší odbornou veřejnost se připravuje i nová verze Manuálu, psaná ne z pohledu anotátora, ale z pohledu uživatele PDT.

Vrátíme se zde alespoň k některým otevřeným otázkám probíraným v článku a podíváme se, jaké (prozatímní) řešení Manuál anotátorům doporučuje. Asi nejčastěji se anotátoři musejí vypořádat s nejistotou při výběru f u n k t o r u (srov. odd. 2). Manuál sice umožňuje vyjádřit nejistotu anotátora vyplněním druhého místa v atributu funktor, a to buď druhým funktorem (např. PAT/DIR1), nebo otazníkem (např. PAT/?), ale zároveň anotátory nabádá, aby této možnosti užívali co nejméně. K zajištění co největší konzistence v přiřazování funktorů slouží průběžně doplňované a aktualizované seznamy příslovcí, předložek, spojovacích výrazů a částic s výčtem možností jejich užití a s doporučeným funktorem a valenční slovníky (zejména valenční slovník sloves vallex, srov. Straňáková-Lopatková – Žabokrtský 2002, i vznikající valenční slovník substantiv).

Problémy při zachycování e l i p s, nastíněné v odd. 3.2, jsou dalším velmi častým zdrojem nejistoty v rozhodování anotátorů. Zvláště např. v koordinovaných konstrukcích s ř í d í c í m substantivem v plurálu je obtížné rozhodnout, zda doplnit druhý výskyt řídicího členu, nebo ne. Tak např. ze spojení *zástupci ministerstva školství a Akademie věd* je při znalosti českých institucí zřejmé, že se jedná o zástupce ministerstva školství a o *zástupce*. ELID Akademie věd (úkol rozhodnout o čísle u obou výskytů slova *zástupce* čeká až na anotátory vzorového souboru, protože ve velkém souboru se gramatémy čísla ještě nevyplňují). Často ale anotátoři ani podle znalostí všeobecných reálií nemohou adekvátně rozhodnout; např. v konstrukci *obchody s dámskou a pánskou konfekcí* je sice zřejmé, že se má doplnit substantivum *konfekce* (jde o dámskou konfekci a o pánskou konfekci), není ale jasné, jestli jde o obchody, kde se prodává obojí konfekce zároveň, nebo zda jde o různé druhy obchodů. Manuál doporučuje dát

v takovém případě přednost jednodušší struktuře, tedy substantivum *obchod* nedoplňovat. Je třeba hledat další (aspoň dílčí) kritéria.

U vět jako *Jirka potkal a pozdravil Marii* je třeba rozhodnout, jestli je dané rozvití společným rozvitím obou (všech) členů koordinace. Pokud ano, zavěsí se společné *r o z v i t í* na spojku, přičemž od skutečných členů koordinace (s nimiž je v tuto chvíli na stejné úrovni) se odlišuje hodnotou atributu 'Memberof'; zatímco slovesa *potkat* a *pozdravit* mají u tohoto atributu hodnotu CO ('Coordination'), substantiva *Jirka* a *Marie* tu mají hodnotu NIL (nejde o členy koordinace). Ovšem i zde existuje množství příkladů, kdy se můžeme pouze dohadovat, zda jde o společné rozvití koordinace. Tak např. u věty *Byl to hodný otec a dědeček* anotátor spíše rozhodne, že adjektivum *hodný* je společným rozvitím spojení *otec* a *dědeček*, ale u příkladů jako *velmi hodný a pečlivý chlapec* nevíme, zda je chlapec nejen velmi hodný, ale i velmi pečlivý, nebo pouze pečlivý a velmi hodný. V případě nejistoty tohoto typu Manuál doporučuje dané rozvití ponechat pouze jako rozvití toho členu koordinace, na kterém jednoznačně závisí.

Mohou se ale vyskytnout i případy, kdy je sice z kontextu zřejmé, že se jedná o společné rozvití koordinačního spojení, tento uzel by však potom zastupoval uzly s různou hodnotou některého z atributů, např. *func* (funktor) nebo *tfa* (aktuální členění). V tomto případě je potřeba doplnit nový uzel s odpovídajícím lematem a s odpovídajícími hodnotami problematických atributů jako rozvití druhého členu koordinace; např. v konstrukci *Jan a Karel se dříve často setkávali a informovali* doplní anotátoři podle instrukcí o zachycování reciprocit (srov. odd. 2.6.2) uzel s lematem *Rcp* (*se si tu navíc konkuruje jako součást lematu *setkávat\_se* a jako aktant slovesa *informovat**), tento uzel však zastupuje ve valenčním rámci slovesa *setkávat\_se* *Patiens* (*někdo.ACT se setkává s někým.PAT*), zatímco ve valenčním rámci slovesa *informovat* zastupuje *Adresát*

(někdo.ACT informuje někoho.ADDR o něčem.PAT). Zatímco tedy všechny ostatní členy této konstrukce (koordinace substantiv *Jan a Karel* i obě příslovce, *dříve* a *často*) jsou chápány jako společná rozvítko koordinace sloves, uzel s lematem Rcp rozvíjí každé sloveso zvlášť s patřičným funktorem.

Společné rozvítko koordinace může být i jiného charakteru, jako je tomu např. při koordinaci *m o d á l n í c h* konstrukcí (srov. odd. 2.10.2). V bezproblémových případech neodpovídá modálnímu slovesu uzel v TGTS, a přestože je v dané konstrukci modální sloveso povrchově vyjádřeno jen jednou, do atributu Deontmod obou plnovýznamových sloves se ve vzorovém souboru vloží příslušný gramatém označující modalitu (např. *chce odpočívat.VOL a poslouchat.VOL hudbu*). Složitější jsou ale např. spojení jako *chtěl prosperovat a aby se mu celkově dobře dařilo*; „povrchové“ *chtít* tu zastupuje dvě různá slovesa *chtít*: v první části koordinace mu odpovídá v atributu Deontmod hodnota VOL, ale v druhé části koordinace, kde nejde o čistě modální (pomocné) sloveso, je potřeba sloveso doplnit a závislou klauzi s *aby* na něj zavěsit jako PAT.

Než se dostatečně prozkoumá problematika konstrukcí, kde je plnovýznamové sloveso negované (např. *může nepřijít*, srov. odd. 2.10.2), doporučuje Manuál zachycovat modální slovesa jako plnovýznamová a slovesa v infinitivu na ně zavěsit jako jejich PAT (negace se pak zachycuje standardním způsobem, pomocí uzlu Neg.RHEM). Pokud jsou negovaná obě slovesa (např. *nemůže nepřijít*), zavěsí anotátoři uzel Neg.RHEM na každé z obou sloves. Ve vzorovém souboru je pak možné tuto konstrukci (jako synonymní se spojením *musí přijít*) zachytit bez negace a s gramatémem DEB.